Artificial Intelligence, Firm Growth, and Product Innovation*

Tania Babina† Anastassia Fedyk‡

Alex He§ James Hodson¶

November 2021

Abstract

We study the use and economic impact of AI technologies among U.S. firms. We propose a new measure of firm-level AI investments, using a unique combination of worker resume and job postings datasets. Our measure reveals a stark increase in AI investments across sectors. AI-investing firms see increased growth in sales, employment, and market valuations. This growth comes primarily through increased product innovation, reflected in trademarks, product patents, and product updates. AI-powered growth concentrates among ex-ante larger firms, leading to higher industry concentration and reinforcing winner-take-most dynamics. Our results highlight that new technologies can contribute to growth through product innovation.

Keywords: artificial intelligence, technological change, technology adoption, economic growth, product innovation, productivity, human capital, superstar firms, industry concentration

JEL codes: D22, E22, J23, J24, L11, O33

*This paper builds on an earlier version of the project titled "Artificial Intelligence, Firm Growth, and Indsutry Concentration", which had its first draft on March 16, 2020. The authors are grateful to Daron Acemoglu, Philippe Aghion, David Autor, Ann Bartel, Jan Bena, Francesco D'Acunto, Nicolas Crouzet, Xavier Giroud, Maarten de Ridder, Larry Katz, Anton Korinek, David Levine, Max Maksimovic, Gustavo Manso, Filippo Mezzanotti, Thomas Philippon, David Sraer, Tano Santos, Bledi Taska, Laura Veldkamp, Shang-Jin Wei, Michal Zator, and seminar participants at AFFECT, Babson College, Columbia University, Federal Reserve Board of Atlanta, FOM Conference, Labor and Finance Group Conference, Michigan State University, NBER Corporate Finance, NBER productivity seminar, NBER SI Macroeconomics and Productivity, NBER CRIW, Northwestern University, NYU WARPFIN, Society for Economic Dynamics, Tilburg University, Triangle Macro-Finance Workshop, Tulane University, Queen Mary University of London, Sao Paulo School of Economics, UC Berkeley, UNC Junior Roundtable, University of Illinois Chicago, University of Maryland, University of Oklahoma, University of Washington, and Vienna Graduate School of Finance for helpful suggestions. The authors thank Cognism Ltd. for providing the employment data; Burning Glass Technologies for providing the job postings data; Bernhard Ganglmair, W. Keith Robinson, and Michael Seeligson for sharing their data on patent (product vs. process) classification; and Gerard Hoberg and Gordon Phillips for providing data on changes in firms' product mix. Nikita Chuvakhin, Joanna Harris, Derek Ho, Gary Nguyen, and Hieu Nguyen provided excellent research assistance. Fedyk gratefully acknowledges financial support from the Clausen Center for International Business and Policy at UC Berkeley.

†Columbia University. Email: tania.babina@gsb.columbia.edu.

‡University of California, Berkeley. Email: fedyk@berkeley.edu.

§University of Maryland. Email: axhe@umd.edu.

¶Cognism; AI for Good Foundation. Email: hodson@ai4good.org.

Technological change is a key driver of economic growth (Romer, 1990; Aghion and Howitt, 1992). The past decade has seen a new technological shift: substantial developments in artificial intelligence (AI) technologies and their wide-spread commercial application, driven by rapid accumulation of data, falling costs of computing power, and methodological breakthroughs (Furman and Seamans 2019; Mihet and Philippon 2019). AI is a prediction technology, and predictions are at the heart of decision-making under uncertainty (Agrawal et al., 2019). AI algorithms allow firms to learn better and faster from vast quantities of data, significantly improving the accuracy of predictions. As such, AI can be a general purpose technology that generates growth through increased productivity and product innovation (Aghion et al., 2017; Cockburn et al., 2018). Indeed, in a survey of executives at companies investing in AI, 70% anticipate that AI will fundamentally transform their companies and industries within the next five years.[1] Yet it remains an open question whether artificial intelligence can transform economies and spur economic growth, as lackluster aggregate productivity growth over the past decade has led to concerns that the benefits of AI may be over-hyped or take a much longer time to materialize (Mihet and Philippon, 2019; Brynjolfsson et al., 2019; Haltiwanger, 2019). To date, the lack of comprehensive data on firms' use of AI technologies has posed the key challenge to understanding the prevalence and the economic impact of AI technologies (Seamans and Raj, 2018).

In this paper, we propose a new measure of investments in AI technologies based on firms' AI-skilled human capital. The heavy reliance of AI on human expertise makes the human-capital-based approach particularly well-suited in this setting. We take advantage of a unique combination of datasets that capture both the *stock* of and the *demand* for AI-skilled employees among U.S. firms: resume data from Cognism Inc, which offer job histories for 535 million individuals globally, and job postings data from Burning Glass, which capture 180 million job vacancies. Our new AI measure allows us to analyze the patterns of AI adoption and its impact on the adopting firms and industries. Our main takeaway is that firms that invest more in AI experience higher growth through increased product innovation, which can be seen in increased trademarks, product patents, and updates to firms' product portfolios (Hoberg and Phillips, 2016). At the aggregate level, growth in AI investments are associated with increased industry concentration and winner-take-most dynamics, as larger firms benefit more from AI investments. Overall, our results suggest that, so far, the first-order effect of AI has been in empowering growth through product innovation, consistent with AI reducing the costs of product development.

Our work offers several innovations over the existing literature. First, while prior research has made progress in studying the impact of AI on labor markets and occupations (e.g., Felten et al., 2019; Acemoglu et al., 2021), we focus on the impact of AI on firms and shed new light on the

[1] See here for a survey by Deloitte in 2018.

ability of AI to drive growth, including the mechanisms through which this growth is achieved. To do so, we measure granular AI investments and their potential effects for a broad sample of AI-*using* firms across a wide range of industries, which complements recent work that focuses on AI-*inventing* firms (Alderucci et al., 2020). Second, in the absence of administrative firm-worker matched U.S. data containing individual workers' occupations, our Cognism resume data provide high coverage of U.S. jobs with detailed job descriptions while representing more than 64% of full-time U.S. employment as of 2018.[2] Third, our paper is also the first to cross-validate AI labor demand identified from job postings with inflows and outflows of AI workers identified from resumes. Fourth, our rich data on firms' employees and their jobs allow us to measure and control for confounding factors, such as the use of non-AI information technologies, and capture the use of external AI solutions and software (e.g., IPSoft Amelia). Finally, our comprehensive cross-industry data provide a first glimpse at the aggregate industry-level implications of AI investments.

Even with our detailed data, identifying firms' investments in AI is challenging due to the multifaceted nature of AI applications.[3] We circumvent this challenge by proposing a new data-driven approach to identify AI-related jobs, which does not depend on pre-specified lists of keywords. Instead, our algorithm learns the AI-relatedness of each job posting empirically. First, we measure the AI-relatedness of each skill in the job postings data, based on that skill's co-occurrence with the core AI skills—machine learning, computer vision, and natural language processing. Second, we obtain a measure of AI-relatedness of each job posting by averaging the AI-relatedness of all skills required by the job posting. Finally, we leverage the most AI-related skills from the job postings data to classify AI workers in the less structured resume data. For each employee, we consider whether skills with the highest AI-relatedness (e.g., "deep learning") appear either in the job title, in the job description, or in any publications, patents, or awards received during that job. This gives us a classification of each employee of each firm at each point in time. We aggregate both job postings data and resume data to the firm level and match to public firms in the Compustat data. Encouragingly, the two measures of AI investments, although based on two independent datasets, are highly correlated and yield consistent results throughout the paper.

We confirm that our human-capital-based measures of AI investments display intuitive properties. First, we manually inspect large samples of AI-classified jobs and confirm that our classification picks up highly AI-skilled positions. Second, given that we do not use job titles to filter AI-skilled jobs, we validate our measure by confirming that the job postings with the highest AI-

[2]For comparison, while the U.S. Census Bureau's Longitudinal Employer-Household Dynamics (LEHD) program provides firm-worker matched data and worker wages, it does not include any information on workers' occupations or their jobs (Abowd et al., 2009; Haltiwanger et al., 2014). Moreover, a typical project using the LEHD data does not have access to all states due to administrative reasons: for example, Babina (2020) has access to about 40% of employment and Babina and Howell (2018)—60%.

[3]Within a single firm, e.g. Caterpillar Inc., AI can have use cases ranging from improving machinery via computer vision to offering a new product line of Internet of Things style analytics to machine operators.

relatedness measures skew heavily towards highly AI-specific job titles. Third, we confirm that our AI measure does not pick up general data-related skills, only those that are specifically associated with AI implementation. Fourth, we provide detailed case studies of specific applications of AI within several firms. Fifth, we confirm that AI-investing firms also increase research and development (R&D) expenditures, consistent with increased experimentation with applying the new AI technologies. Finally, we enrich our baseline measure by incorporating the use of external AI solutions and software and find similar results.

We begin our analysis by describing key patterns in AI investments. In both job postings and employee resume datasets, the fraction of AI jobs has increased dramatically over time, growing more than seven-fold from 2010 to 2018. The share of AI jobs is highest in the technology sector, but the rate of increase in AI investments over time is similar across sectors. At the firm level, growth in AI investments are more pronounced among ex ante larger firms and firms with higher cash holdings and R&D intensity. Looking at the local labor market conditions, we observe that higher-wage and more educated areas experience faster growth in AI-skilled hiring.

We next address the fundamental question of whether AI investments are associated with higher firm growth. As is standard in settings with slow-moving processes like technological change (e.g., Acemoglu and Restrepo, 2020), our primary specification is a long-differences regression of changes in firm outcomes from 2010 to 2018 on changes in the firm-level share of AI workers. This strategy is especially well-suited for our setting, where AI investments accumulate gradually over time and generate effects that may not be immediate. We include a rich set of controls: industry fixed effects and firm-, industry-, and commuting-zone-level characteristics in 2010. We document a strong and consistent pattern of higher growth among firms that invest more in AI: a one-standard-deviation increase in the resume-based measure of AI investments over the 8-year period corresponds to a 20.3% increase in sales, a 21.9% increase in employment, and a 22.4% increase in market valuation. The results are ubiquitous across major industry sectors (e.g., manufacturing, finance, and retail), supporting the idea that AI is a general purpose technology.

While the long-differences specification controls for time-invariant firm characteristics, we perform several tests to address concerns about omitted variables or reverse causality and buttress a causal interpretation of our results. First, we exploit firm-level panel data to examine firm growth dynamically in each year around AI investments using a standard distributed lead-lag model (Aghion et al., 2020). We find no differential trends in firm growth prior to AI investments and a positive effect after a lag of two to three years. Second, the results are robust to the inclusion of controls for past firm and industry growth and future growth opportunities proxied by Tobin's q. Third, we confirm that our results reflect specifically investments in AI, rather than other technologies: the effects of AI investments remain unchanged when controlling for contemporaneous

firm-level investments in robotics, non-AI information technologies, and non-AI data analytics.

We further address concerns regarding unobserved shocks driving both firm growth and AI investments using a novel instrumental variables strategy. We instrument for growth in firm-level AI investments using variation in firms' ex-ante exposure to the future supply of AI talent from universities that are historically strong in AI research. The core idea is that the scarcity of AI-trained labor is one of the most important constraints to firms' AI adoption (e.g., CorrelationOne, 2019), and universities that are historically strong in AI research have been able to train more AI-skilled graduates in recent years, enabling firms that typically hire from those universities to more readily recruit AI talent. To construct the instrument, we compile two new datasets on (i) the ex-ante strength of AI research in each university and (ii) firm-university hiring networks prior to 2010 to measure firms' exposure to AI-strong universities. Consistent with commercial interest in AI becoming widespread only since 2012, we show that firms' connections to AI-strong universities in 2010 were not driven by the need to hire AI-skilled workers. Moreover, we validate the key assumptions underlying this instrument and show that firms' ex-ante connections to AI-strong universities strongly predict ex-post hiring of AI-skilled workers. We then show that the instrumented firm-level growth in AI investments robustly predicts firm growth. Finally, we verify that the results are not driven by other characteristics of AI-strong universities such as strength in general computer science.

To understand the mechanisms through which AI generates firm growth, we examine two key non-mutually-exclusive channels: (i) product innovation and (ii) reduction in operating costs. The first channel is motivated by the extensive literature documenting the importance of product innovation for firm growth (Klette and Kortum, 2004; Hottman et al., 2016; Argente et al., 2021). Theoretically, AI can potentially reduce the costs of product innovation in two ways. First, since product development involves lengthy experimentation with uncertain benefits (Braguinsky et al., 2020), the ability of AI algorithms to quickly learn from large datasets can reduce the uncertainty of experimentation in product development and make the process of learning about promising projects more efficient. For example, at Moderna, AI algorithms have been leveraged in the development of the first COVID-19 vaccine in just 65 days, a process that would previously take years. Moreover, AI algorithms themselves can constitute improved products (e.g., AI-powered trading platforms). Second, AI can contribute to increased product scope by improving firms' ability to learn about customer preferences and tailor product offerings to customer tastes (Mihet and Philippon, 2019). Empirically, we find that firms with larger growth in AI investments see increased product innovation, reflected in more product patents (Ganglmair et al., 2021), trademarks (Hsu et al., 2021), and updates to product portfolios (Hoberg et al., 2014).

The second channel through which AI can stimulate growth is by lowering operating costs

and improving productivity, for example, by replacing human labor for some tasks (Agrawal et al., 2019) or by increasing operational efficiency through better forecasting and more efficient processes (Basu et al., 2001; Farboodi and Veldkamp, 2021). Empirically, we do not find support for this second channel. Growth in AI investments have zero effect on changes in sales per worker, total factor productivity, and process patents—which may reflect long lags in materializing the productivity benefits of AI (David, 1990; Brynjolfsson et al., 2019).

The benefits from AI investments are unevenly distributed across firms, consistent with the hypothesis that AI can increase inequality by favoring large firms with more data, which is a crucial input to AI implementation (Mihet and Philippon, 2019; Farboodi et al., 2019). We estimate the effect of AI investments within groups of firms by initial size and find that the positive relationship between AI investments and firm growth is much stronger among ex-ante larger firms. These results provide a new angle to the endogenous growth literature. For example, Akcigit and Kerr (2018) find that larger firms have higher costs of product innovation, which put constraints on the ability of large firms to scale. Our evidence shows that AI can help large firms overcome these previously documented barriers and scale up more easily.

Our final set of results speak to potential aggregate effects of AI on industry dynamics. We first test whether firm-level growth translates into industry-level growth. It is possible that the positive effects on AI-investing firms are offset or even dominated by negative spillovers to competitors within the industry, and previous work shows that the use of technology can be contractionary at the aggregate level if input use declines (Basu et al., 2006). Nevertheless, we find that industries that invest more in AI experience an overall increase in sales and employment within the sample of Compustat firms. Second, growth in AI investments are associated with increased industry concentration, consistent with our finding that AI favors ex-ante larger firms with more data. This suggests that AI investments can affect industry dynamics by reinforcing winner-take-most dynamics.

Overall, we document that AI leads to higher firm growth, and this growth mainly comes from firms' use of AI technologies for product innovation. This mechanism reflects the nature of AI as a prediction technology. Predictions are essential for firms' decision-making across all aspects of operations and particularly in product development, which requires experimentation and learning about promising projects and customer preferences (Braguinsky et al., 2020). The ability to perform better predictions with AI can create new business opportunities. In this context, our paper offers micro-level evidence and helps to unpack the black box of where "new projects" and investment opportunities come from: new technologies like AI, which allow firms to learn better and faster, can expand the investment opportunity frontier by decreasing firms' product development costs.

Related Literature

Our paper provides one of the first pieces of systematic evidence for the impact of artificial intelligence on firms and economic growth. While some recent work provides evidence on the use of AI technologies in specific industries such as finance (D'Acunto et al., 2019) and retail (Bajari et al., 2019),[4] our comprehensive data allow us to measure AI investments across a wide range of industries, offering a new measure of technology adoption based on firms' human capital (Hall and Jaffe 2018).[5] Recent theoretical literature argues that as a general purpose technology, AI has the potential to stimulate economic growth across a wide range of sectors (e.g., Aghion et al., 2017; Mihet and Philippon, 2019). Our empirical evidence supports this view and offers an additional insight: the mechanism through which AI fuels growth is by empowering product innovation. Product innovation has been considered a key mechanism for firm growth (e.g., Hottman et al., 2016; Argente et al., 2021). Our results suggest that AI can contribute to product innovation by both (i) overcoming supply-side constraints, such as costly experimentation in product development, and (ii) improving firms' ability to learn about consumer preferences. Our findings are consistent with Braguinsky et al. (2020), who argue that experimentation and new technologies are crucial for firm growth, and support the insight by Cockburn et al. (2018) that AI technologies can spur innovation by allowing for faster accumulation of knowledge. Our findings also complement Rock (2019), who shows that the launch of Google's TensorFlow expedited the gain in market valuations associated with firms' exposure to AI, with null effects on productivity.

Furthermore, our results speak to the literature on technology adoption, diffusion, and implications for growth (e.g., Romer, 1990; Aghion and Howitt, 1992; Parente and Prescott, 1994). Recent work by Crouzet et al. (2021) exploits demonetization in India to explore the role of complementarities for technology adoption, and Juhász et al. (2020) highlight that experimentation in applying new technologies can delay their effect on firms. Several previous technologies have been specifically associated with increased product innovation. For example, Basker and Simcoe (2021) document an increase in trademark activity following the introduction of the universal product codes. At the same time, previous waves of IT investment were associated with economically large productivity increases but mixed results on firm growth measures such as market share (e.g., Tambe et al., 2020), and diffusion patterns for these technologies tended to favor smaller firms (e.g., Hobijn and Jovanovic, 2001). By contrast, our evidence shows that AI technologies can stimulate firm growth through product innovation, with effects that are especially pronounced

[4] Another strand of literature measures workers' exposure to AI and its impact on labor market outcomes. See Felten et al.(2018; 2019), Webb (2020), Grennan and Michaely (2020), and Acemoglu et al. (2021), among others.

[5] Our firm-level measures of AI investments based on human capital are complementary to recent work that measures technology adoption using survey data (e.g., Brynjolfsson and McElheran 2016; Acemoglu et al. 2022). To foster further research on the economic impact of AI, all code to generate our AI investment measures and our firm-level AI data will be publicly available on the authors' websites.

in large firms. Given the unique ability of AI to process large amounts of data and the fact that large firms accumulate more data, AI appears to reduce the costs of product development that are especially high for large firms (Akcigit and Kerr, 2018), allowing these firms to scale more easily.

Methodologically, our paper offers a new approach to measure firms' intangible capital based on human capital, with a specific application to capturing investments in AI. Despite ongoing efforts to incorporate more comprehensive measures of intangibles in the U.S. at the national level (Corrado et al. 2016), most firm-level measures of intangible capital use cost items such as R&D and SG&A (e.g., Eisfeldt and Papanikolaou, 2013; Peters and Taylor, 2017; Crouzet and Eberly, 2019; Eisfeldt et al., 2020). Finer firm-level data on intangible investments, which are available for some countries (e.g., New Zealand; see Chappell and Jaffe 2018), are generally not available for U.S. firms. Our methodology offers a new measure of intangibles that is consistent across firms and sectors and can be applied to measure various forms of intangible assets, especially those based on human expertise. For example, while our focus is on AI investments, we are also able to measure firm investments in robotics, non-AI information technology, and non-AI data analytics. More broadly, our AI measure contributes to the growing literature that uses textual analysis to construct measures of intangibles such as human capital and innovation. For example, Hoberg and Phillips (2016) analyze text of 10-K filings to create measures of firms' product markets, Fedyk and Hodson (2019) use textual analysis to measure firms' focus on technical skills, Kogan et al. (2019) construct occupation-specific indicators of technological change using patent text, Argente et al. (2020) employ textual analysis to map patents to products, and Babina et al. (2020) uses patent text to build a new comprehensive measure of technological entrepreneurship.

Finally, we contribute to the active debate on the causes and consequences of rising industry concentration (e.g., Gutiérrez and Philippon, 2017; Syverson, 2019; Covarrubias et al., 2019; Autor et al., 2020). One proposed channel is that intangible assets propel growth of the largest firms, contributing to increased industry concentration (e.g., Crouzet and Eberly, 2019). Our results support this hypothesis and suggest that technologies like AI can contribute to increased concentration by enabling large firms to grow even larger through increased product innovation, reinforcing winner-take-most dynamics.

1 Artificial Intelligence: Background and Mechanisms

According to the Organisation for Economic Co-operation and Development (2019), an AI system is defined as a *"Machine-based system that can, for a given set of human-defined objectives, make predictions, recommendations or decisions influencing real or virtual environments."* We provide a brief overview of the current commercial use and key features of AI, followed by a discussion on eco-

nomic mechanisms through which AI investments might benefit a broad range of firms.

1.1 Artificial Intelligence: A Brief Overview

Commercial applications and investments in AI have increased exponentially over the past decade. While there are no systematic data on AI investments by firms, recent estimates hover around $140 billion globally per year with estimated growth of nearly 100% over the next three years.[6] Historically, the U.S. is considered to be the leader in both academic research and private investments in AI, but other regions including China and the E.U. have recently been spearheading their investments (Knight, 2017). There has also been an expansion of AI investments across industry sectors. While the tech sector was an early adopter of AI, surveys of executives indicate widespread adoption of AI technologies by firms in all industries (see here for a survey by McKinsey).

Academic research in AI has flourished for decades since John McCarthy coined the term in 1955 (McCarthy et al., 1955).[7] The recent explosion of commercial interest in AI in the private sector is driven by supply-side factors: rapid accumulation of data, decreasing costs of computation, and advances in methodologies, including deep learning (Hodson, 2016). In terms of commercial applications, three key areas of artificial intelligence have captured the bulk of private sector investments: machine learning, natural language processing, and computer vision (see here for a survey by Deloitte in 2018).[8] These core techniques are united by their ability to perform high-skilled, non-routine tasks, such as prediction, detection, and classification (Agrawal et al., 2019). Their main distinction from traditional methods of data analysis consists of these techniques' ability to learn from vast quantities of high-dimensional data (including text, speech, and image data; Hauptmann et al., 2015) and significantly improve the accuracy of predictions. For example, the ImageNet challenge in 2012 led to an almost halving of image recognition error rates (relative to traditional methods), which launched large corporate interest in the computer vision space.[9]

AI has several key economic properties. First, AI is a prediction technology, and predictions are at the heart of decision-making under uncertainty—faced by firms in all aspects of their operations. As a result, the ability to perform better predictions with AI can create new business opportunities. Second, economists have argued that AI may be a general purpose technology (GPT): AI can be leveraged across different business segments and sectors to solve a wide range of business problems. Well-known examples of GPT include the steam engine, electricity, and internal combustion. Third, investments in AI center around human expertise, with complementary

[6] See the PitchBook AI & ML Emerging Tech Report 2021 here.

[7] A brief history of AI research can be found here.

[8] While our focus is on artificial intelligence technologies rather than specific automation technologies like ATMs and industrial robots, our measure does incorporate relevant recent robotics technologies (e.g., autonomous vehicles, vision-guided robots) that are highly related to computer vision and machine learning technologies.

[9] See ImageNet Large Scale Visual Recognition Challenge 2012 (ILSVRC2012)

investments in computing technology and data infrastructure. This differs from technologies that require mainly capital investments, such as industrial robots. As such, AI is an intangible asset, reflecting the broader shift towards intangible capital (Mihet and Philippon, 2019). The fourth key feature of AI technologies is that they are information goods with non-rival uses: new algorithms are usually published openly and can be used simultaneously by many firms. However, the extent to which AI can benefit firms depends to a large extent on who owns big data—the key input to AI technologies (Fedyk, 2016; Jones and Tonetti, 2020).

1.2 Artificial Intelligence and Firm Growth: Mechanisms

It is an open question whether and how investments in AI technologies benefit firms. On the one hand, as a potential general purpose technology, AI might spur economic growth. On the other hand, current attention to AI may be over-hyped (Mihet and Philippon, 2019), or AI may still be too early in the adoption cycle to have a meaningful impact on firm growth (Brynjolfsson et al., 2020). Below, we discuss two non-mutually-exclusive channels through which investments in AI can affect firm growth: (i) by increasing product innovation, e.g. through new product creation, improved product quality, and product customization, and (ii) by lowering operating costs, e.g. by replacing labor.

AI as a Driver of Product Innovation. An important mechanism for firm growth is through product innovation and the expansion of product varieties. For example, Hottman et al. (2016) quantify that practically all variation in firm growth in sales can be explained by variation in firms' product appeal to customers (due to either the quality of firms' products or customer tastes) and product scope (the number and variety of products). Importantly, Braguinsky et al. (2020) point out that product variety and product appeal, which are commonly treated as primitives, are actually determined endogenously through experimentation by firms, and other work also stresses the importance of knowledge accumulated through exploration and experimentation for product innovation (Klette and Kortum, 2004; Bustamante et al., 2020).

As a prediction technology, AI can potentially affect both aspects of product innovation highlighted by the prior literature: (i) innovations that overcome supply-side frictions to create new products or products of superior quality, and (ii) demand-driven innovations that increase product scope by better tailoring products to customer tastes. According to surveys of executives, enhancement of existing products and services and the creation of new ones is the top use of AI to date (see here for a survey by Deloitte).

The potential of AI to help create new or better products comes from the insight that product innovation is costly: it requires (often lengthy) experimentation, and its resulting benefits are uncertain, creating supply-side barriers to product creation and improvement (Akcigit and Kerr,

2018; Braguinsky et al., 2020). The ability of AI algorithms to quickly analyze large datasets and learn about the underlying relationships in the input data can potentially lower these barriers by reducing the uncertainty of experimentation and making the learning process more efficient, which leads to more product innovation (Cockburn et al., 2018). In practice, recent years show a number of ways in which AI has enabled or sped up the product innovation process. An important application is drug development, where AI can shorten the drug development life cycle. For example, at Moderna, AI algorithms have been leveraged to design and optimize mRNA constructs, contributing to the development and the production of the first dose of the COVID-19 vaccine in just 65 days, a process that would previously take years.[10] In addition, AI algorithms can help innovate on the quality of existing products and services by building AI models directly into products. For example, in Online Appendix A.1, we offer detailed case studies of the applications of AI, which show examples of AI empowering the introduction of the AI-driven trading platform DeepX at JPMorgan (which allows for faster and cheaper execution of trades) and "smart" machinery at Caterpillar (which improves machine safety and flexibility).

At the same time, AI can also contribute to increased product scope by helping firms learn about customer preferences more efficiently and, therefore, better tailor product and service offerings to customers' tastes and needs. An extensive literature in trade economics highlights the importance of information regarding target market conditions—especially firms' product appeal to the target customers—for firms' decisions to enter new export markets (e.g., Dickstein and Morales 2018; Berman et al. 2019). Analogous hurdles apply to firms' decisions to launch new products or expand their product variety in domestic markets: there is uncertainty regarding what customers want and how customer preferences might change. Using AI to analyze customer data can potentially enable firms to overcome this hurdle, providing "the right product on a hyper-individualized basis" (Hodson, 2016) and overcoming frictions in firms' demand accumulation processes (Foster et al., 2016; Argente et al., 2021). For example, data on individual behaviors, such as web browsing and location history and other digital footprints, can enable better approximations of parameters entering individual demand functions than pure demographic information, leading to more heterogeneity in products tailored to customers with different tastes (Mihet and Philippon, 2019). This application of AI mirrors recent theories of big data as generating more precise forecasts regarding which product lines will yield the highest value (Bustamante et al., 2020; Farboodi and Veldkamp, 2021): by better learning about consumer preferences for different products, AI can help direct firms' efforts towards the development of the most promising

[10]See here, where Dave Johnson, Moderna's VP of Informatics, Data Science, and AI, explains how Moderna was able to develop a COVID vaccine so quickly: "We very purposely designed all this infrastructure that we think of as an AI factory, in order to rapidly deliver algorithms from concept to production, to enable our scientists to leverage the power of AI in their daily jobs. [...] That allows our scientists to design novel mRNA constructs, use AI algorithms to optimize them, and then order them from our high throughput preclinical scale production line."

products.

Overall, AI can theoretically reduce the costs of product innovation and provide a mechanism for the key drivers of firm growth highlighted by Hottman et al. (2016). AI can potentially help firms increase both (i) their product appeal via improvements in product quality and the creation of new products; and (ii) their product scope via improved inference of customer preferences and the customization of products to those preferences. Empirically, if AI enables product innovation, we should observe that AI investments are associated with increased product creation and variety.

AI as a Driver of Lower Operating Costs. Technological innovations also often aim at lowering costs of existing operations and improving productivity (e.g., Basu et al. 2001; Cardona et al. 2013; Acemoglu et al. 2020). When it comes to AI, the technology can lower costs and increase productivity in at least two ways. First, AI can potentially replace human labor for some tasks (Agrawal et al., 2019), cutting per-unit labor costs. Specifically, the ability of AI to aid in the decision-making process and in solving complex cognition problems has led to concerns that AI can disrupt many high-skill and high-wage occupations, in contrast to previous waves of technology adoption (Webb, 2020). Second, AI can increase operational efficiency through better forecasting (Mihet and Philippon, 2019). For example, Bajari et al. (2019) examine the impact of big data in the context of Amazon's retail forecasting system. Theoretically, this aspect is explored by Tanaka et al. (2019), who present a model of firm input choice under uncertainty and costly adjustment, where forecast errors result in under- or over-investment.

The potential to use AI-based forecasting for streamlining firms' existing operations can also be seen in our data. The case studies in Online Appendix A.1 highlight how AI-enabled forecasting improves firm operations across a variety of industries: for example, AI workers at JPMorgan Chase model default of non-performing loans; Caterpillar leverages AI for inventory management; and UnitedHealth uses AI to support efficient medical billing. Empirically, these potential improvements can manifest through lower operating costs or higher firm-level productivity.

2 Data

We propose a new measure of firms' investments in AI based on their intensity of AI-skilled hiring. Relative to the prior literature, which has a dearth of firm-level data on AI investments, we provide a uniquely comprehensive perspective on firm-level AI investments by simultaneously measuring firms' *demand* for AI workers through job postings and the *stock* of AI workers through employment profiles. We detail each dataset and describe our sample construction.

2.1 Job Postings from Burning Glass

The first dataset we use covers over 180 million job postings in the United States in 2007 and 2010–2018. The dataset is provided by Burning Glass Technologies (BG in short) and draws from a rich set of sources. BG examines more than 40,000 online job boards and company websites, aggregates the job postings data, parses them into a systematic, machine-readable form, and creates labor market analytic products. The company employs a sophisticated deduplication algorithm to avoid double counting vacancies that post on multiple job boards. BG data contain detailed information for each job posting, including job title, job location, occupation, and employer name. Importantly, the job postings are tagged with thousands of specific skills standardized from the open text in each job opening. The main advantages of the BG dataset are the breadth of its coverage and the rich detail of the individual job postings. The dataset captures the near-universe of jobs posted online and covers approximately 60–70% of all vacancies posted in the U.S., either online or offline. Hershbein and Kahn (2018) provide a detailed description of the BG data and show that their representativeness is stable over time at the occupation level.

We focus on jobs with non-missing employer names and at least one required skill. About 65% of job postings have employer information and 93% of job postings require at least one skill.[11] We also drop job postings that are internships. We then match the employer firms in the remaining job postings to Compustat firms. This step is necessary to aggregate job postings to the firm level and merge with other firm-level variables. We perform a fuzzy matching between firm names in BG and Compustat after stripping out common endings such as "Inc" and "L.P.". For observations that do not match exactly on firm name, we manually assess the top ten potential fuzzy matches by looking at the firm name, industry, and location. Out of 112 million job postings with non-missing employer names and skills, 42 million (38%) are matched to Compustat firms. This slightly overrepresents employees of publicly listed firms, which constitute just over one fourth of U.S. employment in the non-farm business sector (Davis et al., 2006).

2.2 Employment Profiles from Cognism

We complement job postings with employee resumes, which allow us to measure the actual *stock* of AI workers at each firm and help address potential concerns around job postings: for example, if a firm is not able to hire despite active job postings, if a firm posts numerous job openings for one job, or if a firm on-boards AI talent through acquisitions rather than through direct hiring. We leverage a novel dataset of approximately 535 million individual profiles provided by Cognism, an aggregator of employment profiles for lead generation and client relationship management ser-

[11] Job postings with missing employer names are primarily those listed on recruiting websites that mask the employers' identities.

vices. Cognism obtains the resumes from a variety of sources, including publicly available online profiles, collaborations with recruiting agencies, third party resume aggregators, human resources databases of partner organizations, and direct user contributed data.[12] These data are introduced and described in detail in Fedyk and Hodson (2019). While the data slightly over-represent high-skilled employees, they cover approximately 64% of the entire U.S. workforce as of 2018 and offer a representative breakdown across industries. For each employment record listed by the individual, we see the start and end dates, the job title, the company name, and the job description. Individuals may also list their patents, awards, and publications. Cognism's AI Research department leverages techniques from machine learning and natural language processing, including named entity disambiguation and graph-based modeling methods, to further enrich the resume data by normalizing job titles and occupations, associating employees with functional divisions and teams within each firm, and identifying institutions, degrees, and majors from education records.[13]

We match employer names in the Cognism data to the names of publicly traded firms using a similar approach to matching employers in BG data to Compustat firms. Fedyk and Hodson (2019) provide further details on the procedure as applied to the resume data. The matching of individual resumes to firm entities is performed dynamically to account for acquisitions and divestitures. Of the 657 million US-based person-firm-year employment records between 2007 and 2018, 120 million (18%) are matched to U.S. public firms. This is consistent with approximately 26% of overall U.S. employment being accounted for by publicly listed firms (Davis et al., 2006). The sample of 120 million person-firm-years matched to U.S. public firms is comprised of 19 million distinct individual employees.

2.3 Additional Data Sources

We merge the Burning Glass job postings data and the Cognism resume data to several additional data sources. We collect commuting-zone-level wage and education data from the Census American Community Surveys (ACS), industry-level wages and employment data from the Census Quarterly Workforce Indicators (QWI), and academic publications from the Open Academic Graph (described in detail in Appendix A). Firm-level operational variables (e.g., sales, employment, market value) come from Compustat.

[12] The processing of all profiles is compliant with the applicable GDPR and CCPA regulations.

[13] The data snapshot is from July of 2021. Following Tambe et al. (2020), we only use the years through 2018, because the lag in workers updating their resumes could otherwise add significant noise to our measures.

3 Methodology and Descriptive Evidence

3.1 AI Investments from Job Postings (Burning Glass)

We take advantage of the detailed information on required skills in the job postings data to propose a new data-driven methodology for identifying AI-related jobs. Other work classifies job postings based on the presence of key terms from a pre-specified list,[14] which is likely to suffer from both Type I (incorrectly labeling tangentially-related employees as AI-related) and Type II (missing real AI skills that did not make the initial dictionary) errors due to the arbitrariness of the list of keywords. This is especially relevant in a quickly-evolving domain such as AI, where new emerging skills can easily be missed. Our methodology circumvents these challenges by learning the AI-relatedness of each of approximately 15,000 unique skills directly from the job postings data, based on their empirical co-occurrence (within required lists of skills across job postings) with unambiguous core AI skills. We then aggregate the skill-level measure to the job level by generating a continuous measure of AI-relatedness for each job posting, from which we can classify employees into AI-skilled workers and non-AI-skilled workers.

To measure the AI-relatedness of each skill, we calculate the skill's co-occurrence with Artificial Intelligence (AI) and its three main sub-fields: machine learning (ML), natural language processing (NLP), and computer vision (CV):

$$w_s^{AI} = \frac{\text{\# of jobs requiring skill } s \text{ and (ML, NLP, CV or AI in required skills or in job title)}}{\text{\# of jobs requiring skill } s}$$

Intuitively, this measure captures how correlated each skill s is with the core AI skills. For example, the skill "Tensorflow" has a value of 0.9, which means that 90% of job postings with Tensorflow as a required skill also require one of the core AI skills or contain one of the core AI skills in the job title. Hence, a "Tensorflow" requirement in a job posting is highly indicative of that job being AI-related. On the other hand, the AI-relatedness measure of the skill "Microsoft Office" is only 0.003. We list the skills with the highest AI-relatedness measures in Online Appendix Table A1.

We define the job-level AI-relatedness measure ω_j^{AI} for a given job posting j as the mean skill-level measure w_s^{AI} across all skills required by job posting j. We transform the continuous AI measure into a binary indicator by defining each job posting j as AI-related if the measure ω_j^{AI} is above 0.1, a threshold that captures the full range of AI-related technical jobs while minimizing false positives based on manual inspection of the data. The firm-level measure $Share_{f,t}^{AI}$ is then defined as the fraction of job postings by firm f in year t that are AI-related (i.e. $\omega_j^{AI} > 0.1$).[15]

[14] For example, Hershbein and Kahn (2018) classify jobs as requiring cognitive abilities if any listed skills include at least one of the following terms: "research," "analy-," "decision," "solving," "math," "statistic," or "thinking." Similar bag-of-words approaches with pre-specified search terms are used to identify AI-related employees (e.g., Alekseeva et al., 2020; Acemoglu et al., 2021).

[15] Throughout our empirical analyses, we focus on jobs that are matched to Compustat firms. Online Appendix

We use a discrete classification for ease of interpretability and consistency with the resume-based measure in Section 3.2, but we show in Section 4.1 that the results are robust to: (i) alternative cut-offs (e.g., 0.05 and 0.15), and (ii) using the continuous measure ω_j^{AI} aggregated to the firm level.

Online Appendix Table A2 provides examples of AI and non-AI job postings. For each job, the continuous AI measure is the average AI-relatedness of all required skills. Our measure enables us to capture a wide range of AI-related jobs, from data scientists to speech recognition scientists to autonomous vehicle engineers. While many AI-related jobs are data scientists and similar data-analysis-related jobs, our measure differentiates data-analysis jobs specifically related to AI (job postings numbered 6-10) from data-analysis jobs that are not specific to AI and that focus on more traditional statistical methods (job postings numbered 11-15). In addition, we further ensure that our measure is not picking up general programming or statistics skills not specific to AI by showing (in Section 4.1) the robustness of our results to manually refining our measure. In particular, we screen out skills that represent general programming languages (e.g., Python) or statistics (e.g., linear regression) and only keep skills that relate specifically to AI, including AI methodology or algorithms (e.g., supervised learning) and AI software (e.g., Tensorflow). This process, curated by the AI-trained personnel at the AI for Good Foundation, categorizes the 700 skills that have an AI-relatedness measure above 0.05 and are required in at least 50 job postings into "narrow" and "broad" AI skills. This refinement mainly leaves out skills with relatively lower AI-relatedness measures and empirically has little effect on the results.[16]

3.2 AI Investments from Resumes (Cognism)

In the Cognism resume data, we identify AI-related employees as those whose job positions directly involve AI. We begin with the set of 67 keywords in Online Appendix Table A1, which have the highest skill-level AI-relatedness measures. We then search for these terms in every employment record of each individual in the resume data to see whether: (i) that job (role and description) directly includes any of the identified AI terms; (ii) any patents obtained during the year of interest or the two following years (to account for the time lag between the work and the patent grant) include these AI terms; (iii) any publications during the year of interest or the following year include the AI terms; and (iv) any of the identified AI terms appear in awards received during the year of interest or the following year. If any of these conditions are met, then that person at that firm in that year is classified as an AI-related employee. For example, jobs with titles such as

Figure A1 plots the share of all job postings and the share of AI-related job postings that are matched to Compustat in each year. Although publicly listed firms constitute 38% of all job postings, they account for approximately half of all AI-related job postings. This suggests that, on average, publicly-listed firms hire more AI workers than private firms.

[16]For example, among the 50 skills with the highest AI-relatedness measures, 49 are classified as narrow AI skills (the single exception is "statsmodels," a Python package for general statistical analysis).

"senior **machine learning** developer" or job descriptions such as "develop chatbots using Python with **Tensorflow** and **deep learning** models" are identified as AI jobs.

After classifying each individual in each year, we use the number of AI-related employees and the number of total employees at each firm in each year to compute the percentage of employees of that firm in that year who are classified as AI-related. Given that our empirical analyses focus on U.S.-listed firms, our firm-level measure focuses on the employees who are based in the U.S.

3.3 Summary Statistics and Validation

We examine both of our constructed measures of AI investments, confirm that they display intuitive properties, and discuss how our resume data help address potential limitations of measuring AI investments through job postings. Validating our novel measure is challenging, given the lack of existing firm-level measures of AI investments. However, we show that our measure displays a number of intuitive properties, captures specifically AI investments, and does not suffer from biases such as firms investing in AI by acquiring AI startups.

First, we document that both measures—based on job postings and resumes—display a natural rise over time, increasing more than seven-fold from 2010 to 2018. Panel (a) of Figure 1 shows that the fraction of AI-skilled job postings starts out at 0.1% in 2010, rises monotonically over time (with the increase speeding up from 2014 to 2018), and peaks at 0.8% in 2018. Panel (b) shows analogous patterns in the resume data. The fraction of all employees who are classified as AI-related starts at 0.04% in 2007 and reaches 0.29% in 2018. There is substantial heterogeneity in the growth in AI-skilled labor across individual firms, which provides the variation needed to examine the relationship between AI investments and firm outcomes. For the entire sample of public firms, while a median firm sees an increase of 0% (0%) in the resume-based (job-postings-based) measure, this increase is 0.35% (1.33%) at the 90th percentile, 0.62% (2.99%) at the 95th, and 2.22% (8.11%) at the 99th percentile.

It is helpful to put into perspective the incidence of AI-skilled workers among U.S. employees. While AI workers constitute a relatively small fraction of total employment, skyrocketing demand for AI skills and correspondingly high salaries that they command—on the order of millions of dollars for prominent AI-researchers (Gofman and Jin, 2020)—suggest that AI-skilled workers are similar to other specialized, high-skilled, high-wage jobs. For example, in terms of the technological and innovative nature of their work, AI-skilled workers could be compared to inventors. Inventors also tend to be highly paid and represent around 0.13–0.24% of the U.S. workforce, which is similar in prevalence to AI workers.[17] Overall, while AI workers form a small fraction of the

[17] These estimates come from the USPTO patent data (Babina et al., 2020), where 0.13% is the share of U.S. workers who file patents in a given year, and 0.24% is the share of U.S. workers who file patents over a three-year period.

overall workforce, it is helpful to contextualize their impact against that of executives (Bertrand and Schoar, 2003) and patent inventors (Kline et al., 2019), both of whom are similarly small, high-skilled groups of employees that can nonetheless disproportionately affect firm outcomes.

Second, we document that the increase in AI jobs displays an intuitive distribution across industries. Panel (a) of Figure 2 plots the average share of AI-related jobs in the job postings data for public firms in each of the 2-digit NAICS sectors, separately for the years 2007–2014 and 2015–2018. Panel (b) repeats the same analysis for the share of AI-related employees in the resume data. The figure highlights that the share of AI job postings (resumes) is highest in the "Information" sector, growing from 0.57% (0.15%) in the early years of 2007–2014 to 1.68% (0.50%) in the later period of 2015–2018. However, almost all sectors see a meaningful increase in both measures, supporting the notion that AI is a general purpose technology (Goldfarb et al., 2019). The ability of our measures to pick up AI investments in a broad cross-section of economic sectors highlights a key advantage of our human-capital-based approach.

Third, intuitively, AI investments correlate positively with increased R&D expenditures. For example, changes in the resume-based share of AI workers from 2010 to 2018 display a correlation of 0.27 with changes in log R&D expenditures over the same time period, controlling for industry fixed effects. The pattern of AI-investing firms increasing research and development (R&D) expenditures supports the notion that AI investments involve a great deal of experimentation with applying the new technology (Braguinsky et al., 2020).

Fourth, digging deeper into the skills and jobs with the highest AI-relatedness measures according to our methodology, we observe that our measure is indeed capturing the essence of AI investments by firms. The skills with the highest AI-relatedness measures, presented in Online Appendix Table A1, are highly AI-specific skills, such as "Tensorflow" and "Random Forests," while general data-analytics-related skills have low AI-relatedness measures: for example, the measure is equal to 0.04 for "Data Modeling" and 0.03 for "Quantitative Analysis." Similarly, Online Appendix Table A3 shows that the job titles associated with the highest job-level measures of AI-relatedness are all very relevant postings such as "Artificial Intelligence Engineer" (average AI-relatedness measure of 0.497), "Senior Data Scientist - Machine Learning Engineer" (0.394), and "AI Consultant" (0.369). Since we do not require information contained in job titles of job postings to identify AI-related skills and jobs, these patterns provide additional validation that our measure captures relevant AI positions.

As a further validation, it is worth noting the geographic locations of the identified AI jobs. We aggregate firm-level AI investments of Compustat firms to the commuting zone level and link the commuting-zone-level changes in the share of AI workers from 2010 to 2018 to 2010 commuting zone characteristics from the Census American Community Survey. Online Appendix Figure

A1 displays a heat map of the growth in the average job-level AI relatedness measure from 2010 to 2018 and shows that there is significant variation in AI investments across commuting zones. Online Appendix Figure A1 (a) shows a strong positive relationship between the change in the share of AI workers from 2010 to 2018 and the average commuting-zone-level log wage in 2010. Online Appendix Figure A1 (b) demonstrates that the growth in AI workers is also concentrated in commuting zones with a large fraction of college-educated workers. These patterns are intuitive, given that AI employees tend to be high-skilled technologically-oriented workers, and contrast with investments in robotics, which concentrate in areas with larger shares of manufacturing employment (Acemoglu et al., 2020).

Finally, we observe a high correlation between our two measures of AI investments (see Table 1), which helps address the concern that firms' job postings may not be sufficient to capture firms' actual hiring of AI talent. For example, if a firm is unable to fill AI-related vacancies, the job postings measure will overstate that firm's investments in AI. In practice, this does not appear to be a main driver of the job-postings-based measure, because the two measures of AI investments—using job postings and resumes—yield similar results throughout the remainder of the paper.

High correlations and consistency across our two measures also address the potential concern that job postings do not reflect firms investing in AI by acquiring other firms (e.g., AI startups). Human capital on-boarded through acquisitions *is* captured by the resume data, where employees of acquisition targets are counted as employees of the acquirer subsequent to the acquisition. This argument (that AI-skilled labor is not mainly acquired via acquisitions) is also supported by the smooth increase (without jumps) of most firms' AI workers (representative time series for a few large firms are presented in Online Appendix A.1) and by industry reports estimating that 90% of firms' investments in AI are internal, with only 10% is spent on acquisitions (Bughin et al., 2017).

3.4 Firm-level Determinants of AI Investments

We consider the determinants of investments in AI technologies and document that larger firms and firms with higher markups, cash reserves, and R&D tend to invest in AI more aggressively.

Our focus is on understanding the *use* of AI technologies by a wide range of firms, rather than the invention of new AI tools. For that reason, we exclude firms in the tech sector (2-digit NAICS 51 or 54) from our main empirical analyses in this and the following sections.[18] Our main regression sample is comprised as follows. In 2010, there are 3735 U.S.-listed public firms that have non-missing industry codes, positive sales and employment, and are not in the tech sector.

[18] In later analyses, we confirm that the main effects of AI spurring firm-level growth are also present in these industries. A complementary analysis of the impact of AI on specifically AI-inventing firms is provided by Alderucci et al. (2020).

Among these firms, 2668 are matched to Cognism,[19] and 1933 are matched to Burning Glass. For the Cognism sample, we further restrict to firms with at least 20 U.S. jobs in both 2010 and 2018 to ensure good coverage of the firm's workforce, which leaves us with 1993 firms. For the Burning Glass sample, we further restrict to firms that are also matched to Cognism, so that we can cross-validate with the actual hiring, leaving us with 1192 firms.

In Table 2, we examine which ex ante firm characteristics predict future growth in firm-level AI investments. For each measure of AI investments, we estimate the following specification:

$$\Delta ShareAIWorkers_{i,[2010,2018]} = \beta FirmVariable_{i,2010} + IndustryFE + \epsilon_i, \tag{1}$$

where $\Delta ShareAIWorkers_{i,[2010,2018]}$ denotes the change in the share of firm i's AI-related employees (job postings) from 2010 to 2018 in Panel 1 (Panel 2). All regressions include 2-digit NAICS industry fixed effects. Here and throughout all subsequent analyses, the $\Delta ShareAIWorkers_{i,[2010,2018]}$ variables are standardized to mean zero and standard deviation one to aid in economic interpretation. $FirmVariable_{i,2010}$ represents one of the ex ante firm characteristics of interest measured as of 2010: log firm sales in column 1, the ratio of cash to total assets (Cash/Assets) in column 2, the ratio of R&D expenditures to sales (R&D/Sales) in column 3, revenue total factor productivity (TFP)[20] in column 4, log markup measured as the log of the ratio of sales to cost of goods sold following De Loecker et al. (2020) in column 5, Tobin's Q defined as market value of assets divided by book value of assets in column 6, market leverage measured as total debt divided by market value in column 7, return on assets (ROA) measured as the ratio of net income plus interest expense to assets in column 8, and firm age in column 9. Column 10 includes all variables in a multivariate specification. We winsorize all continuous variables at 1% and 99% to limit the influence of outliers, although we confirm in untabulated analyses that, empirically, our results are little changed by the winsorization. To account for differences in precision in the measurement of AI investments across firms with different numbers of available observations, the estimating equation is weighted by each firm's number of resumes (job postings) in 2010.[21]

The results reported in Table 2 highlight that ex ante larger firms experience higher growth in AI investments. For example, using the Cognism-based measure in Panel 1, a one-standard-deviation increase in log sales in 2010 (which equals 2.1) corresponds to the share of AI workers

[19] Firms that are not matched to Cognism tend to be either ADRs that do not have many U.S. employees or smaller firms with few employees.

[20] We use standard methodology to calculate revenue TFP as the residual from regressing log real sales on log employment and log capital, controlling for firm fixed effects and year fixed effects: $\log y_{it} = \mu_i + \mu_t + \alpha_s^l \log(l_{it}) + \alpha_s^k \log(k_{it-1}) + \varepsilon_{it}$. The regression is estimated using OLS separately for each industry. The capital stock is constructed using the perpetual inventory method. The TFP measure is specific to Cobb-Douglas production functions, while sales per worker measure labor productivity for more general production functions.

[21] Since the numbers of worker resumes and job postings are correlated with firm size, this weighting scheme also roughly weights firms in accordance to their contribution to the economy.

increasing by 23% of the standard deviation from 2010 to 2018, significant at the 1% level. In addition, firms with higher starting Cash/Assets, R&D/Sales, and markups also see greater investments in AI, consistent with contemporaneous work of Alekseeva et al. (2020). By contrast, revenue total factor productivity, firm valuation (Tobin's Q), market leverage, return on assets, and firm age do not robustly predict future AI investments. In all further regressions, we control for the ex-ante firm characteristics that predict firm AI adoption (size, cash, R&D intensity, and markups). Importantly, the patterns for firm-level demand for AI talent measured with Burning Glass data are consistent with the results using Cognism data, reinforcing the high correlations documented in Table 1. This consistency suggests that, in the absence of matched employer-employee data, our methodology for identifying AI investments from the job postings data can be a good proxy for firms' actual AI hiring.

4 AI Investments and Firm Growth

We next document that firms investing in AI technologies grow faster in sales, employment, and market value, and that this effect cannot be explained by alternative explanations, including reverse causality (e.g., firms on faster growth trajectories investing more in AI) and omitted variables (e.g., concurrent investments in other technologies or demand shocks driving both firm growth and AI investments).

4.1 Long-differences Results

We begin the analysis by examining whether firms that invest in AI see faster growth from 2010 to 2018. As is standard in settings with slow-moving processes, such as technological progress (e.g., Acemoglu and Restrepo, 2020), our primary specification is a long-differences regression of changes in firm outcomes from 2010 to 2018 on changes in AI investments proxied by the share of AI workers. This strategy is especially well-suited for our setting because AI investments are gradual over time (with 70% of firms onboarding AI workers over a span of multiple years), with effects that may not be immediate. By taking first differences in independent and dependent variables, the long-differences specification helps to ensure that time-invariant firm characteristics do not drive the results. In Table 3, we report the estimates from the following regression:

$$\Delta FirmVariable_{i,[2010,2018]} = \beta \Delta ShareAIWorkers_{i,[2010,2018]} + Controls'_{i,2010}\gamma + IndustryFE + \epsilon_i, \quad (2)$$

where the main independent variable, $\Delta ShareAIWorkers_{i,[2010,2018]}$, captures the change in the share of AI workers in firm i from 2010 to 2018, standardized to mean zero and standard deviation of one. As in Section 3.4, this analysis focuses on firms in non-tech sectors. $IndustryFE$ are

2-digit NAICS fixed effects.[22] Panel 1 reports the coefficients for the resume-based measure of AI investments, while Panel 2 considers the job-postings-based measure. In columns 1, 3, and 5 we include only industry fixed effects to examine the unconditional relationship between changes in AI investments and firm growth. In columns 2, 4, and 6, we include a rich set of controls that are all measured at the start of the sample period in 2010: (i) the initial firm-level characteristics that predict changes in AI investments in Section 3.4 (log sales, cash/assets, R&D/Sales, and log markup) and the log of the firm's total number of jobs (or job postings)[23]; (ii) characteristics of the commuting zones (CZ) where the firms are located (the share of workers in IT-related occupations, the share of college-educated workers, log average wage, the share of foreign-born workers, the share of routine workers, the share of workers in finance and manufacturing industries, and the share of female workers); and (iii) the log industry-average wage.[24] Out of the 1993 (1192) non-tech firms in the Cognism (Burning Glass) sample in Table 2, 1472 (939) firms have positive sales and employment in 2018, which are necessary to calculate the dependent variables. We further restrict the sample to firms with non-missing control variables throughout, to keep the sample composition stable. This results in a sample of 1052 firms in Cognism and 935 firms in Burning Glass. The results of the regressions without controls are similar when estimated on the entire available sample. Summary statistics on key variables for the main regression sample are provided in Online Appendix Table A4.

In columns 1 and 2 of Table 3, the dependent variable is the firm-level change in log sales from 2010 to 2018. Both measures of changes in AI investments are associated with a significant and economically meaningful increase in sales growth: a one-standard-deviation increase in the share of AI workers over an eight-year period corresponds to an additional 15% to 20% growth in sales, depending on the specification.[25] In columns 3 and 4, we find a positive effect on employment of a similar magnitude to the effect on sales. This suggests that AI is not yet displacing firms' workforces, at least on net, although we do not rule out the reallocation of labor across different job functions or tasks. In untabulated analyses, we confirm that the results are similar when using changes in employee counts in the Cognism resume data rather than Compustat employment. Columns 5 and 6 show that firms investing in AI also see increases in their stock market valua-

[22] In Online Appendix Table A7, we show that our results are robust to controlling for industry at 3-digit NAICS, 4-digit NAICS, and 5-digit NAICS level. The coefficient on $\Delta ShareAIWorkers_{i,[2010,2018]}$ remains stable, and the standard error increases as more granular industry controls absorb more of the variation.

[23] We control for the log number of jobs to address the concern that the share of AI jobs may be more volatile in firms with fewer total jobs. This control ensures that the variation in the share of AI jobs is between firms with similar total numbers of jobs but different numbers of AI jobs.

[24] When firms span multiple commuting zones, we calculate commuting-zone-level variables as the weighted average, using numbers of BG job postings in each commuting zone as weights, which restricts the sample in the Cognism regression analysis to firms that are also matched to the Burning Glass data. The results are similar in magnitude and economic significance if we only include firm-level controls enumerated in list (i).

[25] A one-standard-deviation increase in the share of AI workers is roughly at the 90th percentile of the distribution of changes in the share of AI workers (see Online Appendix Table A4).

tions: a one-standard-deviation increase in the share of AI workers is associated with a 15%–23% increase in the firm's market value.[26] It is worth noting that the inclusion of firm-level, location-level, and industry-level controls in even columns (all measured at the start of the sample period in 2010) generally has little effect on the estimated coefficients. This is consistent with our long-differences specification already controlling for time-invariant firm characteristics.

The magnitude of the effects in Table 3 is economically meaningful (on the order of a 2% increase in annual sales growth per one-standard-deviation increase in the share of AI workers). Our results provide initial evidence that AI-skilled labor can have a strong positive relationship with firm growth. In this context, our results are consistent with prior evidence that certain key, high-skilled employees—including chief executives, inventors, and entrepreneurs—can have a disproportionate effect on firm outcomes.

The positive relationship between increases in AI investments and firm growth is ubiquitous across different sectors of the economy, reinforcing the notion that AI is a general purpose technology. Online Appendix Table A5 displays the results from regressing changes in log sales and log employment on the change in the share of AI workers, separately for the largest 2-digit NAICS sectors: (i) Manufacturing, (ii) Wholesale and Retail Trade, (iii) Finance, and (iv) the remaining non-tech sectors. While we exclude tech sectors from our main analysis, we find that AI also has an positive relationship with growth for firms in the two tech sectors—Information and Professional and Business Services (see Online Appendix Table A6). Overall, we observe that investments in AI are associated with economically significant increases in firms' operations, and these effects are meaningful across key economic sectors.

However, the benefits from AI investments are not evenly distributed across the firm size distribution. Table 4 shows the relationship between changes in AI investments and firm growth, across terciles of firms by employment in 2010 (within the firm's 2-digit NAICS sector), controlling for initial size and sector-by-size-tercile fixed effects. The effect of AI investments on employment, sales, and market value is monotonically increasing in the firm's initial size. The stronger positive relationship between changes in AI investments and growth among the ex ante larger firms is consistent with big data and AI technologies having scale effects that favor large firms, which accumulate large amounts of data as a by-product of their economic activity (Farboodi et al., 2019). Akcigit and Kerr (2018) highlight that larger firms face constraints on their ability to scale due to higher costs of new product innovation. The results in Table 4 suggest that AI may provide a channel through which large firms can combat barriers to innovation and scale by leveraging their data assets. For example, biotech firms that have accumulated large troves of proprietary samples of molecular compounds are able to leverage AI tools to obtain an advantage over competitors

[26] Market value is defined as total assets (*at*), minus the book value of common equity (*ceq*), plus the market value of common equity (*prcc_c* times *csho*).

(see here for the PitchBook AI & ML Emerging Tech Report 2021).

4.2 Robustness

Our granular data allow us to rule out a number of alternative explanations for our results. We first discuss measurement strengths of our data and show robustness to using alternative constructions of the AI measure. We then address several identification concerns and end this section by addressing concerns related to sample selection bias.

Measurement strengths and potential concerns. First, it is worth noting that the resume data address two important potential measurement concerns regarding job postings data: (i) that the job-postings-based measure captures only firms' demand for AI talent and not their actual ability to hire; and (ii) that firms acquire AI expertise through acquisitions, which would not be reflected in job postings. Cognism resume data reflect actual employees, including those onboarded through acquisitions, and the results are consistent across these richer data and the BG job postings, validating the use of job postings for measuring firms' AI-skilled human capital.

Second, while our measure is centered on internal AI investments, our rich resume data allow us to also consider whether firms' use of external AI solutions might affect the interpretation of our results. Even external AI software requires internal data management and implementation guidance by AI-skilled workers to be effective (Fedyk, 2016), and industry reports underscore that AI-skilled labor is the most critical input to successful deployment of AI programs. Nevertheless, we leverage the rich detail of the Cognism resume data to confirm that our approach of focusing on AI workers to identify AI investments is indeed a suitable one. We undertake a deep dive into case studies of individual firms (see Online Appendix A.1 for examples) and observe that (i) our measure captures internal AI investments well, and (ii) the use of external AI software solutions (e.g., IBM Watson, IPSoft Amelia) tends to be complementary to internal AI hiring. In addition, we process individual job descriptions and job titles in our resume data for any mention of external AI software (including IBM Watson Studio, Symphony, AyasdiAI, Salesforce Einstein, and about a hundred other key AI-powered solutions) to construct a proxy for firms' reliance on external AI solutions.[27] In Online Appendix Table A8, we confirm that our results are robust to directly including this proxy in our overall measure of AI investments.

Third, we confirm that our results are not sensitive to different methods of constructing firm-level AI measures. In Online Appendix Table A9, we show that the results are robust to using firm-level average continuous AI-relatedness measures of job postings based on all skills (Panel

[27] The Cognism resume data are especially well-suited to capture the use of external technological solutions, given Cognism's emphasis on developing "technographic data" (defined by Cognism as "the technologies that the employee or company is using"). Cognism advertises these data for two purposes: (i) enhancing technology-providers' targeted marketing of their products, and (ii) improving individual firms' understanding of which technologies are used by their competitors.

1) or refined narrow AI skills (Panel 2), which are defined in Section 3.1. The narrow AI measure excludes non-AI-specific skills such as general programming ("Python") and statistics (e.g., "Linear Regressions"); however, Panel 2 shows that this refinement does not affect our results. In Online Appendix Table A10, we also find similar results when using higher or lower cutoffs for classifying AI job postings based on their continuous measure of AI-relatedness.

Identification concerns. We conduct several tests to show that our results are not driven by reverse causality or omitted variables concerns, such as differential growth trends or investments in other technologies. First, we address the concern that AI-investing firms might already be on higher growth trajectories, leading to a potential positive bias in our estimates. In Online Appendix Table A11, we control for past industry-level and firm-level growth in the decade before our sample period (from 2000 to 2008) and find similar results. In Online Appendix Table A12, we further confirm that the results are robust to the addition of controls for (i) Tobin's Q as of 2010, which proxies for the firm's future growth opportunities, and (ii) state fixed effects, which control for growth opportunities and other potential omitted variables at the state level. Furthermore, Table Online Appendix A13 estimates a predictive regression of firm growth during the later part of our sample (2015–2020) on growth in AI investments during the earlier part of the sample (2010-2015).[28] The estimates are qualitatively and quantitatively similar to those in Online Appendix Table 3, with milder magnitudes corresponding to the shorter estimation period (growth from 2015 to 2020 rather than 2010–2018), pointing against reverse causality driving our results.

Second, we leverage our detailed data to address omitted variable concerns related to firms' potential use of non-AI technologies driving our results—which is an important point, given recent evidence that investments in information technology (IT) are also correlated with firm growth (Tambe et al., 2020). Our rich data allow us to develop measures of investments in non-AI technologies that parallel the measure of AI investments: for each firm, we measure the percentage of job postings in each year requiring IT-, robotics- or data-related skills that are not specific to AI. In Online Appendix Table A14, we test whether our measure of AI investments captures investments in other technologies by estimating the relationship between changes in AI investments and firm-level growth, controlling for growth in: (i) investments in (non-AI) IT, (ii) investments in robots, (iii) investments in non-AI data skills (e.g., "Data Cleaning"), and (iv) investments in non-AI-related data analytics (e.g., "SAS"). The estimated relationship between growth in AI investments and firm growth remains similar with the addition of these controls, confirming that the documented effects on firm growth are specifically driven by AI rather than by other technologies.

Sample selection bias. A remaining concern is that the long-differences specification requires AI-investing firms to be present at the beginning (2010) and the end of the sample (2018), and that

[28] In these regressions, we can use firm growth estimated through 2020, because this specification does not require firm AI data (which end in 2018) to go beyond 2015.

survivorship bias might affect our estimate of the effect of AI. While, by construction, we do not observe growth of firms that are not present in both 2010 and 2018, we perform the industry-level analysis in Section 6 for both: (i) only firms in our main sample, and (ii) including entering and exiting firms. If the composition of firms changes in an important way, then the estimates of the relationship between growth in AI investments and industry-level growth would be significantly different across these two samples—which is not the case empirically. Moreover, in the next section we use panel structure of the data—that does not condition on firms being present over the entire sample period—to show in a dynamic setting that, similar to long-differences specification, firms grow more following AI investments.

4.3 Dynamic Effects

We augment our long-differences specification by estimating firm growth dynamically following AI investments. This analysis not only offers additional evidence against reverse causality concerns and AI-investing firms being on differential growth trajectories prior to AI investments, but also elucidates the lag between AI investments and their realized effects.

We use firm-level panel data to estimate firm growth dynamically around the years of AI investments in a distributed lead-lag model, which allows for continuous variation in the treatment variable (Aghion et al., 2020; Stock and Watson, 2015). This specification is especially well-suited to our setting, because firms tend to invest in AI on a continuous basis, rather than make lumpy investments in a single year, which precludes us from examining dynamic effects in a standard event-study framework with discontinuous treatment (e.g., before and after a lumpy investment).[29] The standard distributed lead-lag model is specified as:

$$Y_{it} = \sum_{k=-2}^{5} \delta_k \Delta ShareAIWorkers_{i,t-k} + \mu_i + \lambda_{nt} + \theta_{st} + \epsilon_{it} \tag{3}$$

where $\Delta ShareAIWorkers_{i,t-k}$ is the annual change in the share of AI workers from year $t-k-1$ to year $t-k$, normalized to mean zero and standard deviation of one, and Y_{it} is either log sales or log employment in year t. We include firm fixed effects μ_i to absorb firm-specific time-invariant factors, and industry-year fixed effects λ_{nt} and state-year fixed effects θ_{st} to control for industry-specific and state-specific trends. Each lead-lag coefficient δ_k captures the cumulative response of the outcome variable in year t to AI investments in year $t-k$, holding fixed the path of AI investments in all other years. As such, specification (3) incorporates both immediate and delayed responses of firm size to firms' AI investments.[30] The estimated coefficients for the leads can be

[29] The percentage of AI-investing firms that only invest in a single year is 29.5%, compared to 70.6% for robotics (Humlum, 2019).

[30] For each firm-year observation of sales or employment between 2010 and 2016, we consider five lags and two leads,

used as a pre-trend test: if firms investing in AI are on similar growth trends as other firms prior to AI investments, δ_k with $k < 0$ should be statistically indistinguishable from zero.[31]

Figure 3 reports the coefficients from the lead-lag regressions. The top panel shows that sales increase following AI investments, but not immediately—it takes two to three years for firms to realize the benefits from AI investments. The cumulative effect of a one-standard-deviation increase in annual AI investments on log annual sales is 1.5%–2% and remains steady five years out. This is consistent with the long-differences estimates in Section 4.1, where a one-standard-deviation increase in AI investments is associated with a 15%–20% increase in sales over eight years. The bottom panel shows that AI investments have a similar positive effect on firms' employment. Importantly, there is no evidence of pre-trends in either outcome variable: conditional on the controls we include, firms that invest more in AI in any given year show comparable sales and employment paths in prior years and start diverging only afterwards. This provides additional evidence that our results are not capturing the reverse effect of firm growth on AI investments or the effect of omitted variables placing AI-investing firms on differential growth trajectories, helping to bolster a causal interpretation of our main results.

4.4 Instrumenting AI Investments

To further address endogeneity and measurement concerns, we instrument firm-level changes in AI investments using variation in firms' ex-ante exposure to the supply of AI talent from universities that are historically strong in AI research. The core idea is that the scarcity of AI-trained labor is one of the most important constraints to firms' AI adoption (e.g., CorrelationOne, 2019), and universities that are historically strong in AI research have been able to train more AI-skilled graduates in recent years, enabling firms that typically hire from those universities to more readily attract AI talent. Since commercial interest in AI became widespread only around 2012, we argue (and offer empirical support) that firms' connections to AI-strong universities in 2010 were not driven by the need to hire AI-skilled workers, especially for the sample of non-tech firms that are the focus of this paper. To construct the instrument, we compile two datasets on: (i) the ex-ante strength of AI research in each university, and (ii) firm-university hiring networks. To the best of our knowledge, there is no comprehensive historical data on either of these two aspects. We now

so that we estimate the cumulative impact of AI investments on firm growth from two years before the investments to five years after the investments. Since the data on AI investments end in 2018, we include only two leads to keep all firm-year observations up to 2016. We obtain similar results when including only one lead or no leads at all. Furthermore, the analysis of dynamic effects focuses on the Cognism resume data, because these data offer full coverage of AI investments going back to 2005. By contrast, Burning Glass job postings data have a more limited time series, where including 2 leads and 5 lags would restrict the sample to only firm-year observations in 2015 and 2016.

[31] It is worth noting that, given that the independent variables in this distributed lead-lag model are changes in continuous AI investments instead of period dummies as would be the case in a standard event-study framework, we cannot normalize the estimates to an exact zero for any given period.

briefly discuss the construction of both datasets, while Appendix A provides a detailed discussion of these issues.

To identify universities strong in AI research before 2010, we use data from the Open Academic Graph (OAGv2), which provides the most comprehensive openly available repository of scholarly work since 1870 (Sinha et al., 2015; Tang et al., 2008). We match 689 research institutions in the National Science Foundation's Higher Education Research and Development Survey (HERDS) to researchers in the OAGv2 and work with the field experts at the AI for Good Foundation to identify AI-related publications. We classify each AI researcher based on the share of AI publications in that researcher's overall portfolio, and we classify universities as AI-strong if their number of AI researchers is at the top of the distribution over 2005–2009.

A key concern with our instrument is that AI-strong universities are also likely to be strong in the broader field of computer science (CS), producing more CS-skilled graduates, which might affect firm outcomes through channels other than AI investments. To address this concern, we also measure the number of CS researchers in each university to include as a control. In addition, we verify that AI-strong universities are not strongly correlated with the overall university rankings.

We construct the firm-university hiring networks by leveraging our resume data to observe the universities granting the degrees of each firm's employees.[32] For the firm-university hiring networks to provide the necessary variation for our instrumental variable strategy, different firms need to hire from different sets of universities, and these networks need to be persistent over time. Our data show evidence of both: each firm tends to concentrate its hiring in a small number of universities, and ex-ante networks (i.e., which universities each firm hired from before 2010) strongly predict the universities from which firms hire after 2010 (see Appendix Table 9).

We define our instrument for each firm i as: $IV_i = \sum_u s_{iu}^{2010} AIstrong_u$, where s_{iu}^{2010} is the share of STEM workers in firm i in 2010 who graduated from university u, and $AIstrong_u$ equals one if university u is identified as an AI-strong university based on pre-2010 publications.[33] To control for the effects of general computer science (and not specifically AI), we construct an analogous measure of firms' exposure to ex-ante CS-strong universities: $\sum_u s_{iu}^{2010} CSstrong_u$.

We examine an important identification concern regarding this instrument: if firms anticipated the surge in demand for AI, they could have started building their connections to AI-strong universities before 2010, making firm-university hiring networks in 2010 endogenous to firms' de-

[32] Aggregated to the university-year level, our resume data cover, on average, 59% of all degrees conferred by each university according to IPEDS data, and the number of fresh graduates in the resume data is highly correlated with the total number of degrees conferred (correlation=0.73) in the IPEDS data. Confirming the relevance of our measure of AI-strong universities, Appendix Figure 5 shows that the increase in AI-trained graduates during the 2010s was much more pronounced in ex-ante AI-strong universities than in non-AI-strong universities.

[33] We use firm-university hiring networks based on STEM workers to account for potential segmentation in firms' hiring networks, where business employees may be hired from different universities than technically skilled employees. However, empirically, firm-university hiring networks constructed from all workers yield similar results.

mand for AI-trained students. This is unlikely, given the lack of both commercial interest in AI by firms and AI-skilled graduates by universities prior to 2010 (see Appendix Figure 5). Moreover, we confirm empirically that firms connected to AI-strong universities in 2010 did *not* increase their share of hired fresh graduates from those universities from 2005 to 2010 (see Appendix Table 10).

Appendix Table 11 shows the first-stage results with industry fixed effects and the CS control included throughout. We sequentially add (i) baseline controls (firm-, industry-, and commuting-zone-level controls), (ii) pre-period firm sales and employment growth between 2000 and 2008 to address unobservable firm characteristics that might simultaneously drive firms' growth trajectories and their hiring of AI workers, and (iii) state fixed effects to control for local labor market characteristics that might drive both firms' AI hiring and and their growth. The instrument has a strong first stage with F-statistics well above the conventional level of 10 for all specifications using the Cognism resume data. The F-statistics are also above 10 for two out of four specifications using the Burning Glass job postings data. Intuitively, the first stage is stronger in the Cognism data because the data generating process for the instrument is based on Cognism resumes and captures the *supply* of AI-skilled labor to firms. As a result, we focus on Cognism resume data in the second stage in Appendix Table 5. The results show a robust and significant effect of AI investments on sales (columns 1–4), employment (columns 5–8), and market value (columns 9–12). Online Appendix Table A15 shows similar results for the job postings data.

5 Mechanisms

We examine the drivers of AI-fueled firm growth by considering the two non-mutually-exclusive mechanisms detailed in Section 1. We document that AI-investing firms are able to significantly increase their product innovation and find no evidence of reductions in operating costs.

5.1 AI as a Driver of Product Innovation

As we outline in Section 1, AI can contribute to firm growth via product innovation by: (i) facilitating the creation of new and improved products, and (ii) increasing product scope through improved tailoring of products to customer tastes. To explore this empirically, we need firm-level data on products and services, which are challenging to obtain, especially across different sectors. We overcome this challenge by using three proxies for firms' product innovation.

First, we examine whether AI-investing firms experience increases in trademarks, which are registered whenever new products or services are ready for commercialization and therefore offer a good proxy for the creation of new products and services (Hsu et al., 2021). Columns 1 and 2 in Table 6 present the results from long-differences regressions of changes in firms' USPTO

trademarks against growth in their AI investments, showing that AI-investing firms significantly increase their trademark portfolios.[34] Second, columns 3 and 4 reveal a similar relationship between AI investments and the number of product patents, which are patents specifically focusing on product innovations.[35] While trademarks are registered with the creation of new products, product patents reflect both new product creation and innovations in the quality of existing product lines. We find that a one-standard-deviation increase in the share of AI workers based on resumes (job postings) over eight years corresponds to a 21% (20%) increase in the number of product patents.

Finally, we build a measure of changes in firms' product mix based on the self-fluidity measure in Hoberg et al. (2014). Using firm 10K filings, Hoberg et al. (2014) take the cosine similarity between word vectors describing a firm's product offerings in two adjacent years to measure the extent to which the firm's product offerings changed in a given year. These changes reflect both the creation of new products and the tailoring of existing products to evolving consumer tastes.[36] In columns 5 and 6, we find that growth in AI investments is associated with increased changes in firms' product mix from 2010 to 2018. For robustness, Online Appendix Table A16 shows that the instrumented firms' AI investments also have a positive effect on the number of trademarks, the number of product patents, and the change in product offerings (although not always significant). Overall, the results point towards firms utilizing AI to expand product variety and customization, consistent with surveys of corporate executives, who highlight product improvement and creation as top uses of AI (see here).

Our findings provide a first piece of evidence for how AI technologies can stimulate growth for a broad set of firms: the unique reliance of AI on big data reduces the uncertainty of exploration (Cockburn et al., 2018), facilitates the discovery process for new or better products, and enables the tailoring of products to customer tastes. Moreover, AI algorithms themselves can be used as an ingredient in product development and improvement (e.g., AI-powered trading platforms or self-driving cars). These results are consistent with evidence from technological innovation during industrialization, where new technologies have been shown to help firms expand through

[34] The dependent variable is the change in log(1 + number of trademarks) from 2010 to 2018, so that the regression takes into account firms with zero trademarks in either 2010 or 2018. The results are also robust to using the inverse hyperbolic sine transformation (i.e., $\ln(x + \sqrt{(1 + x^2)})$). The regression sample is smaller than our baseline sample, because not all public firms file trademarks (we include firms with at least one trademark in 2009-2018).

[35] See Ganglmair et al. (2021) for the methodology to distinguish between product patents and process patents. The regression sample is smaller than our baseline sample, because not all public firms file patents, and we only include firms with at least one patent during 2005–2018. The dependent variable is the change in log(1 + number of product patents) from 2010 to 2018.

[36] We use the same word vectors as Hoberg et al. (2014) and construct our measure as follows: for each year, we calculate the angle between the two word vectors indicating firms' product offerings in that year and the previous year. For example, the measure equals 0 if the product offerings remain exactly the same and $\pi/2$ if the product offerings change completely. We sum up of angle of each year over eight years from 2010 to 2018 to measure the total change in firms' product portfolios from 2010 to 2018.

product innovation (Braguinsky et al., 2020).

5.2 AI as a Driver of Lower Operating Costs

We next test whether the increase in firm growth from AI investments could reflects AI technologies lowering firms' operating costs and increasing firm-level productivity. First, in columns 1 to 4 of Table 7 we look at costs directly by considering how growth in firms' AI investments relate to changes in costs of goods sold (COGS) and operating expenses. AI investments are associated with increases in costs that are similar in magnitude to the growth in firm sales, suggesting that AI is not associated with lower operating costs.

Second, columns 5 to 8 of Table 7 consider two measures of productivity: sales per worker (i.e., labor productivity) and revenue total factor productivity (TFP). The relationship between AI investments and both productivity measures is not significant. The lack of growth in labor productivity is consistent with the results in Section 4 that AI investments predict similar increases in sales and employment, challenging the view that the primary effect of AI is to replace jobs.[37] Furthermore, in columns 9 and 10, we bring another proxy for efficiency gains that complements revenue-based measures of productivity: process patents, which reflect process innovations and potential improvements in efficiency. We find a zero relationship between AI investments and process innovation, in contrast to the positive increase in product patents documented in Table 6.

Overall, we do not find evidence that investments in AI help firms cut their operating expenses and achieve productivity improvements. This speaks to the broader debate on the timing of productivity gains from general purpose technologies. Our evidence is consistent with long-standing arguments in the literature that the adoption of general purpose technologies leads to delayed productivity benefits (David, 1990; Brynjolfsson et al., 2020). In Online Appendix Table A17, we examine the effect of changes in AI investments during the first half of the period (2010–2014) on productivity growth through 2018 and do not find any significant positive effect. Hence, even with a lag of a few years, AI investments are not yet associated with productivity improvements.

6 AI Investments and Industry-level Outcomes

To shed light on the potential aggregate effects of AI, we examine the relationship between industry-level variation in AI investments and: (i) industry growth; (ii) industry concentration.

[37] It is worth noting that both sales per worker and revenue TFP are revenue-based measures of productivity and may not fully reflect actual physical productivity. For example, sales per work and revenue TPF may provide downward-biased estimates of actual productivity changes if quantities produced increase to such an extent that lower prices are charged (Foster et al., 2008; Garcia-Marin and Voigtländern, 2019; Caliendo et al., 2020). To consider this possibility, in untabulated analyses we find that there are no changes in AI-investing firms' markups.

While AI-investing firms grow faster, the gains in industry sales and employment may be zero-sum if the use of AI technologies creates a business-stealing effect on competitors (Bloom et al., 2013). For example, negative spillovers have been shown to dominate positive firm-level effects in the case of robotics, leading to an overall negative effect on aggregate employment (Acemoglu et al., 2020). Hence, signing the aggregate effect of AI investments is an empirical question. We estimate the following long-differences regression at the industry level:

$$\Delta \ln y_{j,[2010,2018]} = \gamma \Delta ShareAIWorkers_{j,[2010,2018]} + IndustrySectorFE + \epsilon_j \tag{4}$$

where $\Delta \ln y_{j,[2010,2018]}$ is the change in total sales or employment for all Compustat firms (including those that entered the sample after 2010 or exited before 2018) in 5-digit NAICS industry j, and $\Delta ShareAIWorkers_{j,[2010,2018]}$ is the change in the share of AI workers among Compustat firms in industry j from 2010 to 2018. Analogously to the firm-level tests, the regressions are weighted by the total number of resumes (or job postings) in each industry in 2010.

Columns 1–4 of Table 8 show that AI investments are associated with a robust increase in employment and sales at the industry level. In both panels (Panel 1: resume-based AI measure, and Panel 2: job-postings-based AI measure), odd columns estimate the unconditional relationship (with 2-digit NAICS fixed effects only), and even columns add controls for log employment, log sales, and log average wages at the industry level in 2010. For example, with the full set of controls, a one-standard-deviation increase in the industry-level share of AI workers in the resume data is associated with a 19.9% increase in sales and a 23.4% increase in employment. Importantly, in Online Appendix Table A18, we show that the results remain similar when we restrict the sample to firms that are in the Compustat sample both in 2010 and 2018 (i.e., excluding entrants and exits). This confirms that sample selection issues are not driving our main results for publicly traded firms. While we cannot speak to growth effects outside of publicly traded firms (where sales data are not reported), the fact that our effects concentrate among the largest firms in the Compustat sample (Table 4) suggests that the net effects on industry growth of all (public and private) firms are likely milder than those documented in Table 8.[38]

We next examine whether the higher AI-fueled growth among larger firms is substantial enough to translate into increased industry concentration. We link industry-level growth in AI investments to contemporaneous changes in industry concentration from 2010 to 2018. Following Autor et al. (2020), we use the Herfindahl-Hirschman Index (HHI) to measure industry concentration. To examine winner-take-most dynamics, we also consider the fraction of sales accruing to

[38] A caveat with these results is that the Compustat sample assigns each firm to a single main industry, even for firms that might have operations in several industries. This caveat is unlikely to affect the interpretation of our results, given that prior research using U.S. Census micro data shows that for a typical U.S. public firm the large majority of its operations fall within one main industry (Babina, 2020).

the largest firm in each 5-digit NAICS industry among the Compustat firms. Columns 5–8 of Table 8 show a positive relationship between industry-level growth in AI investments and changes in industry concentration.

Overall, our results support the argument by Crouzet and Eberly (2019) that investments in intangible assets are responsible for the rise in industry concentration observed in the U.S. data. Our results suggest that, as a general purpose technology that can be applied across many industries, AI has the potential to further increase concentration across a broad range of industries by facilitating product innovation and the expansion for the largest firms.

7 Conclusion

In this paper, we study how firms invest in and benefit from one of the most important new technologies of the last decade—artificial intelligence. We introduce a novel measure of investments in AI technologies at the firm level using two detailed datasets on human capital: job postings from Burning Glass Technologies, which indicate each firm's demand for particular skills, and resume data from Cognism, which reveal the actual composition of a firm's workforce. Our unique measure allows us to examine both the determinants and the consequences of AI investments by firms across a wide range of sectors. We find a positive feedback loop between AI investments and firm size: AI investments concentrate among the largest firms, and as firms invest in AI, they grow larger, gaining sales, employment, and market share. This AI-fueled growth does not appear to stem from cost-cutting; instead, AI-investing firms expand through product innovation and increased product offerings.

Our findings highlight important differences between the adoption of AI technologies and the adoption of information technology (IT) in the 1980s and 1990s.[39] Much of the previous literature finds that IT investments were associated with economically large productivity increases but mixed results on firm growth measures such as market share. By contrast, we observe increased growth for AI-investing firms, along with increased product innovation, but no evidence (yet) of higher firm-level productivity. Our results also show higher AI adoption and larger gains from AI investments for larger firms, which contrasts with prior work on diffusion patterns for IT (Hobijn and Jovanovic, 2001). These differences underscore the distinctive features of AI relative to previous waves of IT: as a prediction technology, AI facilitates product innovation and creates new business opportunities by enabling firms to learn better and faster from big data.

Our findings imply that the benefits from AI depend to a large extent on who owns big data—the key input to AI technologies (Fedyk, 2016). While data are non-rival (data can be used by any

[39] See Dedrick et al. (2003) and Cardona et al. (2013) for reviews of that literature.

number of firms simultaneously), recent theoretical work suggests that, fearing creative destruction, firms may choose to hoard data they own, leading to inefficient use of nonrival data; and that giving the data property rights to consumers can generate allocations that are close to optimal (Jones and Tonetti, 2020). While our empirical work does not directly speak to the optimality of data ownership, our results suggest that AI contributes to the increase in industry concentration and the rise of "superstar" firms documented in recent work (Gutiérrez and Philippon, 2017; Autor et al., 2020). Further understanding how AI affects production processes, corporate strategies, and organizational structure of firms and assessing the distribution of gains from investing in AI technologies across firms and workers are fruitful avenues for future research.

References

Abowd, John M, Bryce E Stephens, Lars Vilhuber, Fredrik Andersson, Kevin L McKinney, Marc Roemer, and Simon Woodcock, "5. The LEHD Infrastructure Files and the Creation of the Quarterly Workforce Indicators," in "Producer dynamics," University of Chicago Press, 2009, pp. 149–234.

Acemoglu, Daron and Pascual Restrepo, "Robots and jobs: Evidence from U.S. labor markets," *Journal of Political Economy*, 2020, *128* (6), 2188–2244.

_ **, Claire LeLarge, and Pascual Restrepo**, "Competing with robots: Firm-level evidence from France," *AEA Papers and Proceedings*, 2020, *110*, 383–388.

_ **, David Autor, Jonathon Hazell, and Pascual Restrepo**, "AI and Jobs: Evidence from Online Vacancies," 2021.

_ **, Gary Anderson, David N Beede, Cathy Buffington, Eric Childress, Emin Dinlersoz, Lucia Foster, Nathan Goldschlag, John C. Haltiwanger, Zachary Kroff, Pascual Restrepo, and Nikolas Zolas**, "Technology, Firms, and Workers: Evidence from the 2019 Annual Business Survey," Technical Report 2022.

Aghion, P. and P. Howitt, "A model of growth through creative destruction," *Econometrica*, 1992, *60*, 323–351.

Aghion, Philippe, Benjamin F. Jones, and Charles I. Jones, "Artificial intelligence and economic growth," *NBER Working paper w23928*, 2017.

_ **, Céline Antonin, Simon Bunel, and Xavier Jaravel**, "What are the labor and product market effects of automation? New evidence from France," *CEPR Discussion Paper No. DP14443*, 2020.

Agrawal, Ajay, Joshua S. Gans, and Avi Goldfarb, "Artificial intelligence: the ambiguous labor market impact of automating prediction," *Journal of Economic Perspectives*, 2019, *33* (2), 31–50.

Akcigit, Ufuk and William R Kerr, "Growth through heterogeneous innovations," *Journal of Political Economy*, 2018, *126* (4), 1374–1443.

Alderucci, Dean, Lee Branstetter, Eduard Hovy, Andrew Runge, and Nikolas Zolas, "Quantifying the impact of AI on productivity and labor demand: Evidence from U.S. Census microdata," *Working paper*, 2020.

Alekseeva, Liudmila, José Azar, Mireia Gine, Sampsa Samila, and Bledi Taska, "The demand for AI skills in the labor market," *CEPR Discussion Paper No. DP14320*, 2020.

Argente, David, Doireann Fitzgerald, Sara Moreira, and Anthony Priolo, "How do firms build market share?," *Working paper*, 2021.

_ **, Salomé Baslandze, Douglas Hanley, and Sara Moreira**, "Patents to Products: Product Innovation and Firm Dynamics," *Working paper*, 2020.

Autor, David, David Dorn, Lawrence F Katz, Christina Patterson, and John Van Reenen, "The fall of the labor share and the rise of superstar firms," *Quarterly Journal of Economics*, 2020, *Forthcoming*.

Babina, Tania, "Destructive creation at work: How financial distress spurs entrepreneurship," *Review of Financial Studies*, 2020, *33* (9), 4061–4101.

_ **and Sabrina T Howell**, "Entrepreneurial spillovers from corporate R&D," Technical Report, National bureau of economic research 2018.

_ **, Asaf Bernstein, and Filippo Mezzanotti**, "Crisis innovation," Technical Report, National Bu-

reau of Economic Research 2020.

Bajari, P., V. Chernozhukov, A. Hortaçsu, and J. Suzuki, "The impact of big data on firm performance: An empirical investigation," *AEA Papers and Proceedings*, 2019, *109*, 33–37.

Basker, Emek and Timothy Simcoe, "Upstream, downstream: Diffusion and impacts of the universal product code," *Journal of Political Economy*, 2021, *129* (4), 1252–1286.

Basu, Susanto, John G Fernald, and Matthew D Shapiro, "Productivity growth in the 1990s: technology, utilization, or adjustment?," in "Carnegie-Rochester conference series on public policy," Vol. 55 Elsevier 2001, pp. 117–165.

_ , _ , and Miles S Kimball, "Are technology improvements contractionary?," *American Economic Review*, 2006, *96* (5), 1418–1448.

Berman, Nicolas, Vincent Rebeyrol, and Vincent Vicard, "Demand learning and firm dynamics: Evidence from exporters," *Review of Economics and Statistics*, 2019, *101* (1), 91–106.

Bertrand, Marianne and Antoinette Schoar, "Managing with style: The effect of managers on firm policies," *Quarterly Journal of Economics*, 2003, *118*, 1169–1208.

Bloom, Nicholas, Mark Schankerman, and John Van Reenen, "Identifying technology spillovers and product market rivalry," *Econometrica*, 2013, *81* (4), 1347–1393.

Braguinsky, Serguey, Atsushi Ohyama, Tetsuji Okazaki, and Chad Syverson, "Product innovation, product diversification, and firm growth: Evidence from Japan's early industrialization," *NBER Working Paper w26665*, 2020.

Brynjolfsson, Erik and Kristina McElheran, "The rapid adoption of data-driven decision-making," *American Economic Review*, 2016, *106* (5), 133–39.

_ , Daniel Rock, and Chad Syverson, "Artificial intelligence and the modern productivity paradox," *The Economics of Artificial Intelligence: An Agenda*, 2019, p. 23.

_ , _ , and _ , "The productivity J-curve: How intangibles complement general purpose technologies," *NBER Working Paper w25148*, 2020.

Bughin, Jacques, Eric Hazan, Sree Ramaswamy, Michael Chui, Tera Allas, Peter Dahlström, Nicolaus Henke, and Monica Trench, "Artificial intelligence the next digital frontier?," *McKinsey and Company Global Institute*, 2017.

Bustamante, Maria Cecilia, Julien Cujean, and Laurent Fresard, "Knowledge Cycles and Corporate Investment," *Available at SSRN 3418171*, 2020.

Caliendo, Lorenzo, Giordano Mion, Luca David Opromolla, and Esteban Rossi-Hansberg, "Productivity and organization in Portuguese firms," *Journal of Political Economy*, 2020, *128* (11), 4211–4257.

Cardona, Melisande, Tobias Kretschmer, and Thomas Strobel, "ICT and productivity: conclusions from the empirical literature," *Information Economics and Policy*, 2013, *25* (3), 109–125.

Chappell, Nathan and Adam Jaffe, "Intangible investment and firm performance," *Review of Industrial Organization*, 2018, *52* (4), 509–559.

Cockburn, Iain M., Rebecca Henderson, and Scott Stern, "The impact of artificial intelligence on innovation," *NBER Working Paper w24449*, 2018.

Corrado, Carol, Jonathan Haskel, Cecilia Jona-Lasinio, and Massimiliano Iommi, "Intangible investment in the E.U. and U.S. before and since the great recession and its contribution to productivity growth," *EIB Working Papers No. 2016/08*, 2016.

CorrelationOne, *Future of Data Talent, 2019 Annual Report* 2019.

Covarrubias, Matias, Germán Guriérrez, and Thomas Philippon, "From good to bad concentration? U.S. industries over the past 30 years," *NBER Working paper w25983*, 2019.

Crouzet, Nicolas and Janice C Eberly, "Understanding weak capital investment: The role of market concentration and intangibles," *NBER Working Paper w25869*, 2019.

_ , Apoorv Gupta, and Filippo Mezzanotti, "Shocks and technology adoption: Evidence from electronic payment systems," *Working paper*, 2021.

D'Acunto, Francesco, Nagpurnanand Prabhala, and Alberto G Rossi, "The promises and pitfalls of robo-advising," *Review of Financial Studies*, 2019, *32* (5), 1983–2020.

David, Paul, "The dynamo and the computer: an historical perspective on the modern productivity paradox," *AEA Papers and Proceedings*, 1990, *80* (2), 355–61.

Davis, Steven J., John Haltiwanger, Ron Jarmin, Javier Miranda, Christopher Foote, and Eva Nagypal, "Volatility and dispersion in business growth rates: Publicly traded versus privately held firms," *NBER Macroeconomics Annual*, 2006, *21*, 107–179.

Dedrick, Jason, Vijay Gurbaxani, and Kenneth L Kraemer, "Information technology and economic performance: A critical review of the empirical evidence," *ACM Computing Surveys (CSUR)*, 2003, *35* (1), 1–28.

Dickstein, Michael and Eduardo Morales, "What do exporters know?," *Quarterly Journal of Economics*, 2018, *133* (4), 1753–1801.

Eisfeldt, Andrea L. and Dimitris Papanikolaou, "Organization capital and the cross-section of expected returns," *Journal of Finance*, 2013, *68*, 1365–1406.

_ , Edward Kim, and Dimitris Papanikolaou, "Intangible value," *NBER Working Paper w28056*, 2020.

Farboodi, Maryam and Laura Veldkamp, "A growth model of the data economy.," *NBER Working Paper w28427*, 2021.

_ , Roxana Mihet, Thomas Philippon, and Laura Veldkamp, "Big data and firm dynamics.," *AEA Papers and Proceedings*, 2019, *109*, 38–42.

Fedyk, Anastassia, "How to tell if machine learning can solve your business problem," *Harvard Business Review*, 2016.

_ and James Hodson, "Trading on talent: Human capital and firm performance," *Working paper*, 2019.

Felten, Edward, Manav Raj, and Robert Seamans, "The variable impact of artificial intelligence on labor: The role of complementary skills and technologies," *Working paper*, 2019.

Felten, Edward W., Manav Raj, and Robert Seamans, "A method to link advances in artificial intelligence to occupational abilities," *AEA Papers and Proceedings*, 2018, *108*, 1–4.

Foster, Lucia, John Haltiwanger, and Chad Syverson, "Reallocation, firm turnover, and efficiency: Selection on productivity or profitability?," *American Economic Review*, 2008, *98* (1), 394–425.

_ , _ , and _ , "The slow growth of new plants: Learning about demand?," *Economica*, 2016, *83*, 91–129.

Furman, Jason and Robert Seamans, "AI and the economy," *Innovation Policy and the Economy*, 2019, *19*, 161–191.

Ganglmair, Bernhard, W. Keith Robinson, and Michael Seeligson, "The Rise of Process Claims: Evidence from a Century of U.S. Patents," *Working paper, University of Mannheim*, 2021.

Garcia-Marin, Alvaro and Nico Voigtländern, "Exporting and plant-level efficiency gains: It's in the measure," *Journal of Political Economy*, 2019, *127* (4), 1777–1825.

Gofman, Michael and Zhao Jin, "Artificial intelligence, human capital, and innovation," *Working Paper*, 2020.

Goldfarb, Avi, Bledi Taska, and Florenta Teodoridis, "Could machine learning be a general-purpose technology? Evidence from online job postings," *Working Paper*, 2019.

Grennan, Jillian and Roni Michaely, "Artificial Intelligence and High-Skilled Work: Evidence from Analysts," *Available at SSRN*, 2020.

Gutiérrez, Germán and Thomas Philippon, "Declining competition and investment in the U.S.," *NBER Working Paper w23583*, 2017.

Hall, Bronwyn H and Adam B Jaffe, *Measuring science, technology, and innovation: A review*, now Publishers Incorporated, 2018.

Haltiwanger, John C, "Synthesizing micro and macro evidence on the US economy," *NBER Reporter*, 2019, (3), 14–17.

Haltiwanger, John, Henry R Hyatt, Erika McEntarfer, Liliana Sousa, and Stephen Tibbets, "Firm age and size in the longitudinal employer-household dynamics data," *US Census Bureau Center for Economic Studies Paper No. CES-WP-14-16*, 2014.

Hauptmann, Alexander G., James Hodson, Juanzi Li, Nicu Sebe, and Achim Rettinger, "Cross-lingual cross-media content linking: Annotations and joint representations.," *Dagstuhl Reports*, 2015, *5.*

Hershbein, Brad and Lisa B. Kahn, "Do recessions accelerate routine-biased technological change? Evidence from vacancy postings," *American Economic Review*, 2018, *108* (7), 1737–72.

Hoberg, Gerald, Gordon Philips, and Nagpurnanand Prabhala, "Product Market Threats, Payouts, and Financial Flexibility," *Journal of Finance*, 2014, *69* (1), 293–324.

Hoberg, Gerard and Gordon Phillips, "Text-based network industries and endogenous product differentiation," *Journal of Political Economy*, 2016, *124* (5), 1423–1465.

Hobijn, Bart and Boyan Jovanovic, "The information-technology revolution and the stock market: Evidence," *American Economic Review*, 2001, *91.*

Hodson, James, "How to make your company machine learning ready," *Harvard Business Review*, 2016.

Hottman, Colin J, Stephen J. Redding, and David E. Weinstein, "Quantifying the sources of firm heterogeneity," *Quarterly Journal of Economics*, 2016, *131* (3), 1291–1364.

Hsu, Po-Hsuan, Dongmei Li, Qin Li, Siew Hong Teoh, and Kevin Tseng, "Valuation of new trademarks," *Management Science*, 2021.

Humlum, Anders, "Robot adoption and labor market dynamics," *Working Paper*, 2019.

Jones, Charles I. and Christopher Tonetti, "Nonrivalry and the economics of data," *American Economic Review*, September 2020, *110* (9), 2819–58.

Juhász, Réka, Mara Squicciarini, and Nico Voigtländer, "Technology adoption and productivity growth: Evidence from industrialization in France," *Working Paper*, 2020.

Klette, Tor Jacob and Samuel Kortum, "Innovating Firms and Aggregate Innovation," *Journal of*

Political Economy, 2004, *112* (5), 986–1018.

Kline, Patrick, Neviana Petkova, Heidi Williams, and Owen Zidar, "Who profits from patents? rent-sharing at innovative firms," *The Quarterly Journal of Economics*, 2019, *134* (3), 1343–1404.

Knight, Will, "The U.S. Leads in Artificial Intelligence, but for How Long?," *MIT Technology Review*, 2017.

Kogan, Leonid, Dimitris Papanikolaou, Lawrence Schmidt, and Bryan Seegmiller, "Technological Change and Occupations over the Long Run," *Available at SSRN 3585676*, 2019.

Loecker, Jan De, Jan Eeckhout, and Gabriel Unger, "The rise of market power and the macroeconomic implications," *Quarterly Journal of Economics*, 2020, *Forthcoming.*

McCarthy, John, Marvin L Minsky, Nathaniel Rochester, and Claude E Shannon, "A proposal for the Dartmouth summer research project on artificial intelligence (1955)," *Reprinted online at http://www-formal. stanford. edu/jmc/history/dartmouth/dartmouth. html*, 1955.

Mihet, Roxana and Thomas. Philippon, "The economics of big data and artificial intelligence," *Disruptive Innovation in Business and Finance in the Digital World (International Finance Review)*, 2019, *20*, 29–43.

Organisation for Economic Co-operation and Development, "Artificial intelligence in society," *OECD Publishing: Paris*, 2019.

Parente, Stephen L. and Edward C. Prescott, "Barriers to technology adoption and development," *Journal of Political Economy*, 1994, *102* (2), 298–321.

Peters, Ryan and Lucien A. Taylor, "Intangible capital and the investment-q relation," *Journal of Financial Economics*, 2017, *123*, 251–272.

Rock, Daniel, "Engineering value: The returns to technological talent and investments in artificial intelligence," *Working paper*, 2019.

Romer, Paul M, "Endogenous technological change," *Journal of political Economy*, 1990, *98* (5, Part 2), S71–S102.

Seamans, Robert and Manav Raj, "AI, labor, productivity and the need for firm-level data," *NBER Working paper w24239*, 2018.

Sinha, Arnab, Zhihong Shen, Yang Song, Hao Ma, Darrin Eide, Bo-June Paul Hsu, and Kuansan Wang, "An Overview of Microsoft Academic Service (MAS) and Applications," in "International World Wide Web Conferences" May 2015.

Stock, James H and Mark W Watson, "Introduction to econometrics (3rd updated edition)," *Age (X3)*, 2015, *3* (0.22).

Syverson, Chad, "Macroeconomics and market power: Facts, potential explanations and open questions," *Brookings Institution, Washington DC*, 2019.

Tambe, P., L. Hitt, D. Rock, and E. Brynjolfsson, "Digital capital and superstar firms," *NBER Working Paper w28285*, 2020.

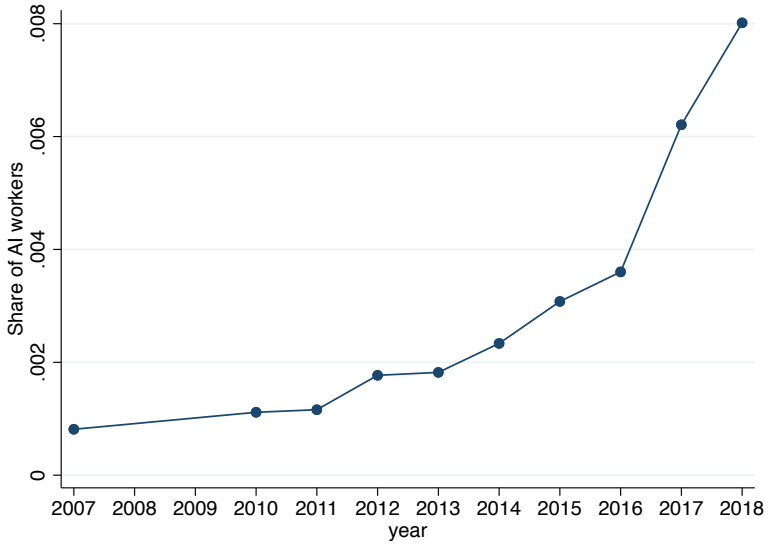
Tanaka, Mari, Nicholas Bloom Bloom, Joel M. David, and Maiko Koga, "Firm performance and macro forecast accuracy.," *Journal of Monetary Economics*, 2019, *forthcoming.*

Tang, Jie, Jing Zhang, Limin Yao, Juanzi Li, Li Zhang, and Zhong Su, "ArnetMiner: Extraction and Mining of Academic Social Networks," in "in" KDD '08 Association for Computing Machinery New York, NY, USA 2008, p. 990â998.

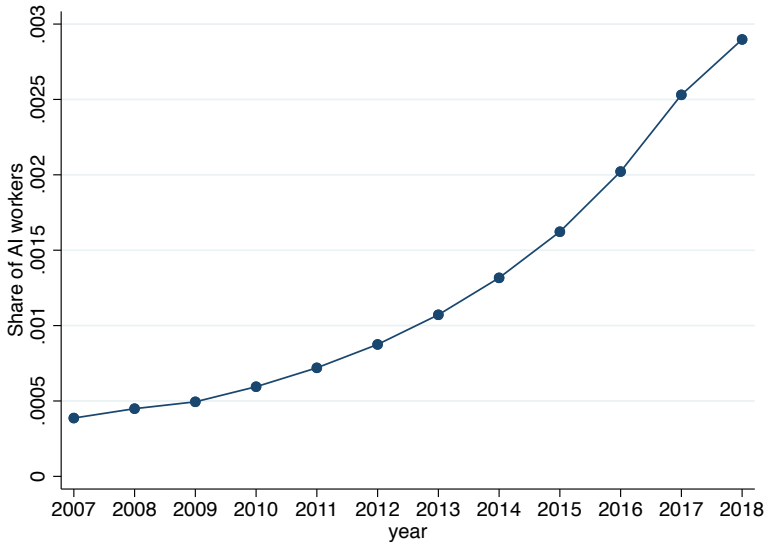
Webb, Michael, "The impact of artificial intelligence on the labor market," *Working Paper*, 2020.

Figure 1. Time Series of AI Investments

This figure shows the time series of the two measures of AI investments. Panel (a) reports the fraction of jobs with continuous AI measure above 0.1 for 2007 and 2010-2018, based on the job postings in Burning Glass with employer firms matched to public firms in Compustat. Panel (b) shows the fraction of all employees (across all public firms) in a given year who are classified as holding directly AI-related positions in the Cognism resume data from 2007 to 2018.



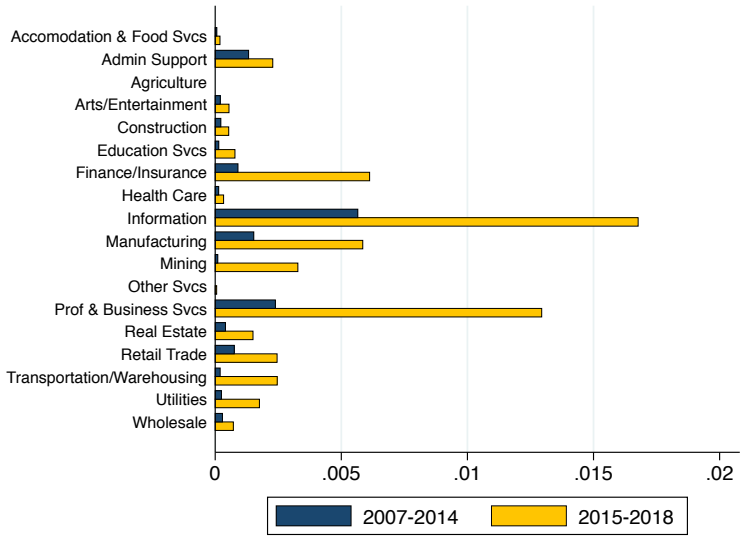
(a) Job posting data (Burning Glass)



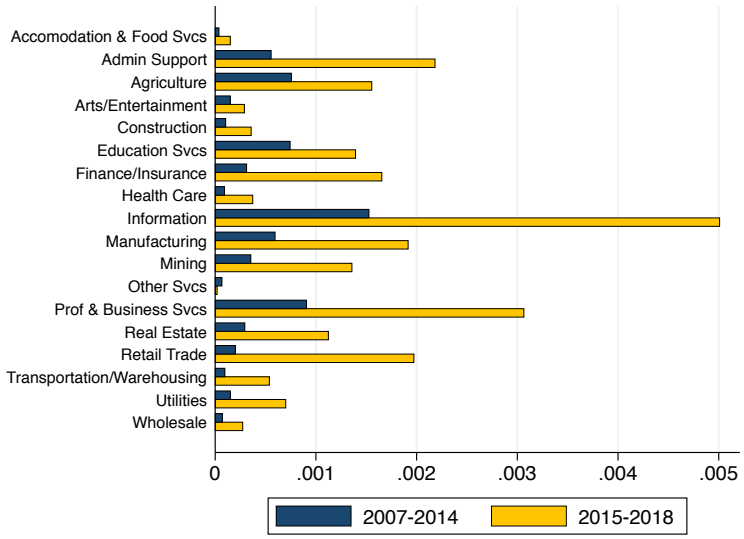
(b) Resume data (Cognism)

Figure 2. AI Investments by Industry Sector

This figure presents the average share of AI jobs at the industry level, based on the sample of public firms. For each sector (based on NAICS-2 digit industry codes), we compute the average share of AI-related job postings (with job-level continuous AI measure above 0.1) across all job postings (in Panel (a)) and the fraction of AI-related employees in the resume data (in Panel (b)). The statistics are computed across all public firms in each sector across two sub-periods: 2007–2014 and 2015–2018.



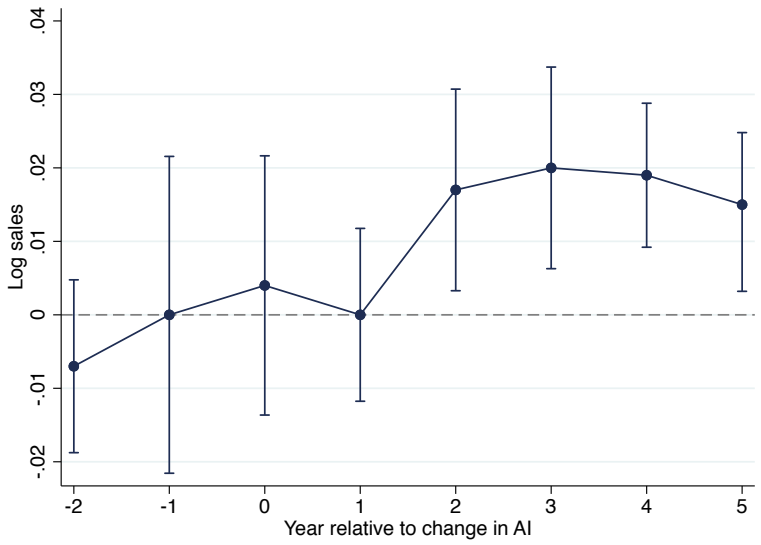
(a) Job posting data (Burning Glass)



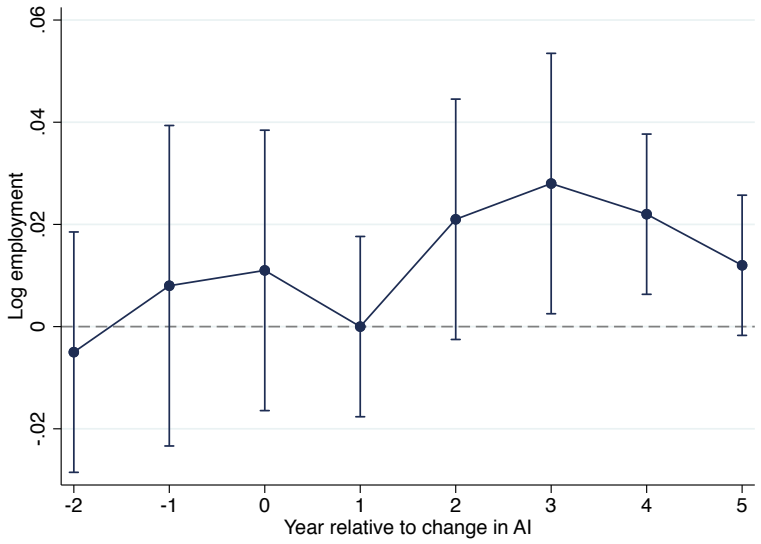
(b) Resume data (Cognism)

Figure 3. AI Investments and Firm Growth Over Time

This figure plots the coefficients from the distributed lead-lag model. The dependent variable is annual log sales in Panel (a) and log employment in Panel (b). The independent variable is the annual change in the share of AI workers in Cognism resume data, standardized to have a mean of zero and a standard deviation of one. Regressions include firm-level sales (or employment) observations between 2010 and 2016 and control for firm fixed effects, 2-digit NAICS industry-by-year fixed effects, and state-by-year fixed effects. Regressions are weighted by the number of workers in Cognism resume data. The vertical bars indicate 95% confidence intervals. Standard errors are clustered at the 5-digit NAICS level.



(a) Sales



(b) Employment

Table 1. Correlations between Job-posting-based and Resume-based AI Measures

This table reports, for each year from 2010 to 2018, the Spearman rank correlations between three pairs of firm-level variables: (i) the absolute number of AI job postings in Burning Glass against the absolute number of AI employees in resumes from Cognism; (ii) the fraction of employees classified as AI-related in the two datasets; and (iii) the fraction of AI employees in Cognism against the average continuous AI measure in Burning Glass. Panel 1 shows raw correlations, and Panel 2 displays correlations conditional on industry sector fixed effects and the baseline controls all measured as of 2010: firm-level characteristics (log sales, cash/assets, R&D/sales, log markup, and log number of jobs—resumes in Panel 1 and job postings in Panel 2), log industry wage, and characteristics of the commuting zones where the firms are located (the share of workers in IT-related occupations, the share of college-educated workers, log average wage, the share of foreign-born workers, the share of routine workers, the share of workers in finance and manufacturing industries, and the share of female workers). All correlations are computed over the cross-section of firms with at least 20 total employees in the Cognism resume data in each year of the sample.

Panel 1: Raw Correlations

	Correlations between:		
Year	Numbers of AI jobs	Fractions of AI Jobs	Cognism fraction & BG continuous measure
2010	0.320	0.272	0.374
2011	0.341	0.288	0.390
2012	0.338	0.291	0.388
2013	0.424	0.363	0.447
2014	0.468	0.410	0.484
2015	0.474	0.405	0.496
2016	0.503	0.421	0.499
2017	0.564	0.474	0.531
2018	0.574	0.484	0.538

Panel 2: Correlations Conditional on Baseline Controls

	Correlations between:		
Year	Numbers of AI jobs	Fractions of AI Jobs	Cognism fraction & BG continuous measure
2010	0.825	0.650	0.470
2011	0.822	0.622	0.476
2012	0.801	0.583	0.487
2013	0.784	0.569	0.498
2014	0.757	0.526	0.551
2015	0.729	0.467	0.513
2016	0.702	0.475	0.501
2017	0.687	0.507	0.510
2018	0.670	0.502	0.513

Table 2. Firm-level Determinants of AI Investments

This table reports the coefficients from regressions of cross-sectional changes in AI investments by U.S. public firms (in non-tech sectors) from 2010 to 2018 on the following ex-ante firm characteristics measured in 2010: log sales in column 1, cash/assets in column 2, R&D/sales in column 3, revenue TFP in column 4, log markup measured following De Loecker et al. (2020) in column 5, Tobin's Q in column 6, market leverage in column 7, return on assets (ROA) in column 8, and firm age in column 9. In Panel 1, the dependent variable is the growth in the share of AI workers from 2010 to 2018 using the resume data from Cognism. In Panel 2, the dependent variable is the growth in the share of AI workers from 2010 to 2018 using the job postings data from Burning Glass. Regressions are weighted by the number of Cognism resumes in 2010 in Panel 1, and by the number of Burning Glass job postings in 2010 in Panel 2. All specifications control for industry sector fixed effects. The dependent variable is normalized to have a mean of zero and a standard deviation of one. Standard errors are clustered at the 5-digit NAICS industry level and reported in parentheses. *, **, and *** denote statistical significance at the 10%, 5%, and 1% levels, respectively.

Panel 1: AI measure from resume data

	Δ Share of AI Workers, 2010–2018									
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
Log Sales 2010	0.108***									0.151***
	(0.028)									(0.026)
Cash/Assets 2010		3.819***								1.957***
		(1.140)								(0.516)
R&D/Sales 2010			3.607***							2.144**
			(1.225)							(0.840)
Revenue TFP 2010				1.343						-0.407
				(0.995)						(0.256)
Log Markup 2010					0.401*					0.375
					(0.220)					(0.258)
Tobin's Q 2010						0.134***				-0.030
						(0.049)				(0.074)
Market Leverage 2010							-0.873			0.142
							(0.644)			(0.342)
ROA 2010								1.248		1.335
								(0.830)		(0.813)
Firm Age 2010									-0.003	-0.001
									(0.004)	(0.002)
Industry FE	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y
Adj R-Squared	0.154	0.301	0.171	0.148	0.144	0.194	0.127	0.130	0.120	0.367
Observations	1,993	1,993	1,993	1,818	1,992	1,653	1,863	1,972	1,993	1,539

Panel 2: AI measure from job postings data

	Δ Share of AI Workers, 2010–2018									
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
Log Sales 2010	0.159***									0.224***
	(0.034)									(0.048)
Cash/Assets 2010		3.431***								1.414***
		(1.213)								(0.488)
R&D/Sales 2010			2.667**							2.806**
			(1.310)							(1.106)
Revenue TFP 2010				0.549						-0.542
				(0.965)						(0.334)
Log Markup 2010					0.400**					0.403*
					(0.194)					(0.222)
Tobin's Q 2010						0.172*				-0.002
						(0.097)				(0.091)
Market Leverage 2010							-0.921			-0.238
							(0.677)			(0.328)
ROA 2010								2.729**		2.462
								(1.334)		(1.565)
Firm Age 2010									-0.005	-0.003
									(0.004)	(0.002)
Industry FE	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y
Adj R-Squared	0.186	0.264	0.165	0.145	0.163	0.243	0.159	0.171	0.149	0.405
Observations	1,192	1,192	1,192	1,120	1,192	1,013	1,139	1,188	1,192	959

Table 3. AI Investments and Firm Growth: Long-differences Estimates

This table reports the coefficients from long-differences regressions of firm growth from 2010 to 2018 on the contemporaneous firm-level changes in AI investments among U.S. public firms (in non-tech sectors). We consider three measures of firm growth: changes in log sales (columns 1 and 2), changes in log employment (columns 3 and 4), and changes in log market value (columns 5 and 6). The dependent variables are measured as growth from 2010 to 2018. The main independent variable is the growth in the share of AI workers from 2010 to 2018, standardized to mean zero and standard deviation of one. Panel 1 considers the resume-based measure of the share of AI workers, while Panel 2 looks at the job-postings-based measure. Regressions are weighted by the number of Cognism resumes in 2010 in Panel 1 and the number of Burning Glass job postings in 2010 in Panel 2. All specifications control for industry sector fixed effects. Columns 2, 4, and 6 also include the baseline controls all measured as of 2010: firm-level characteristics (log sales, cash/assets, R&D/sales, log markup, and log number of jobs—resumes in Panel 1 and job postings in Panel 2), log industry wage, and characteristics of the commuting zones where the firms are located (the share of workers in IT-related occupations, the share of college-educated workers, log average wage, the share of foreign-born workers, the share of routine workers, the share of workers in finance and manufacturing industries, and the share of female workers). Standard errors are clustered at the 5-digit NAICS industry level and reported in parentheses. *, **, and *** denote statistical significance at the 10%, 5%, and 1% levels, respectively.

Panel 1: AI measure from resume data

	Δ Log Sales		Δ Log Employment		Δ Log Market Value	
	(1)	(2)	(3)	(4)	(5)	(6)
Δ Share AI Workers	0.202***	0.203***	0.239**	0.219***	0.231**	0.224***
	(0.069)	(0.060)	(0.097)	(0.077)	(0.094)	(0.077)
Industry FE	Y	Y	Y	Y	Y	Y
Controls	N	Y	N	Y	N	Y
Adj R-Squared	0.221	0.428	0.237	0.418	0.221	0.364
Observations	1,052	1,052	1,052	1,052	1,010	1,010

Panel 2: AI measure from job postings data

	Δ Log Sales		Δ Log Employment		Δ Log Market Value	
	(1)	(2)	(3)	(4)	(5)	(6)
Δ Share AI Workers	0.150***	0.160***	0.161**	0.127**	0.146*	0.189***
	(0.055)	(0.045)	(0.076)	(0.050)	(0.083)	(0.068)
Industry FE	Y	Y	Y	Y	Y	Y
Controls	N	Y	N	Y	N	Y
Adj R-Squared	0.273	0.452	0.320	0.526	0.338	0.461
Observations	935	935	935	935	903	903

Table 4. Heterogeneous Relationship between AI Investments and Firm Growth by Initial Firm Size

This table reports the coefficients from long-differences regressions of firm growth from 2010 to 2018 on contemporaneous changes in AI investments among US public firms (in non-tech sectors), separately for each tercile of initial firm size. Firms in each 2-digit NAICS sector are divided into terciles based on employment in 2010. We consider three measures of firm-level growth for the dependent variable: changes in log sales (columns 1 and 2), changes in log employment (columns 3 and 4), and changes in log market value (columns 5 and 6). The main independent variable is the growth in the share of AI workers from 2010 to 2018, standardized to mean zero and standard deviation of one. Panel 1 considers the resume-based measure of AI workers, while Panel 2 looks at the job-posting-based measure. Regressions are weighted by the number of Cognism resumes in 2010 in Panel 1 and the number of Burning Glass job postings in 2010 in Panel 2. All specifications control for industry sector by initial firm size tercile fixed effects. Columns 2, 4, and 6 also include the baseline controls all measured as of 2010: firm-level characteristics (log sales, cash/assets, R&D/sales, log markup, and log number of jobs—resumes in Panel 1 and job postings in Panel 2), log industry wage, and characteristics of the commuting zones where the firms are located (the share of workers in IT-related occupations, the share of college-educated workers, log average wage, the share of foreign-born workers, the share of routine workers, the share of workers in finance and manufacturing industries, and the share of female workers). Standard errors are clustered at the 5-digit NAICS industry level and reported in parentheses. *, **, and *** denote statistical significance at the 10%, 5%, and 1% levels, respectively.

Panel 1: AI measure from resume data

	Δ Log Sales		Δ Log Employment		Δ Log Market Value	
	(1)	(2)	(3)	(4)	(5)	(6)
Δ Share AI Workers*Size Tercile 1	0.046**	0.003	0.041**	-0.012	0.059**	0.021
	(0.023)	(0.020)	(0.018)	(0.027)	(0.028)	(0.041)
Δ Share AI Workers*Size Tercile 2	0.219***	0.186***	0.217***	0.177***	0.202***	0.171***
	(0.054)	(0.051)	(0.048)	(0.060)	(0.045)	(0.054)
Δ Share AI Workers*Size Tercile 3	0.223***	0.213***	0.260**	0.227***	0.250**	0.235***
	(0.077)	(0.066)	(0.105)	(0.081)	(0.103)	(0.087)
Industry*Size tercile FE	Y	Y	Y	Y	Y	Y
Controls	N	Y	N	Y	N	Y
Adj R-Squared	0.248	0.427	0.253	0.430	0.214	0.342
Observations	1,044	1,044	1,044	1,044	1,003	1,003
T-test statistic	3.7	8.8	3.7	7.4	4.0	8.5
T-test p value	0.054	0.003	0.054	0.007	0.047	0.004

Panel 2: AI measure from job postings data

	Δ Log Sales		Δ Log Employment		Δ Log Market Value	
	(1)	(2)	(3)	(4)	(5)	(6)
Δ Share AI Workers*Size Tercile 1	-0.068***	-0.079***	-0.056***	-0.109**	-0.081***	-0.075*
	(0.021)	(0.029)	(0.015)	(0.049)	(0.029)	(0.039)
Δ Share AI Workers*Size Tercile 2	0.040	0.061	0.040	0.035	0.006	0.087*
	(0.052)	(0.038)	(0.043)	(0.050)	(0.040)	(0.052)
Δ Share AI Workers*Size Tercile 3	0.166***	0.168***	0.176**	0.130**	0.165*	0.204***
	(0.055)	(0.048)	(0.078)	(0.051)	(0.085)	(0.073)
Industry*Size tercile FE	Y	Y	Y	Y	Y	Y
Controls	N	Y	N	Y	N	Y
Adj R-Squared	0.292	0.458	0.324	0.534	0.334	0.465
Observations	927	927	927	927	896	896
T-test statistic	19.2	22.8	8.1	9.6	7.3	11.3
T-test p value	0.000	0.000	0.005	0.002	0.007	0.001

Table 5. AI Investments and Firm Growth: IV Estimates Using Resume Data

This table estimates the relationship between AI investments and firm growth from 2010 to 2018 for U.S. public firms (in non-tech sectors), where firm AI investments are instrumented with firm-level ex-ante exposure to AI-skilled graduates from AI-strong universities. The independent variable is the change in the share of AI workers from 2010 to 2018 based on the resume data. Regressions are weighted by the number of Cognism resumes in 2010. The independent variable and the instrument are standardized to mean zero and standard deviation of one. We consider changes in log sales in columns 1 to 4, log employment in columns 5 to 8, and log market value in columns 9 to 12. All specifications control for industry sector fixed effects and ex-ante exposure to universities that are strong in computer science research. Columns 2–4, 6–8, and 10–12 also control for the baseline controls all measured as of 2010: firm-level characteristics (log sales, cash/assets, R&D/sales, log markup, and log number of jobs—resumes in Panel 1 and job postings in Panel 2), log industry wage, and characteristics of the commuting zones where the firms are located (the share of workers in IT-related occupations, the share of college-educated workers, log average wage, the share of foreign-born workers, the share of routine workers, the share of workers in finance and manufacturing industries, and the share of female workers). Columns 3–4, 7–8, and 11–12 additionally control for firm-level changes in log sales and log employment from 2000 to 2008. Columns 4, 8, and 12 add state fixed effects. Standard errors are clustered at the 5-digit NAICS industry level and reported in parentheses. The first-stage F-statistics of the instrument are reported for all specifications. *, **, and *** denote statistical significance at the 10%, 5%, and 1% levels, respectively.

	Δ Log Sales				Δ Log Employment				Δ Log Market Value			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
Δ Share AI Workers	0.311***	0.446***	0.407***	0.261*	0.426***	0.690***	0.556***	0.270*	0.345**	0.391**	0.319**	0.180
	(0.097)	(0.123)	(0.127)	(0.149)	(0.140)	(0.214)	(0.177)	(0.160)	(0.136)	(0.162)	(0.161)	(0.185)
Industry FE	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y
CS Control	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y
Baseline Controls	N	Y	Y	Y	N	Y	Y	Y	N	Y	Y	Y
Control Pre-trend	N	N	Y	Y	N	N	Y	Y	N	N	Y	Y
State FE	N	N	N	Y	N	N	N	Y	N	N	N	Y
F Statistic	14.7	19.4	20.4	29.2	14.7	19.4	20.4	29.2	14.8	19.6	20.3	30.2
Observations	1,001	1,001	777	773	1,001	1,001	777	773	963	963	753	746

Table 6. AI Investments and Product Innovation

This table reports the coefficients from long-differences regressions of the changes in measures of product innovation from 2010 to 2018 on the contemporaneous changes in AI investments by U.S. public firms (in non-tech sectors). The dependent variables are the change in log(1+number of trademarks) in columns 1 and 2; the change in log(1+number of product patents) in columns 3 and 4; and the change in the product mix in columns 5 and 6. Product patents are patents with over 50% of the claims being product claims, following the categorization in Ganglmair et al. (2021). The change in the product mix is measured as the sum of annual changes from 2010 to 2018, where each annual change is the angle between the two word vectors indicating firms' product offerings in that year and the previous year (the word vectors are constructed as in Hoberg et al. (2014). For the main independent variable, Panel 1 considers the resume-based measure of the growth in the share of AI workers from 2010 to 2018, while Panel 2 looks at the job-posting-based measure. Both measures are standardized to mean zero and standard deviation of one. Regressions are weighted by the number of Cognism resumes in 2010 in Panel 1 and the number of Burning Glass job postings in 2010 in Panel 2. All specifications control for industry sector fixed effects. Columns 2, 4, and 6 also include the baseline controls all measured as of 2010: firm-level characteristics (log sales, cash/assets, R&D/sales, log markup, and log number of jobs—resumes in Panel 1 and job postings in Panel 2), log industry wage, and characteristics of the commuting zones where the firms are located (the share of workers in IT-related occupations, the share of college-educated workers, log average wage, the share of foreign-born workers, the share of routine workers, the share of workers in finance and manufacturing industries, and the share of female workers). Standard errors are clustered at the 5-digit NAICS industry level and reported in parentheses. *, **, and *** denote statistical significance at the 10%, 5%, and 1% levels, respectively.

Panel 1: AI measure from resume data

	Δ Log Number of Trademarks		Δ Log Number of Product Patents		Change in Product Mix	
	(1)	(2)	(3)	(4)	(5)	(6)
Δ Share AI Workers	0.144**	0.150*	0.221***	0.212***	0.149***	0.114***
	(0.065)	(0.077)	(0.035)	(0.037)	(0.036)	(0.035)
Industry FE	Y	Y	Y	Y	Y	Y
Controls	N	Y	N	Y	N	Y
Observations	550	550	619	619	958	958

Panel 2: AI measure from job postings data

	Δ Log Number of Trademarks		Δ Log Number of Product Patents		Change in Product Mix	
	(1)	(2)	(3)	(4)	(5)	(6)
Δ Share AI Workers	0.132**	0.059	0.169***	0.198***	0.109***	0.074**
	(0.057)	(0.063)	(0.029)	(0.050)	(0.033)	(0.036)
Industry FE	Y	Y	Y	Y	Y	Y
Controls	N	Y	N	Y	N	Y
Observations	505	505	560	560	860	860

Table 7. AI Investments and Operating Costs

This table reports the coefficients from long-differences regressions of changes in firm operating costs and firm productivity from 2010 to 2018 on contemporaneous changes in AI investments by U.S. public firms (in non-tech sectors). The main independent variable is the change in the share of AI workers from 2010 to 2018, standardized to mean zero and standard deviation of one. We look at two measures of operating costs: log COGS in columns 1 and 2 and log operating expenses in columns 3 and 4. We consider two measures of productivity: log sales per worker (columns 5–6) and revenue TFP (columns 7–8). Revenue TFP is the residual from regressing log revenue on log employment and log capital (constructed using the perpetual inventory method), with separate regressions for each industry sector. In columns 9 and 10, the dependent variable is the change in log(1+number of process patents), where process patents are patents with over 50% of the claims being process claims, following the categorization in Ganglmair et al. (2021). Panel 1 considers the resume-based measure of the share of AI workers, while Panel 2 looks at the job-posting-based measure. Regressions are weighted by the number of Cognism resumes in 2010 in Panel 1 and the number of Burning Glass job postings in 2010 in Panel 2. All specifications control for industry sector fixed effects. Columns 2, 4, 6, and 8 also include the baseline controls all measured as of 2010: firm-level characteristics (log sales, cash/assets, R&D/sales, log markup, and log number of jobs—resumes in Panel 1 and job postings in Panel 2), log industry wage, and characteristics of the commuting zones where the firms are located (the share of workers in IT-related occupations, the share of college-educated workers, log average wage, the share of foreign-born workers, the share of routine workers, the share of workers in finance and manufacturing industries, and the share of female workers). Standard errors are clustered at the 5-digit NAICS industry level and reported in parentheses. *, **, and *** denote statistical significance at the 10%, 5%, and 1% levels, respectively.

Panel 1: AI measure from resume data

	Δ Log COGS		Δ Log Operating Expense		Δ Log Sales per Worker		Δ Revenue TFP		Δ Log Number of Process Patents	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
Δ Share AI Workers	0.195***	0.182***	0.206***	0.204***	-0.082	-0.058	-0.049	-0.026	-0.010	-0.024
	(0.052)	(0.046)	(0.066)	(0.056)	(0.055)	(0.040)	(0.046)	(0.037)	(0.039)	(0.066)
Industry FE	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y
Controls	N	Y	N	Y	N	Y	N	Y	N	Y
Adj R-Squared	0.213	0.393	0.237	0.427	0.244	0.391	0.222	0.347	0.700	0.766
Observations	1,052	1,052	1,052	1,052	1,052	1,052	977	977	619	619

Panel 2: AI measure from job postings data

	Δ Log COGS		Δ Log Operating Expense		Δ Log Sales per Worker		Δ Revenue TFP		Δ Log Number of Process Patents	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
Δ Share AI Workers	0.153***	0.143***	0.148***	0.151***	-0.039	-0.018	-0.018	-0.012	0.031	0.064
	(0.039)	(0.035)	(0.051)	(0.040)	(0.050)	(0.031)	(0.042)	(0.031)	(0.048)	(0.064)
Industry FE	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y
Controls	N	Y	N	Y	N	Y	N	Y	N	Y
Adj R-Squared	0.246	0.390	0.251	0.410	0.375	0.555	0.287	0.422	0.732	0.760
Observations	935	935	935	935	935	935	874	874	560	560

Table 8. AI Investments and Changes in Industry Growth and Concentration

This table reports the coefficients from industry-level long-differences regressions of the changes in industry sales, employment, and concentration on contemporaneous changes in industry-level AI investments. All industry-level variables are calculated for all firms in Compustat (regardless of whether they are in our main regression sample in Table 3 or not). Each observation is a 5-digit NAICS industry, and (as in our main analysis) we exclude tech sectors. The independent variable is the change in the share of AI workers from 2010 to 2018, standardized to mean zero and standard deviation of one. Panel 1 considers the resume-based measure of the share of AI workers, while Panel 2 looks at the job-posting-based measure. Regressions are weighted by the total (industry-level) number of Cognism resumes in 2010 in Panel 1 and the total (industry-level) number of Burning Glass job postings in 2010 in Panel 2. The dependent variables are the changes, from 2010 to 2018, in log total sales in columns 1 and 2, log total employment in columns 3 and 4, the Herfindahl-Hirschman Index (HHI) in columns 5 and 6, and the market share of the top firm in an industry in columns 7 and 8. All specifications control for industry sector fixed effects. Regressions in columns 2, 4, 6, and 8 also include industry-level controls for log total employment, log total sales, and log average wage in 2010. Standard errors are robust against heteroskedasticity and reported in parentheses. *, **, and *** denote statistical significance at the 10%, 5%, and 1% levels, respectively.

Panel 1: AI measure from resume data

	Δ Log Sales		Δ Log Employment		Δ HHI		Δ Top Firm Market Share	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Δ Share AI Workers	0.173***	0.199***	0.205***	0.234***	0.019***	0.011	0.023***	0.013*
	(0.045)	(0.038)	(0.063)	(0.045)	(0.007)	(0.007)	(0.007)	(0.008)
Industry Sector FE	Y	Y	Y	Y	Y	Y	Y	Y
Controls	N	Y	N	Y	N	Y	N	Y
Observations	275	275	275	275	267	267	267	267

Panel 2: AI measure from job postings data

	Δ Log Sales		Δ Log Employment		Δ HHI		Δ Top Firm Market Share	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Δ Share AI Workers	0.159***	0.159***	0.158	0.139	0.036***	0.041***	0.040***	0.047***
	(0.058)	(0.058)	(0.106)	(0.097)	(0.006)	(0.006)	(0.006)	(0.007)
Industry Sector FE	Y	Y	Y	Y	Y	Y	Y	Y
Controls	N	Y	N	Y	N	Y	N	Y
Observations	261	261	261	261	254	254	254	254

A Appendix on Instrument Construction

We instrument firm-level AI investments using variation in firms' ex-ante exposure to the supply of AI-trained graduates from universities that are historically strong in AI. The core idea is that the scarcity of AI-trained labor is one of the most important barriers to firms' AI adoption (e.g., CorrelationOne (2019)). Universities are a key source of skilled labor, and universities historically strong in AI research are able to train more AI-skilled graduates following the wide-spread rise of commercial interest in AI in the 2010s. This enables firms with more ex-ante connections to AI-strong universities (e.g., via alumni networks) to more readily attract AI talent from those universities in the 2010s. It is important to note that while AI research flourished in universities long before 2010 (research in AI and machine learning goes back to the 1950s), commercial interest in AI applications started around 2012, driven by rapid accumulation of data, decreasing costs of computation, and methodological advances in applying techniques such as deep learning.[40] Moreover, universities did not set up specialized data science programs until the mid-2010s. For example, Columbia's Data Science Institute (described as a "*trailblazer in the field*"; see here) was established in 2012. Therefore, in 2010, firms' connections to AI-strong universities were not driven by the need to hire AI-skilled workers, but rather by other pre-existing connections such as alumni networks (e.g., the CEO having graduated from a particular university), especially for the sample of non-tech firms that are the focus of this paper.

To construct the instrument, we need two different datasets. The first is a measure of the strength of AI research in each university in the pre-period. The second, even more difficult to construct, is a measure of firm-university hiring networks in the pre-period. To the best of our knowledge, there is no comprehensive historical data on either of these two aspects. To construct the first measure, we group all universities into those that are ex-ante strong in AI research and those that are not, based on the number of researchers producing AI-related publications in each university before 2010. A key concern with this measure for our instrument is that AI-strong universities are likely to also be strong in the broader field of computer science (CS), producing more CS-skilled graduates, which might affect firm outcomes through channels other than AI investments. To address this concern, we also collect information on the number of CS researchers in each university in each year to be included as a control. To construct the second measure (firm-university hiring networks), we leverage our resume data to observe which universities the stock of a given firm's employees as of 2010 graduated from. To validate the data, we also measure: (i) the number of fresh graduates in each year from each university hired by each firm to confirm that ex-ante firm-university networks predict ex-post hiring, and (ii) the number of AI-trained graduates from each university to validate our premise that ex-ante AI-strong universities produce

[40] A brief history of AI research can be found here.

more AI-trained graduates following the increase in commercial interest in AI.

Data Construction. First, to identify universities that are ex-ante strong in AI, we use data from the Open Academic Graph (OAGv2) to measure AI-related publications associated with each university. OAGv2 provides a unified view of two large-scale databases of academic paper metadata, abstracts, citations, and author links: (i) the Microsoft Academic Graph (part of the Microsoft Academic Service infrastructure in Sinha et al. (2015)), and (ii) ArnetMiner (Tang et al., 2008). Together, these two datasets provide the most comprehensive openly available repository of scholarly work starting from the 1870s and allow us to track research articles and faculty across the near-universe of academic and commercial institutions. The Open Academic Graph contains hundreds of millions of papers from 366M distinct author names and lists author affiliations where available. We use a keyword-based matching procedure to link 689 research institutions (or 99%) in the Higher Education Research and Development Survey (HERDS) data to faculty information in the OAGv2. HERDS data are collected by the National Science Foundation and cover all universities in the U.S. that have at least $150,000 in R&D expenditures in each fiscal year. Our strict matching procedure requires that the full formal university name, or an official shortened variant thereof, be found in full form within the institutional affiliations in the OAGv2 paper metadata files, with only common "stop-words" (such as “and,” “the,” and “in”) removed from both sides of the match. A manual review of the resulting linked data shows over 96% precision in matching author affiliations from the Open Academic Graph to HERDS data, with the remaining incorrect entries manually adjusted to ensure full correctness. For each university matched to HERDS, we consider all publications in the Open Academic Graph in each year that have at least one co-author affiliated with that university.

We work with the field experts at the AI for Good Foundation to identify AI-related publications.[41] First, we identify a small set of “seed” journals and conference proceedings that explicitly include terms like “artificial intelligence” and “machine learning” in their title (e.g., *Journal of Machine Learning Research* and *Proceedings of the International Joint Conference on Artificial Intelligence*). Second, to identify potential additional AI-related journals and conference proceedings, we look at all other journals and proceedings that have published work by the authors of the papers in the seed journals and proceedings. We manually filter this broader set of journals and conference proceedings to the ones that focus predominantly on AI, leading to a final list of 355 journals and conference proceedings globally.

To make sure that our results are not driven by firms’ exposure broader (non-AI) CS-skilled workers, we control for firms’ ex-ante exposure, via their hiring networks, to CS-strong universities. In particular, we construct an analogous measure of computer science publications by start-

[41] Learn more about the AI for Good Foundation here.

ing with a set of seed journals and conference proceedings across different fields of computer science (those with the terms "compilers," "databases," "cryptography," "computation," "software," "programming," "informatics," "robotics," or "information security" in their titles) and then manually screen all other journals and conference proceedings that publish papers by the same authors. We exclude any journal or conference proceeding that we classify as AI-related, leaving a total of 796 non-AI computer science journals and conference proceedings.

After identifying the set of AI-related and CS-related journals and conference proceedings, we classify the focus area of each researcher r as either AI, computer science, or neither. If at least one third of all publications co-authored by r are in either AI or CS journals and conference proceedings, then r is considered a candidate researcher. If r is a candidate researcher *and* at least half of r's AI/CS publications are in specifically AI journals and proceedings, then r is marked as an AI researcher. If more than half of r's AI/CS publications are in non-AI computer science journals and proceedings, then r is considered a non-AI CS researcher. Finally, if more than two thirds of r's publications are outside of the set of identified AI and CS journals and proceedings, then r is classified as a researcher in other (unrelated) fields.

At the university level, we compute the percentage of researchers in each year who are classified as AI researchers and the percentage of researchers who are classified as CS researchers. Researchers in other unrelated fields are included in the denominators of both measures. To reduce noise, we assume that each researcher is employed at the respective university in a non-publishing year if that researcher is employed at that university in both the following and the preceding year. For example, if researcher r is identified as affiliated with university u in both 2005 and 2007 but has no publications in 2006, then r is still considered to be affiliated with university u in year 2006. We then classify whether each university is AI-strong. We define a university as being strong in AI if it satisfies one of the following two criteria in at least one year between 2005 and 2009: (i) the number of AI researchers is in the top 5% of the distribution across all universities in a given year; (ii) the number of AI researchers is in the top 10% of the distribution, and the share of AI researchers (the number of AI researchers divided by the number of other researchers in the OAGv2 data) is in the top 5% of the distribution across all universities in a given year. We use the second criterion because there are some smaller, tech-oriented colleges that could potentially have a large share of researchers in AI but do not necessarily have large departments. Our results are robust to using other cutoffs and earlier years.

We verify that the OAGv2 publication data provide a reliable measure of university research. In Figure 4, we plot the log number of (all) researchers in each university in the OAGv2 data against the log R&D expenditure in the HERDS data. We find a strong positive correlation of 0.83. Furthermore, the top universities we identify as AI-strong include top AI departments, such as

Carnegie Mellon University, UCLA, Stanford University, UIUC, New York University, and University of Maryland College Park, but are not strongly correlated with the overall highest-ranked universities based on the U.S. World & News Report. For example, only 50% (39%) of the top 20 (top 50) universities are AI-strong universities, and among AI-strong universities, only 25% (56%) are ranked in the top 20 (top 50) universities in the U.S. World & News Report.

To construct the second ingredient for our instrument—firm exposure to AI-strong universities via the ex-ante firm-university hiring networks—we use our Cognism resume data. In these data, we observe the granting institutions of all degrees that workers list on their resumes. We disambiguate university names and match them to HERDS data. We define an individual i as a graduate of university u if i's resume lists at least one degree (undergraduate or graduate) from university u. We define an individual i as a *fresh graduate* from university u in year t if i joined a firm in year t and graduated from university u in year t or year $t-1$. These data offer comprehensive coverage of universities; for example, in 2010 there 668 of the 716 universities in the HERDS dataset have at least one fresh graduate in our resume data. Since the firms' hiring patterns might be different for STEM versus non-STEM workers (e.g., if a firm has a hiring relationship with an economics department for economic policy talent and with a business school for management talent), we also consider the firm-university hiring networks based specifically on STEM workers, in case such networks are more relevant for hiring AI workers. We define STEM workers as employees who have at least one degree with a major in either engineering (e.g., electrical, chemical, mechanical), physical sciences (e.g., math, physics, chemistry, computer science, statistics), or biological sciences (e.g., biology, pharmacology).

We compare the coverage of our university graduates data with official statistics from universities and show that our resume data cover a sizable proportion of university graduates in the U.S. In particular, we aggregate the data to university-year level by calculating the total number of fresh graduates from each university in each year. We compare these numbers with the total numbers of all degrees (bachelors, masters, and PhDs) conferred by each university in each year, using the Integrated Postsecondary Education Data System (IPEDS) data, which contain the total enrollment and the number of degrees conferred each year for all post-secondary institutions in the U.S. As of 2012 (the latest year of the IPEDS data), our resume data cover, on average, 59% of all fresh graduates at each university. The number of fresh graduates in the resume data is also highly correlated with graduates in the cross-section of universities (correlation=0.73).

Finally, we use our Cognism resume data to measure the share of all fresh university graduates from each university who get AI-skilled jobs in each year between 2006 and 2018. These data allow us to validate our premise that ex-ante AI-strong universities are able to increase the supply of AI-skilled graduates following the increase in commercial applications in AI in the first half of the

2010s, discussed below.

Instrument Validation. We first validate several core assumptions underlying the intuition behind our instrument. Confirming our key argument, we show that the increase in AI-trained graduates during the 2010s was much more pronounced in AI-strong universities than in non-AI-strong universities. Figure 5 plots the share of fresh graduates that are AI-trained from AI-strong and non-AI-strong universities from 2006 to 2018. In 2006, there were few AI graduates across the board, with the share of AI graduates below 0.3% for both AI-strong and non-AI-strong universities. Even in 2012, the share of AI graduates remained below 0.5% in both groups of universities. From 2012 to 2018, however, the share of AI graduates tripled (to about 1.5%) in AI-strong universities, while the share of AI graduates remained under 0.5% in non-AI-strong universities.

We then examine whether firm-university hiring networks provide the necessary variation for our instrumental variable strategy. First, our instrument leverages the variation in exposure to AI-strong universities across firms. Therefore, it requires that firms do not hire uniformly from the same universities. Empirically, most firms in our data concentrate their hiring in a small number of universities. On average, a firm hires 18% of its fresh graduates from the single main university in its network, 44% from its five main universities, and 59% from its 10 main universities. By contrast, the largest university produces only 1.6% of all fresh graduates, the largest five universities produce 7.1% of all fresh graduates, and the largest 10 universities produce 12.9% of all fresh graduates. Firms also hire disproportionately from universities located in the same state as their headquarters: on average, firms hire 38% of all fresh graduates, 37% of STEM fresh graduates, and 42% of AI fresh graduates from universities located in the same state. Second, in order for the ex-ante firm-university network to predict ex-post hiring of AI-skilled labor, firm-university networks need to be persistent over time. In column 1 of Table 9, we regress the share of fresh graduates hired from each university after 2010 on the share of fresh graduates hired from each university before 2010. We find a strong positive relationship, suggesting that firm-university networks are correlated over time. In column 2, we use the share of all workers employed in a firm in 2010 who graduated from each university to predict the share of fresh graduates hired from that university after 2010, again finding a strong positive correlation. The persistence of firm-university hiring networks also manifests in AI hiring. Columns 3 and 4 show that the universities from which a firm hired before 2010 also strongly predict the universities from which the firm will hire its AI-skilled workers after 2010. Finally, in columns 5 and 6, we show that pre-2010 firm-university hiring networks based only on STEM workers also strongly predict the universities from which firms hire their AI workers after 2010.

Our instrument is defined as follows for each firm i:

$$IV_i = \sum_u s_{iu}^{2010} AIstrong_u,$$

where s_{iu}^{2010} is the share of STEM workers in firm i in 2010 who graduated from university u, and $AIstrong_u$ equals one if university u is identified as an AI-strong university based on pre-2010 publications. We use firm-university hiring networks based on STEM workers in the firm as of 2010, because the instrument based on this measure has a stronger first stage; however, the results are very similar when we construct firm-university hiring networks using all workers in the firm as of 2010. To reduce noise, for each firm's hiring network, we consider the 50 universities from which the firm has the most workers in 2010. To control for the effects of general computer science (and not specifically AI), we construct an analogous measure of firms' exposure to CS-strong universities: $\sum_u s_{iu}^{2010} CSstrong_u$, where the weights s_{iu}^{2010} are firms' 2010 STEM hiring shares, and $CSstrong_u$ is the average share of (non-AI) CS researchers (the number of CS researchers divided by the number of all other researchers) at university u between 2005 and 2009.

Before proceeding, we examine an important identification concern regarding our instrument: if firms anticipated the surge in demand for AI, they might have started building their connections to AI-strong universities before 2010, making firm-university hiring networks in 2010 endogenous to firms' ability to hire AI-trained students ex-post. This is unlikely, given the lack of both commercial interest in AI by firms and training of AI-skilled graduates by universities (Figure 5) prior to 2010. Indeed, we are able to confirm empirically that firms connected to AI-strong universities did not increase their share of hired fresh graduates from those universities from 2005 to 2010. Specifically, in Table 10, we find no significant relationship between the change in the share of fresh graduates from AI-strong universities in the pre-period (from 2005 to 2010) and our instrument.

First Stage. Table 11 presents the first stage of the instrument, where we regress our key independent variable—firm-level changes in the share of AI-skilled workers from 2010 to 2018—on the instrument, which measures ex-ante firm-level exposure to the supply of AI-trained university graduates from AI-strong universities. We control for firm-specific ex-ante exposure to CS-strong universities and industry fixed effects in all specifications. In column 2, we additionally control for our baseline controls measured as of 2010: (i) firm-level variables (log employment, cash/assets, log sales, R&D/Sales, and log markups), (ii) the characteristics of the commuting zones where the firms are located in 2010 (the share of workers in IT-related occupations, the share of college-educated workers, log average wage, the share of foreign-born workers, the share of routine workers, the share of workers in finance and manufacturing industries, and the share of female workers); and (c) the log industry-average wage. The inclusion of these controls helps

to address the concern that firms' ex-ante exposure to AI-strong universities might be correlated with other firm characteristics that can drive AI adoption and firm growth. In column 3, we also control for firms' pre-period sales and employment growth between 2000 and 2008 to address unobservable firm characteristics that might simultaneously drive firms' growth trajectories and their hiring of AI workers. In column 4, we further add state fixed effects to control for local labor market characteristics that might drive both universities' ability to produce AI graduates and firm growth. The first stage F-statistics are well above the conventional level of 10 for all specifications using the Cognism resume data. The F-statistic is also above 10 for two out of four specifications using the Burning Glass job postings data. Intuitively, the first stage is stronger in the Cognism data because: (i) the data generating process is similar for the firm-level AI investments and the firm-university hiring networks (they are both based on our resume data), whereas job-postings-based AI investments are measured using different data than the instrument, and (ii) our instrument captures the *supply* of AI-skilled labor to firms, which is likely to correlate more strongly with firms' actual hiring of AI workers in the Cognism resume data than with firms' *demand* for AI workers in the Burning Glass job postings data.

Figure 4. Correlation Between the Number of University Researchers and University R&D Expenditures

This figure is a binned scatterplot of the log number of researchers in each university against the log R&D expenditure in each university in 2010. Each dot represents roughly the same number of universities, and the solid line is the fitted regression line. The number of researchers in each university is the number of authors from that university with at least one publication in the OAGv2 data. The R&D expenditure of each university is from the NSF's HERDS data.

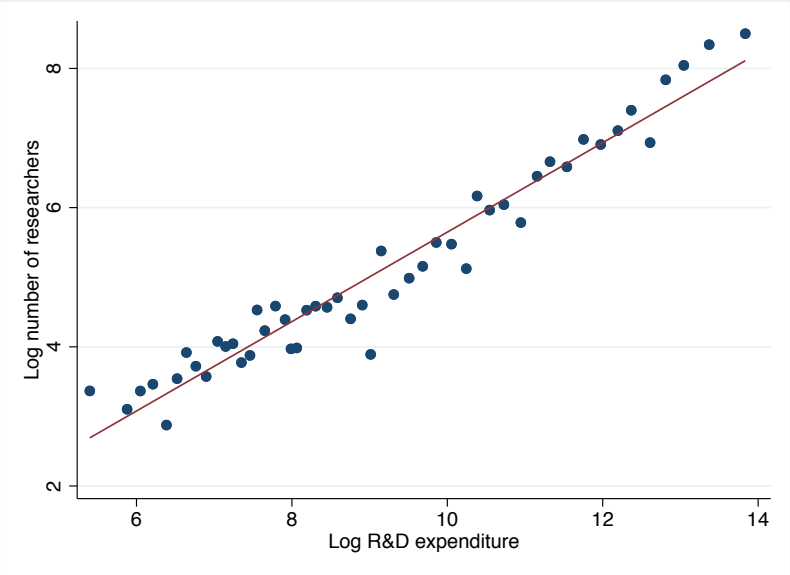


Figure 5. Time Series of the Share of AI-trained Fresh Graduates from Ex-ante AI-strong Universities and Other Universities

This figure plots the average share of AI-trained fresh graduates out of all fresh graduates from 2006 to 2018, separately for ex-ante AI-strong universities and non-AI-strong universities. We define a university as an AI-strong university if it satisfies one of the following two criteria in at least one year between 2005 and 2009: (i) the number of AI researchers is in the top 5% of the distribution across all universities in a given year; (ii) the number of AI researchers is in the top 10% of the distribution, and the share of AI researchers is in the top 5% of the distribution across all universities in a given year. We define an individual i as a fresh graduate from university u in year t if individual i joined a firm in year t and graduated from university u in year t or year $t-1$. An individual is considered an AI-trained fresh graduate in year t if the individual is a fresh graduate in year t and that individual's first job after graduation is an AI-skilled job. AI-skilled jobs are defined based on the methodology described in Section 3.2 and used throughout the paper.

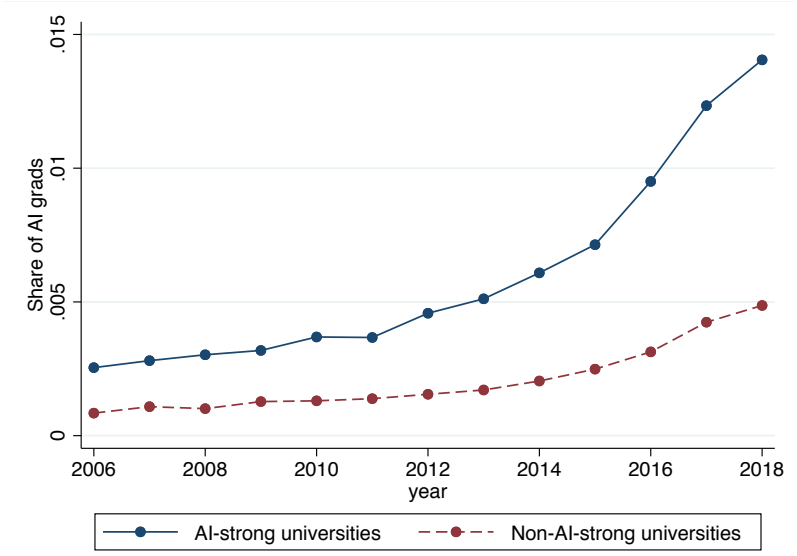


Table 9. Persistence of Firm-University Hiring Networks

This table reports the coefficients from regressing the share of each firm's fresh graduates hired from each university after 2010 on the pre-2010 firm-university network. Each observation is a firm-university pair. The dependent variable, constructed using the Cognism resume data, is the share of all fresh graduates hired from each university after 2010 in columns 1 and 2 and the share of AI-trained fresh graduates hired from each university after 2010 in columns 3–6. In columns 1 and 3, the independent variable is the share of all fresh graduates hired from each university between 2005 and 2010. We define an individual i as a fresh graduate from university u in year t if individual i joined a firm in year t and graduated from university u in year t or year $t-1$. In columns 2 and 4, the independent variable is the share of all workers in the firm in 2010 who graduated from each university. In column 5, the independent variable is the share of all STEM fresh graduates hired from each university before 2010. We define STEM workers as employees who have at least one degree with a major in either engineering (e.g., electrical, chemical, mechanical), physical sciences (e.g., math, physics, chemistry, computer science, statistics), or biological sciences (e.g., biology, pharmacology). In column 6, the independent variable is the share of STEM workers in the firm in 2010 who graduated from each university. All columns control for firm fixed effects and university fixed effects. Standard errors are clustered at the university level and reported in parentheses. *, **, and *** denote statistical significance at the 10%, 5%, and 1% levels, respectively.

	Share of Post-2010 Hires		Share of Post-2010 AI Hires			
	(1)	(2)	(3)	(4)	(5)	(6)
Share of Pre-2010 Hires	0.465***		0.550***			
	(0.017)		(0.054)			
Share of 2010 Workers		0.147***		0.236***		
		(0.006)		(0.028)		
Share of Pre-2010 STEM Hires					0.342***	
					(0.040)	
Share of 2010 STEM Workers						0.197***
						(0.021)
Firm FE	Y	Y	Y	Y	Y	Y
University FE	Y	Y	Y	Y	Y	Y
Observations	327,313	327,313	177,097	177,097	177,097	177,097

Table 10. Changes in Hiring from Ex-ante AI-strong Universities during the Pre-period (2005–2010)

This table reports the coefficients from regressing the change in the share of fresh graduates from AI-strong universities from 2005 to 2010 on the instrument (the share of STEM workers in the firm in 2010 who graduated from AI-strong universities). The independent variable is standardized to mean zero and standard deviation of one. Columns 2–5 control for ex-ante exposure to universities that are strong in CS research. Columns 3–5 also control for industry sector fixed effects. Columns 4 and 5 add the baseline controls all measured as of 2010: firm-level characteristics (log sales, cash/assets, R&D/sales, log markup, and log number of jobs—resumes in Panel 1 and job postings in Panel 2), log industry wage, and characteristics of the commuting zones where the firms are located (the share of workers in IT-related occupations, the share of college-educated workers, log average wage, the share of foreign-born workers, the share of routine workers, the share of workers in finance and manufacturing industries, and the share of female workers). Column 5 additionally controls for state fixed effects. Regressions are weighted by the number of Cognism resumes in 2010. Standard errors are clustered at the 5-digit NAICS industry level and reported in parentheses. *, **, and *** denote statistical significance at the 10%, 5%, and 1% levels, respectively.

	Δ Share of Fresh Graduates Hired from AI Hubs 2005-2010				
	(1)	(2)	(3)	(4)	(5)
Instrument	0.023	0.020	0.039	0.012	0.069
	(0.088)	(0.097)	(0.088)	(0.107)	(0.112)
CS Control	N	Y	Y	Y	Y
Industry FE	N	N	Y	Y	Y
Baseline Control	N	N	N	Y	Y
State FE	N	N	N	N	Y
Observations	830	830	829	829	825

Table 11. First Stage of the Instrument

This table reports the first stage of the instrument, where we regress our key independent variable—firm-level changes in the share of AI-skilled workers from 2010 to 2018—on the instrument, which measures ex-ante firm-level exposure to the supply of AI-trained university graduates from AI-strong universities. The independent variable is the share of STEM workers in the firm in 2010 who graduate from ex-ante AI-strong research universities. For the dependent variable, Panel 1 considers the resume-based measure of the growth in the share of AI workers from 2010 to 2018, while Panel 2 looks at the job-posting-based measure. Both the independent variable and the dependent variable are standardized to mean zero and standard deviation of one. Regressions are weighted by the number of Cognism resumes in 2010 in Panel 1 and the number of Burning Glass job postings in 2010 in Panel 2. All specifications control for industry sector fixed effects and ex-ante firm-level exposure to universities that are historically strong in CS research. Columns 2–4 also include the baseline controls measured as of 2010: firm-level characteristics (log sales, cash/assets, R&D/sales, log markup, and log number of jobs—resumes in Panel 1 and job postings in Panel 2), log industry wage, and characteristics of the commuting zones where the firms are located (the share of workers in IT-related occupations, the share of college-educated workers, log average wage, the share of foreign-born workers, the share of routine workers, the share of workers in finance and manufacturing industries, and the share of female workers). Columns 3 and 4 add controls for firm-level pre-trends: changes in log sales and log employment from 2000 to 2008. Column 4 adds state fixed effects. Standard errors are clustered at the 5-digit NAICS industry level and reported in parentheses. The first-stage F-statistics of the instrument are reported in both tables for all specifications. *, **, and *** denote statistical significance at the 10%, 5%, and 1% levels, respectively.

Panel 1: AI measure from resume data

	Δ Share of AI Workers			
	(1)	(2)	(3)	(4)
Instrument	0.594***	0.383***	0.412***	0.451***
	(0.155)	(0.087)	(0.091)	(0.084)
Industry FE	Y	Y	Y	Y
CS Control	Y	Y	Y	Y
Baseline Controls	N	Y	Y	Y
Control Pre-trend	N	N	Y	Y
State FE	N	N	N	Y
F Statistic	14.7	19.4	20.4	29.2
Observations	1,001	1,001	777	773

Panel 2: AI measure from job postings data

	Δ Share of AI Workers			
	(1)	(2)	(3)	(4)
Instrument	0.575***	0.446***	0.459***	0.458***
	(0.190)	(0.146)	(0.139)	(0.130)
Industry FE	Y	Y	Y	Y
CS Control	Y	Y	Y	Y
Baseline Controls	N	Y	Y	Y
Control Pre-trend	N	N	Y	Y
State FE	N	N	N	Y
F Statistic	9.2	9.4	10.9	12.5
Observations	889	889	702	696

A Online Appendix

A.1 Case Studies on Firms' AI Investments

In order to illustrate the wide range of applications of AI technologies by individual firms, we provide detailed summaries of the investment patterns and uses of AI technologies within four firms in four different industries.

A.1.1 UnitedHealth Group

UnitedHealth Group (UNH) is a large managed healthcare company based in Minnetonka, Minnesota. The group includes a healthcare arm (UnitedHealthcare) established in 1977 and a new technology arm founded in 2011 (Optum). While the UnitedHealthcare arm makes use of AI techniques to optimize operations ranging from cost projections to fraud detection in medical claims, the launch of Optum highlights the way in which firms such as UNH can leverage AI technologies to expand operations by creating new products and entering new market segments. UNH is one of very few companies with access to detailed patient, patient-physician, and drug-patient interaction data for large portions of the U.S. and many additional global locations, making it perfectly placed to harness AI in its operations.

AI use cases and product impact. Most of the AI investments and impact at UNH center around its Optum arm. The traditional UnitedHealthcare part of UNH uses AI in a limited capacity for predictive analytics that inform business decisions and safeguards for vulnerabilities such as fraud detection. The launch of Optum in 2011 has enabled UNH to leverage AI technologies to deliver new products across several healthcare markets. At its core, Optum is a vast data store of proprietary and 3rd party datasets linked together to enable machine-learning-based analysis. Specifically, the AI-powered Optum products include: (i) statistics on drugs and potential alternatives through the pharmaceutical platform Optum Rx; (ii) analysis of electronic medical records through the Optum One platform for physicians; and (iii) the Optum Population Health Management platform for larger institutions (including employers and federal and state agencies) to optimize costs and accessibility to care. The AI-powered OptumIQ system, which is leveraged throughout the Optum solutions, also targets machine-learning-based prediction and diagnostics for diseases such as atrial fibrillation.

Timeline of AI investments at UNH. The use of AI technologies at UNH traces further back than at most firms. As early as the 1990s, UNH piloted AdjudiPro, an AI-powered platform for processing claims from physicians. However, the presence of AI-skilled labor at UNH remained low throughout the 1990s and 2000s, noticeably picking up in 2011 with the launch of the Optum platform. Thereafter, UNH's investment in AI human capital rose steadily throughout the 2010s. The Optum arm of the firm released the Optum360 and Impact Pro products in 2013 and the

Optum One Analytics Platform in 2014, prompting a further acceleration in the rise of UNH's AI human capital in the second half of the decade. The timeline of AI investments at UNH is displayed in Figure A1.

Internal structure of UNH's AI workforce. UNH has a centralized approach to AI integration, with strategic decisions primarily coming from the headquarters in Minnesota and regional offices handling specific applications. Correspondingly, the majority of UNH's AI workforce concentrates in Minneapolis and Minnetonka, including senior personnel heading AI and machine learning efforts, automation/deployment, consumer analytics, and Optum enterprise analytics. Locations outside of the headquarters tend to employ predominantly engineering and general IT personnel to support the AI efforts.

A.1.2 JPMorgan Chase & Co

JPMorgan Chase & Co (JPM) is the largest bank in the U.S., based in New York City, NY, with consumer banking that has relationships with more than half of U.S. households, a commercial banking arm, a large investment banking business, and a sizable asset management arm. The bank stores hundreds of petabytes of data ranging from credit card transactions and loan applications, to financial news and market data, to alternative data sources.

AI use cases and product impact. The main use cases for AI at JPM fall into the following categories: (i) risk modeling and management ranging from internal cybersecurity to fraud detection in consumer banking and assessment of geo-political risks; (ii) quantitative analysis and algorithmic investment products, including the Algo Central, LOXM, and DeepX programs aimed at executing trades at both maximal speed and optimal prices; (iii) general analytics for Big Data use in broad internal applications including recruiting; and (iv) product development, including enhancements to mobile apps and customer support through AI-powered virtual assistants. In addition, JPM also employs AI in more peripheral applications, for example, with methods for processing of alternative data such as satellite images and mapping contingency plans for AI-driven workforce disruptions. The use of AI at JPM is aimed at both cutting costs (e.g., through risk assessment) and creating new products (e.g., machine-learning-powered trading platforms such as DeepX).

Timeline of AI investments at JPM. As highlighted by Figure A1, investments in AI at JPM began at the turn of the century, with a steady increase through the first decade turning into an exponential growth in the second decade. The explosion in AI investments at JPM during the 2010s is marked by the acquisition of the multimedia recommendations patent in 2011; an underscoring of the risks associated with data security following a data leak in 2016; and finally the establishment of a dedicated AI research initiative (Machine Learning Center for Excellence) spearheaded by Dr. Manuela Veloso (previously the Chair of the Machine Learning Department at Carnegie Mellon

University) in 2018.

Internal structure of JP Morgan's AI workforce. AI efforts at JPM are centered in the New York location, with peripheral AI expertise throughout the U.S., in London, and in India. JPM has taken a top-down approach to AI investments, with involvement from the highest levels of management and the establishment of a dedicated AI research team in 2018. At the same time, JPM's investments in AI have seen not only the formation of dedicated AI hubs, but also a different approach to corporate structuring. Specifically, the firm's approach relies heavily on small skilled and responsive AI "task-forces" specializing in different sectors (quantitative analysis, user experience, etc.), which can alternatively work on experimental projects (e.g., satellite imagery analysis) or coordinate together to work on core applications (trading algorithms, firm-wide cybersecurity).

A.1.3 Caterpillar Inc.

Caterpillar Inc. is a large construction manufacturing firm headquartered in Deerfield, IL, with a variety of additional business activities including financial products and insurance. The firm has correspondingly varied applications for AI, ranging from inventory management to part recognition, to credit scoring for machinery financing.

AI use cases and product impact. AI investments at Caterpillar are organized along several key verticals. First, the Data Innovation Lab at UIUC conducts core projects in demand forecasting (unstable demand anticipation) and inventory management, in part identification (using techniques from image recognition), and in tracking and tracing technology for fleet management. Second, Caterpillar's asset intelligence efforts include a product line of Internet of Things (IoT) style analytics for managers and machine operators, which facilitates data collection, interpretation, predictive maintenance, and integration. Lastly, smaller targeted efforts at Caterpillar also employ AI techniques in other parts of the business, including leveraging sensor-based data for equipment management and using drone data to optimize job site organization. Caterpillar's uses of AI serve to modernize the firm's machinery, streamline operations and reduce waste through better forecasting and inventory management, and expand the product offerings with the IoT product line and efficient long-term service contracts.

Timeline of AI investments at Caterpillar. Caterpillar began employing workers with AI expertise at the turn of the century, but the growth in the firm's AI workforce went hand-in-hand with the growth in the firm's overall workforce throughout the 2000s (with a dip during the financial crisis). The share of AI employees at Caterpillar noticeably picked up only in mid-2010s, with the CEO Douglas Obenhelmer underscoring the importance of capitalizing on the firm's vast available data resources. Since 2014, Caterpillar has aggressively pursued the development of "smart" machinery, connecting it to predictive IoT-style networks and developing better models for demand prediction. In 2015, Caterpillar established the Analytics and Innovation Division headed

by Greg Foley, and in 2016, the firm hired Morgan Vawters as the Chief of Analytics. The timeline of Caterpillar's investments in AI human capital is presented in Figure A1.

Internal structure of Caterpillar's AI workforce. The majority of the AI employees at Caterpillar are in the firm's Technology division, with notable presence also in Business and Production departments. The major locations setting the trend for Caterpillar's AI adoption are the company centers in Chicago and Peoria, Illinois, with projects percolating through the dedicated research centers such as the Champaign Innovation Center and production centers such as the manufacturing plant in Aurora, Illinois.

A.1.4 Qualcomm Inc.

Qualcomm Inc. is a wireless telecommunications firm headquartered in San Diego, CA. The firm produces a number of products including semiconductors, hardware, software, and other services related to wireless technology. Device manufacturers such as Apple are Qualcomm's primary clients.

AI use cases and product impact. The principal use of AI at Qualcomm over the past decade and a half has been the improvement of its core products. This includes optimization of chips within mobile devices, improvements to the camera using techniques from computer vision for face recognition and auto-adjustments, audio and video processing, physical sensitivity, power use, and location tracking capabilities. More recently, Qualcomm made a large investment in the development of the Snapdragon Neural Processing Engine (SNPE) platform, which offers a combination of hardware and software on android devices that allows developers to more easily create AI-powered or assisted applications. With the exception of a few stand-alone projects for internal data processing efficiency (e.g., improving internal servers), Qualcomm does not appear to be heavily invested in applying AI for applications such as sales or supply chain optimization, unlike Caterpillar Inc. described above.

Outside of its core businesses, Qualcomm has invested in a number of side products at more exploratory or proof-of-concept stages, such as general work on autonomous vehicles, or enterprise partnerships, for example with Accenture and Kellogg on virtual reality tracking of customers for marketing purposes. This highlights the broad scope of AI technologies that facilitate firms entering new markets: for example, the autonomous vehicle work at Qualcomm makes use of the efforts aimed at enhancing smartphone components, only applied to a different domain.

Timeline of AI investments at Qualcomm. As can be seen from the timeline in Figure A1, the presence of AI employees at Qualcomm began earlier than in the other firms, and by 2007 the firm initiated dedicated AI research projects in its research arm. The ramp up continued through 2013, marked by collaborations with outside partners such as Brain Corp and internal projects on problems such as face detection. After 2013, Qualcomm saw notable consequences of the earlier

investments, including the first release of SNPE and the formation of an organizationally separate AI research group, but the share of Qualcomm's overall workforce that is skilled in AI remained approximately flat from 2013 to 2018.

Internal structure of Qualcomm's AI workforce. Between 2000 and 2018, the majority of Qualcomm's AI employees have been engineers focused on the improvement of the core product being developed at each point in time, supported by an auxiliary staff of patent counsels and data scientists. In 2018, Qualcomm established a separate AI research group, which is bringing about increased centralization of its AI workforce. Specifically, AI efforts at Qualcomm are organized around the San Diego headquarters, with leadership on overall AI strategy, the newly formed AI research group, and teams spanning nearly every project from computer vision R&D to GPU architecture. Smaller AI offices, scattered mostly throughout the U.S. and Canada, tend to focus on single elements of Qualcomm's AI initiative (for example, SNPE in Toronto and positioning sensors in Santa Clara).

Figure A1. Timeline of AI investments by UnitedHealth Group

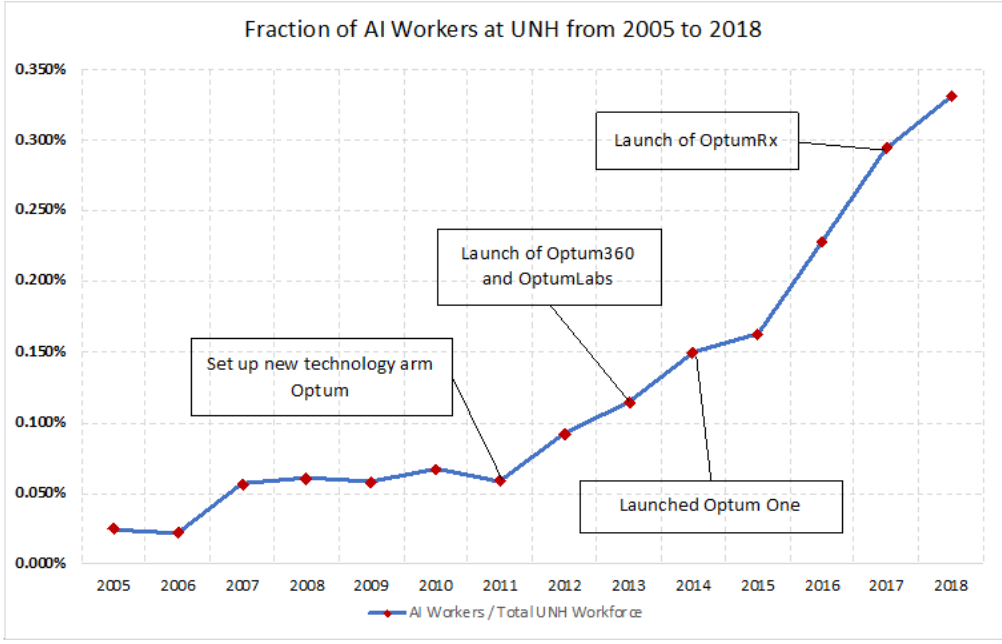


Figure A1. Timeline of AI investments by JPMorgan Chase & Co

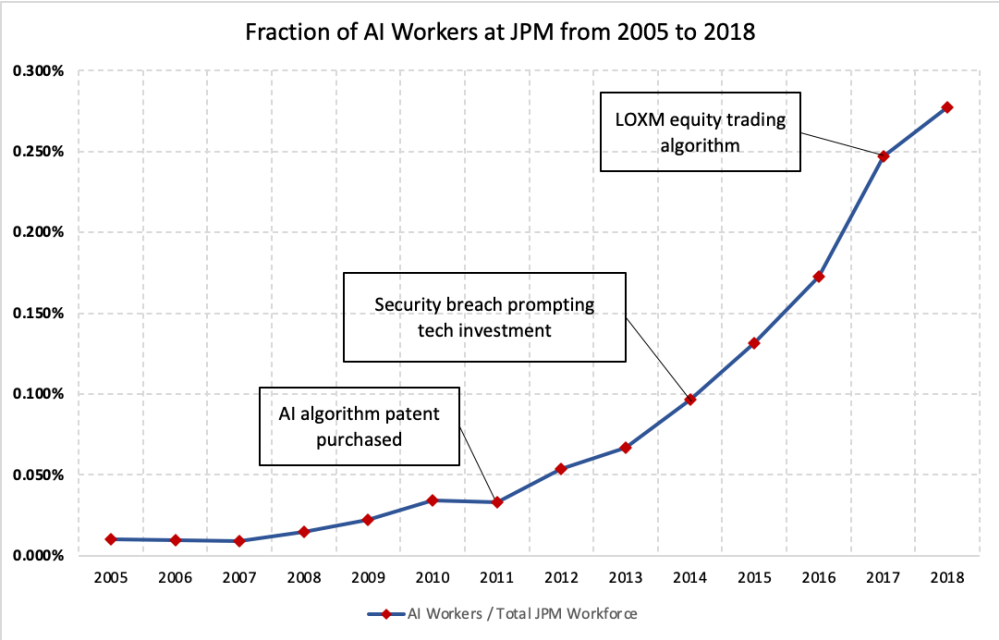


Figure A1. Timeline of AI investments by Caterpillar Inc

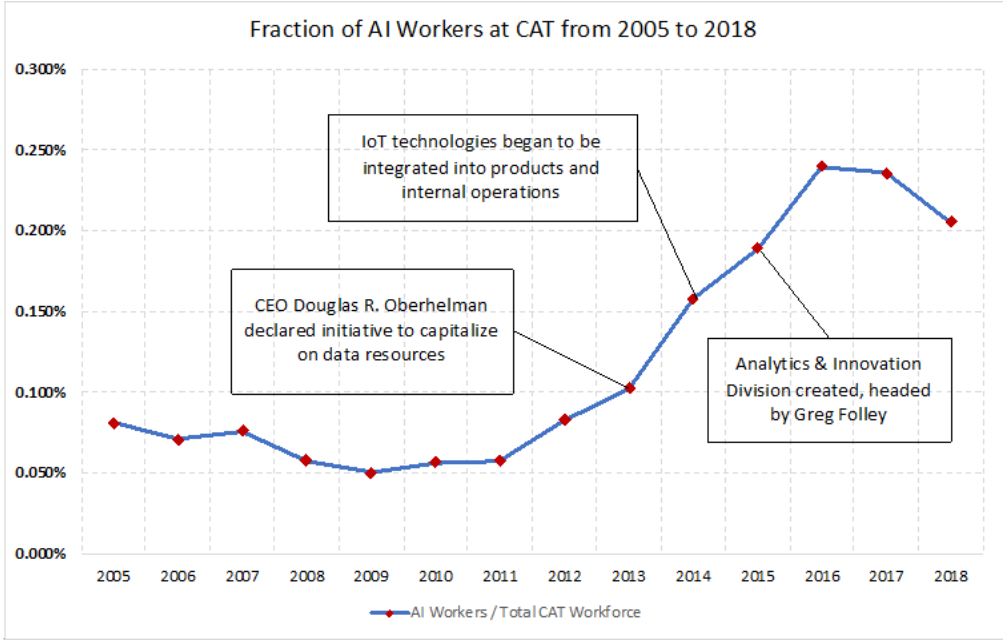
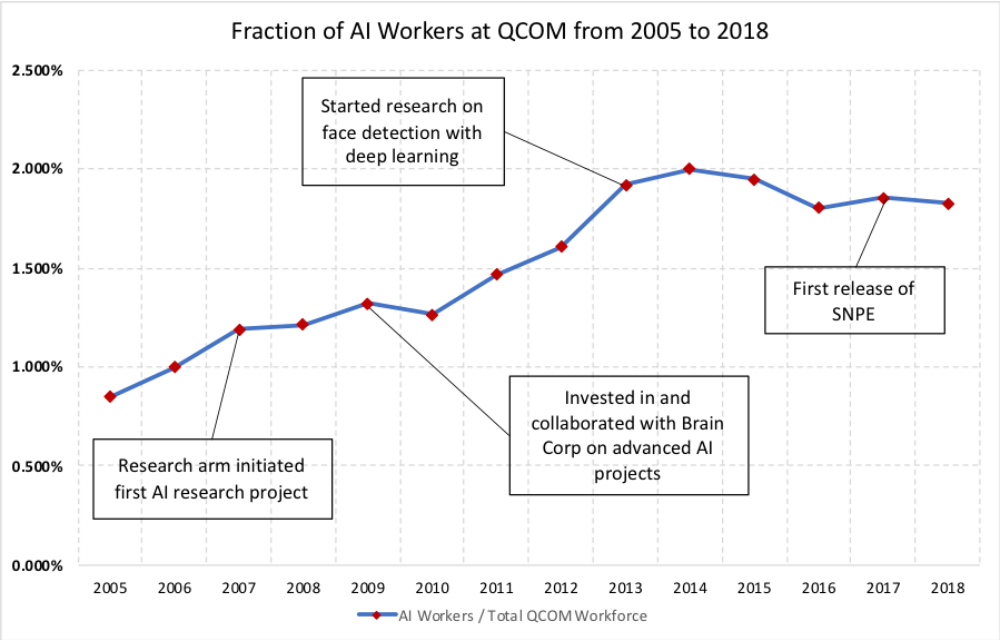


Figure A1. Timeline of AI investments by Qualcomm Inc



A.2 Additional Figures and Tables

Figure A1. Matching Rate to Compustat in Job Postings Data

This figure shows the time series of the share of all job postings and the share of AI job postings (job postings with continuous measure ω_j^{AI} above 0.1) that are matched to Compustat firms in the Burning Glass data in 2007 and in each year from 2010 to 2018.

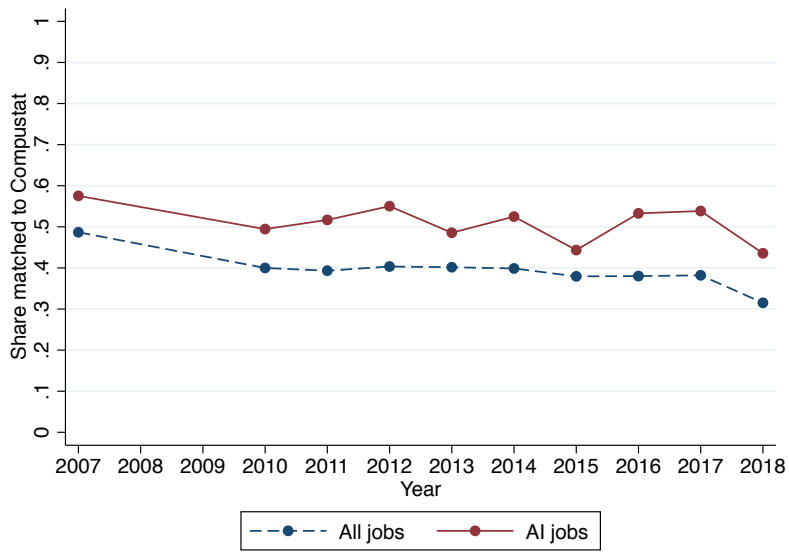


Figure A1. Distribution of AI Investments across U.S. Geographies

This figure shows a heat map of changes in the job-posting-based measure of AI investments across geographies in the U.S. The figure plots the change in the average AI-relatedness measure (w_j^{AI}) of job postings of public firms in each commuting zone from 2010 to 2018.

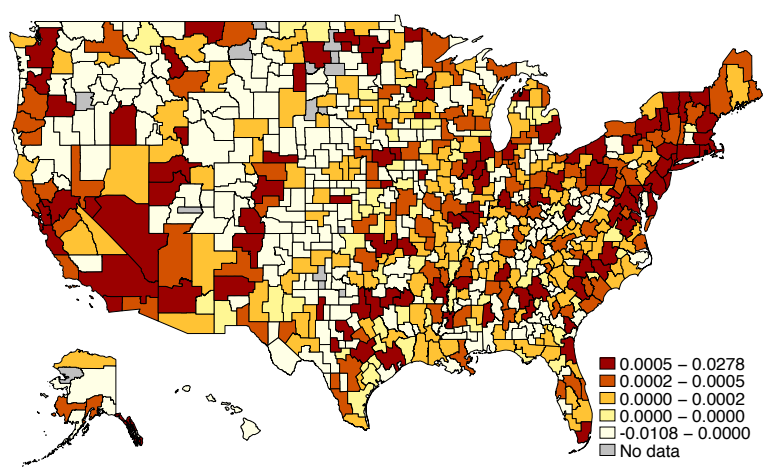
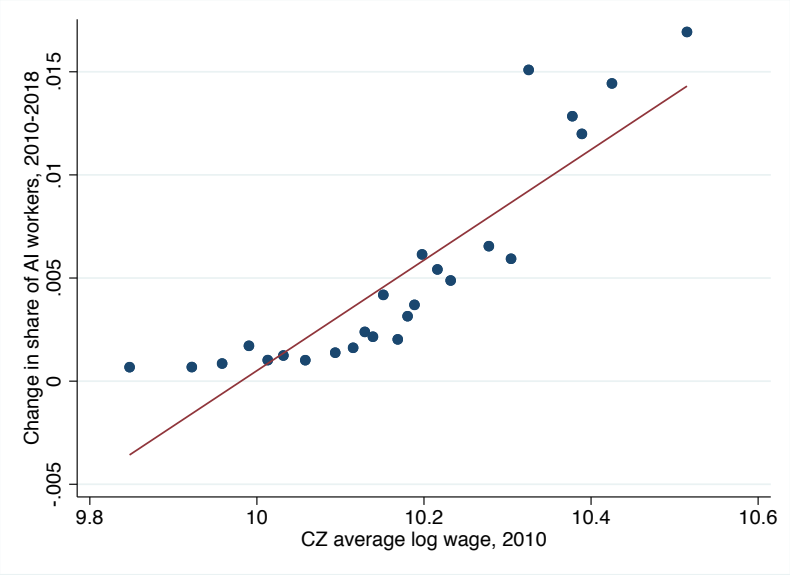
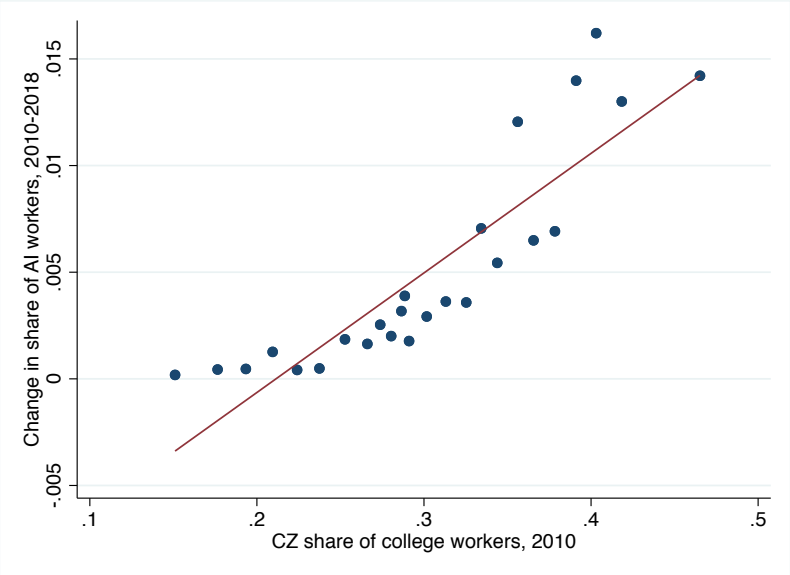


Figure A1. AI Investments and Local Conditions

This figure presents binned scatterplots of commuting-zone-level AI investments against local conditions. Solid line are the fitted regression line, where the regressions are weighted by commuting zones' populations in 2010. The y-axis is the change in AI investments (i.e., the change in the share of AI workers) from 2010 to 2018, using the Burning Glass data (based on the sample of public firms). The x-axis in Panel (a) is the average log wage of a commuting zone in 2010. The x-axis in Panel (b) is the share of college educated workers in a commuting zone in 2010. The log wage and the share of college-educated workers are from the Census American Community Survey. The t-statistic on the regression slope is 23.8 in Panel (a) and 24.1 in Panel (b).



(a) AI Investments and Local Average Wage



(b) AI Investments and Local Share of College-educated Workers

Table A1. Skills with Highest AI-Relatedness Measures in Burning Glass Job Postings

This table lists the top skills in the Burning Glass data ranked by the skill-level AI measure w_s^{AI}. For each skill, we report the percentage of jobs requiring that skill that also require one of the four core AI skills—artificial intelligence, machine learning, computer vision, and natural language processing. For example, for jobs that require "Recurrent Neural Network (RNN)," 96.5% also require one of the four core AI-skills. Only skills that appear in least 50 job postings are included.

#	Skills	AI-relatedness Score
1	Artificial Intelligence	1.000
2	Computer Vision	1.000
3	Machine Learning	1.000
4	Natural Language Processing	1.000
5	ND4J (software)	0.980
6	Kernel Methods	0.979
7	Microsoft Cognitive Toolkit	0.975
8	Xgboost	0.972
9	Sentiment Classification	0.971
10	Long Short-Term Memory (LSTM)	0.971
11	Libsvm	0.968
12	Semi-Supervised Learning	0.968
13	Recurrent Neural Network (RNN)	0.965
14	Word2Vec	0.956
15	MXNet	0.953
16	Caffe Deep Learning Framework	0.950
17	Autoencoders	0.949
18	MLPACK (C++ library)	0.942
19	Keras	0.941
20	Theano	0.938
21	Torch (Machine Learning)	0.932
22	Wabbit	0.929
23	Boosting (Machine Learning)	0.905
24	TensorFlow	0.904
25	Vowpal	0.903
26	Convolutional Neural Network (CNN)	0.897
27	Jung Framework	0.894
28	OpenNLP	0.894
29	Natural Language Toolkit (NLTK)	0.892
30	Unsupervised Learning	0.891
31	Dlib	0.891
32	Scikit-learn	0.889
33	Latent Semantic Analysis	0.889
34	Latent Dirichlet Allocation	0.889
35	Stochastic Gradient Descent (SGD)	0.881
36	Gradient boosting	0.872
37	Dimensionality Reduction	0.861
38	Deep Learning	0.859
39	DBSCAN (Density-Based Spatial Clustering of Applications with Noise)	0.855
40	AI ChatBot	0.844
41	Recommender Systems	0.842
42	Random Forests	0.840
43	Deeplearning4j	0.839
44	Support Vector Machines (SVM)	0.817
45	Unstructured Information Management Architecture	0.806
46	Apache UIMA	0.805
47	Maximum Entropy Classifier	0.799
48	Hidden Markov Model (HMM)	0.796
49	Pybrain	0.786
50	Computational Linguistics	0.780
51	Naive Bayes	0.768
52	H2O (software)	0.763
53	Expectation-Maximization (EM) Algorithm	0.763
54	WEKA	0.761
55	Clustering Algorithms	0.740
56	Matrix Factorization	0.739
57	Object Recognition	0.727
58	Classification Algorithms	0.721
59	Information Extraction	0.709
60	Image Recognition	0.706
61	Bayesian Networks	0.705
62	Supervised Learning (Machine Learning)	0.695
63	OpenCV	0.688
64	K-Means	0.683
65	Sentiment Analysis / Opinion Mining	0.679
66	Machine Translation (MT)	0.655
67	Neural Networks	0.640

Table A2. Examples of AI and Non-AI Job Postings in Burning Glass

This table displays examples of job postings and their continuous AI measure ω_j^{AI}. Jobs 1–10 are examples of AI-skilled jobs, with the first five being non-data-specific and the last 5 being data-specific. Jobs 11–20 are examples of non-AI-skilled jobs, with the first five being data specific and the last 5 being non-data-specific. The AI relatedness score of each skill is listed in parentheses.

	Job Title	Employer	Skills	Score
	AI jobs			
1	Research Engineer - Natural Language Processing	InterActiveCorp	Machine Learning (1), Natural Language Processing (1), Natural Language Toolkit (0.895), Computational Linguistics (0.777), WEKA (0.760), Information Extraction (0.709), Mahout (0.593), Information Retrieval (0.360), Apache Hadoop (0.204), Lucene (0.188), SOLR (0.142), C++ (0.067), Software Engineering (0.043), Python (0.116), Lexical Semantics (0.625), Ontologies (0.326), Java (0.040), PERL Scripting Language (0.034), Relational Databases (0.024), SQL (0.023), Search Analytics (0.022), Shell Scripting (0.020), Web Analytics (0.012), Research (0.011), Online Research (0.010), Extensible Markup Language (0.010)	0.31
2	Computer Vision & Image Processing Researcher	Rambus Incorporated	Computer Vision (1), Object Recognition (0.725), OpenCV (0.689), Pattern Recognition (0.442), CUDA (0.362), Image Processing (0.179), Troubleshooting Technical Issues (0.006), C++ (0.067), Communication Skills (0.003), MATLAB (0.113), Self-Motivation (0.002), Optical System Design and Analysis (0.019), Research (0.011), Writing (0.004), OpenGL (0.117), Prototyping (0.042), Very Large Scale Integration (0.037), Creativity (0.007)	0.21
3	Algorithm Developer	IBM	Natural Language Processing (1), Machine Learning (1), IBM Watson (0.125), Java (0.040), Software Development (0.027), Candidate Generation (0.013), Creativity (0.007), Troubleshooting (0.003), English (0.002)	0.25
4	Senior Autonomous Vehicle Localization Software Engineer	Nvidia Corporation	Computer Vision (1), Deep Learning (0.859), Linear Algebra (0.187), OpenGL (0.117), C++ (0.067), Software Engineering (0.043), Geometry (0.009), Motor Vehicle Operation (0.006), Teamwork / Collaboration (0.005), Calibration (0.004)	0.230
5	Speech Recognition Scientist	Vocera Communications	Computational Linguistics (0.780), Automatic Speech Recognition (0.457), Speech Recognition (0.215), Experiments (0.045), Performance tuning (0.011), Research (0.011), Written Communication (0.003)	0.217
6	Data Scientist	Zappos	Machine Learning (1), Natural Language Processing (1), Boosting (Machine Learning) (0.902), Support Vector Machines (0.816), Naive Bayes (0.759), Matrix Factorization (0.738), Classification Algorithms (0.718), Data Science (0.379), Data Mining (0.159), NoSQL (0.119), Clustering (0.103), Data Structures (0.069), Relational DataBase Management System (0.028), SQL (0.023), Attribution Modeling (0.072), Detail-Oriented (0.002), Revenue Projections (0.003), Traffic Maintenance (0.002)	0.384
7	Data Mining Engineer	Apple Inc.	Artificial Intelligence (1), Natural Language Processing (1), Machine Learning (1), Unsupervised Learning (0.891), Supervised Learning (0.696), Mahout (0.593), Pattern Recognition (0.442), Apache Hadoop (0.204), Image Processing (0.179), Data Mining (0.159), NoSQL (0.119), Data Collection (0.008), Communication Skills (0.003), Java (0.040), Detail-Oriented (0.002), MATLAB (0.113), SQL (0.023), Network Engineering (0.007), Research (0.011), Python (0.116), Meeting Deadlines (0.002), R (0.248), Predictive Models (0.243)	0..309
8	Big Data Engineer	Socialwire	Machine Learning (1), Recommender Systems (0.843), MapReduce (0.285), Apache Hadoop (0.204), Big Data (0.196), Facebook (0.006), R (0.248), Pinterest (0.003), Writing (0.004), MATLAB (0.113)	0.290
9	Big Data Senior Data Scientist	AT&T	Machine Learning (1), WEKA (0.760), Clustering Algorithms (0.738), Mahout (0.593), Data Science (0.379), Big Data (0.196), Data Mining (0.159), Clustering (0.103), Simulation (0.028), Experimental Testing (0.039), R (0.248), SPSS (0.067), Creativity (0.007), SAS (0.053), Information Systems (0.007), Experiments (0.045), Presentation Skills (0.006), Research (0.011), Data Quality (0.025)	0.235
10	Data Scientist	Warby Parker	Natural Language Processing (1), Natural Language Toolkit (0.895), Random Forests (0.839), Pandas (0.498), Data Science (0.379), PIG (0.290), Apache Hadoop (0.204), Data Mining (0.159), Data Visualization (0.136), Tableau (0.074), Pentaho (0.058), NumPy (0.552), SQL (0.023), Python (0.116), Java (0.040), DevOps (0.039), Agile Development (0.030), Creativity (0.007), Django (0.039), Apache Webserver (0.034), Predictive Models (0.243), Relational Databases (0.024), Data Modeling (0.037)	0.249
	Non-AI jobs			
11	Director Of Business Intelligence	Odesus Incorporated	Data Science (0.379), Data Transformation (0.060), SQL (0.023), Communication Skills (0.003), SQL Server Reporting Services (0.009), SQL Server (0.009), SQL Server Analysis Services (0.034), Budgeting (0.001), Microsoft Sharepoint (0.002), Data Warehousing (0.025), MySQL (0.028), Key Performance Indicators (0.006), Problem Solving (0.005), Web Analytics (0.012), Market Research (0.006), Data Modeling (0.037), Business Intelligence (0.026), Creativity (0.007)	0.037
12	Director, Data & Analytics	Decision Resources	Big Data (0.196), Business Intelligence (0.026), Business Intelligence Industry Knowledge (0.020), Teamwork / Collaboration (0.005), Biopharmaceutical Industry Knowledge (0.004), Communication Skills (0.003)	0.042
13	Senior Healthcare Economics Data Analyst	UnitedHealth Group	Tableau (0.076), Advanced Statistics (0.149), SAS (0.053), Data Analysis (0.026), SQL (0.023), Economics (0.016), Database Design (0.014), Clinical Data Analysis (0.012), Clinical Data Review (0.010), Business Process (0.006)	0.039
14	Data Analyst	United Technologies Corporation	Data Analysis (0.026), Data Quality (0.025), Data Management (0.018), Database Design (0.014), Proposal Writing (0.007), Product Improvement (0.007), Business Planning (0.002)	0.014
15	Sas Database Administrator	Pitney Bowes	SAS (0.053), SQL (0.023), Business Strategy (0.009), Teradata DBA (0.005), Self-Starter (0.004), Database Administration (0.004), Pivot Tables (0.004), Market Analysis (0.004), Technical Support (0.002), Microsoft Excel (0.002)	0.011
16	Delivery Driver And Technician	Rotech Healthcare	Physical Abilities (0.000), Lifting Ability (0.000), Caregiving (0.000), Patient Contact (0.000), Patient Transportation and Transfer (0.000), HAZMAT (0.000), Hazardous Materials Endorsement (0.000)	0
17	Vice President Underwriting	Morgan Stanley	Workflow Management (0.005), Written Communication (0.003), Detail-Oriented (0.002), Financial Analysis (0.001), Mortgage Underwriting (0.001), Staff Management (0.001)	0.002
18	Quality Assurance Engineer	Amazon	Computer Engineering (0.034), Software Development (0.027), User Interface (UI) Design (0.016), Software Quality Assurance (0.010), Black-box testing (0.009), Quality Assurance and Control (0.003), Consumer Electronics (0.002)	0.014
19	Sales Associate	GNC	Sales (0.001), Retail Industry Knowledge (0.000), Retail Sales (0.000), Basic Mathematics (0.000)	0
20	Dog And Cat Department Manager	Petco	Creativity (0.007), Leadership (0.003), Budgeting (0.001), Sales Goals (0.001), Retail Industry Knowledge (0.000), Physical Abilities (0.000), Inventory Management (0.000)	0.002

Table A3. Job Titles with the Highest Average AI-relatedness Measures

This table reports the job titles in Burning Glass with the highest average job-level AI measure ω_j^{AI}. We only include job titles that have at least 50 job postings and are matched to Compustat firms.

	Job Title	Avg. Continuous AI Measure
1	Artificial Intelligence Engineer	0.497
2	Senior Data Scientist - Machine Learning Engineer	0.394
3	Lead Machine Learning Scientist - Enterprise Products	0.369
4	AI Consultant	0.369
5	AI Senior Analyst	0.358
6	Machine Learning Engineer	0.315
7	Technician Architecture Delivery Senior Analyst AI	0.311
8	Artificial Intelligence Analyst	0.308
9	Software Engineer, Machine Learning	0.307
10	Artificial Intelligence Architect	0.303
11	Machine Learning Researcher	0.300
12	Computer Vision Engineer	0.293
13	Senior Machine Learning Engineer	0.286
14	Senior Machine Learning Scientist	0.281
15	Senior Software Engineer - Machine Learning	0.278
16	Senior Engineer II - Data Scientist	0.265
17	Senior Machine Learning Researcher	0.264
18	Artificial Intelligence Consultant	0.263
19	Computer Vision Scientist	0.256
20	Lead Machine Learning Researcher	0.255
21	Senior AI Engineer	0.248
22	Senior Applied Scientist	0.245
23	Senior Engineer - Machine Learning	0.243
24	Senior Risk Modeler	0.241
25	Data Scientist - Engineer	0.238
26	Artificial Intelligence Manager	0.237
27	Machine Learning Scientist	0.230
28	Applied Scientist	0.230
29	Software Engineer - Data Mining/Data Analysis/Machine Learning	0.229
30	Senior Associate, Data Scientist	0.223
31	Director, Data Scientist	0.222
32	Big Data Hadoop Consultant	0.214
33	Vice President- Data Analytics	0.211
34	Data Scientist Specialist	0.210
35	Applied Researcher	0.209
36	Data Scientist, Junior	0.205
37	Senior Staff Data Scientist	0.204
38	Principal Data Scientist	0.204
39	Director, Data Science	0.203
40	Research And Development Engineer - Data Mining/Data Analysis/Machine Learning	0.195
41	Manager, Data Scientist	0.192
42	Big Data Scientist	0.191
43	Architect - Relevance Infrastructure	0.191
44	Director Of Data Science	0.189
45	Senior Manager, Data Science	0.189
46	Data Science Specialist	0.188
47	Data Scientist II	0.188
48	Senior Data Science Engineer	0.187
49	Staff Data Scientist	0.186
50	Lead Data Scientist	0.186

Table A4. Summary Statistics

This table reports summary statistics for the sample of firms in our baseline regressions (including 1052 firms in the regressions with the resume-based measure and 935 firms in the regressions with the job-postings-based measure). All changes in variables are computed over 2010–2018. Changes in the numbers of trademarks and patents are measured as changes in log(1+number) to take into account firms with zero trademarks or patents. We follow the methodology proposed by Ganglmair et al. (2021) to distinguish between product patents and process patents. The number of product and process patents are de-trended using the entire patent sample due to the truncation of the patent sample in recent years. The change in the product mix is measured as the sum of annual changes from 2010 to 2018, where each annual change is the angle between the two word vectors indicating firms' product offerings in that year and the previous year following Hoberg et al. (2014). Log markup is measured as the log of the ratio of sales to cost of goods sold, following De Loecker et al. (2020). For each variable, we report the number of observations, the mean, the standard deviation, the median, and 1st, 5th, 10th, 25th, 75th, 90th, 95th, and 99th percentiles.

Variable Name	**N**	**Mean**	**Std. Deviation**	**p1**	**p5**	**p10**	**p25**	**p50**	**p75**	**p90**	**p95**	**p99**
Change in share of AI workers (Cognism)	1052	.0010	.0022	-.0018	-.00011	0	0	.00024	.0011	.0029	.0046	.014
Change in share of AI workers (Burning Glass)	935	.0046	.012	-.0033	0	0	0	0	.0034	.013	.026	.076
Change in log sales	1052	.45	.66	-1.2	-.61	-.24	.073	.39	.77	1.3	1.7	2.6
Change in log employment	1052	.3	.64	-1.8	-.67	-.35	-.036	.26	.61	1.0	1.4	2.2
Change in log market value	1010	.52	.72	-1.2	-.78	-.38	.1	.51	.94	1.4	1.7	2.6
Change in log number of trademarks	553	-.13	1.2	-3.0	-2.3	-1.8	-.92	0	.41	1.4	2.0	3.2
Change in log number of product patents	621	-.20	1.1	-2.9	-2.2	-1.8	-.94	.062	.85	.85	.85	1.5
Change in log number of process patents	621	-.067	.99	-3.0	-2.0	-1.5	-.63	.067	.76	.76	.76	1.5
Change in product mix	958	4.3	1.3	2.1	2.6	3.0	3.4	4.1	5.0	6.0	6.9	7.9
Change in log sales per worker	1052	.14	.39	-1.0	-.44	-.24	-.038	.12	.30	.56	.79	1.6
Change in revenue TFP	977	.0065	.36	-1.2	-.51	-.36	-.17	.0016	.17	.38	.55	1.2
Change in log COGS	1052	.38	.67	-1.6	-.80	-.34	.038	.35	.72	1.2	1.6	2.3
Change in log operating expense	1052	.42	.61	-1.3	-.57	-.22	.073	.37	.74	1.2	1.6	2.1
Employment in 2010 (thousands)	1052	23	55	.066	.21	.39	1.2	4.5	16	57	109	294
Sales in 2010 (millions)	1052	9104	24318	16	69	146	491	1601	5965	20732	50272	125805
Cash / Assets in 2010	1052	.16	.17	.00054	.0062	.014	.041	.11	.23	.41	.55	.70
R&D / Sales in 2010	1052	.055	.17	0	0	0	0	0	.029	.14	.23	1.1
Log markup in 2010	1052	.56	.47	-.34	.082	.14	.25	.44	.74	1.2	1.5	2.3

Table A5. AI Investments and Firm Growth across Non-Tech Sectors

This table reports the coefficients from long-differences regressions of firm growth from 2010 to 2018 on the contemporaneous changes in AI investments among U.S. public firms (in non-tech sectors), separately by broad industry sector. Columns 1 and 2 consider firms in the manufacturing sector (2-digit NAICS = 31, 32, 33), columns 3 and 4 consider firms in the wholesale and retail trade sectors (2-digit NAICS = 42, 44, 45), columns 5 and 6 look at firms in the finance sector (2-digit NAICS = 52), and columns 7 and 8 include firms in the other non-tech sectors (all 2-digit NAICS sectors, except those listed above and 51 and 54). Panel 1 considers the resume-based measure of the share of AI workers, while Panel 2 looks at the job-posting-based measure. The changes in the AI measures are standardized to mean zero and standard deviation of one within each sample. We consider two measures of firm growth: changes in log sales in odd columns and changes in log employment in even columns. Regressions are weighted by the number of Cognism resumes in 2010 in Panel 1 and the number of Burning Glass job postings in 2010 in Panel 2. All regressions include industry sector fixed effects and the baseline controls all measured as of 2010: firm-level characteristics (log sales, cash/assets, R&D/sales, log markup, and log number of jobs—resumes in Panel 1 and job postings in Panel 2), log industry wage, and characteristics of the commuting zones where the firms are located (the share of workers in IT-related occupations, the share of college-educated workers, log average wage, the share of foreign-born workers, the share of routine workers, the share of workers in finance and manufacturing industries, and the share of female workers). Standard errors are clustered at the 5-digit NAICS industry level and reported in parentheses. *, **, and *** denote statistical significance at the 10%, 5%, and 1% levels, respectively.

Panel 1: AI measure from resume data

	Manufacturing		Wholesale & Retail		Finance		Other	
	Δ Log Sales	Δ Log Employment	Δ Log Sales	Δ Log Employment	Δ Log Sales	Δ Log Employment	Δ Log Sales	Δ Log Employment
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Δ Share AI Workers	0.135**	0.125*	0.321***	0.357***	0.239**	0.264**	0.177***	0.125*
	(0.057)	(0.072)	(0.061)	(0.061)	(0.107)	(0.103)	(0.061)	(0.067)
Industry FE	Y	Y	Y	Y	Y	Y	Y	Y
Controls	Y	Y	Y	Y	Y	Y	Y	Y
Adj R-Squared	0.321	0.281	0.817	0.857	0.473	0.478	0.473	0.363
Observations	516	516	109	109	149	149	278	278

Panel 2: AI measure from job postings data

	Manufacturing		Wholesale & Retail		Finance		Other	
	Δ Log Sales	Δ Log Employment	Δ Log Sales	Δ Log Employment	Δ Log Sales	Δ Log Employment	Δ Log Sales	Δ Log Employment
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Δ Share AI Workers	0.135**	0.110	0.259***	0.291***	0.195***	0.141*	0.141	0.040
	(0.065)	(0.081)	(0.035)	(0.035)	(0.067)	(0.068)	(0.092)	(0.095)
Industry FE	Y	Y	Y	Y	Y	Y	Y	Y
Controls	Y	Y	Y	Y	Y	Y	Y	Y
Adj R-Squared	0.319	0.366	0.673	0.771	0.594	0.605	0.570	0.731
Observations	458	458	102	102	123	123	252	252

Table A6. AI Investments and Firm Growth in Tech Sectors

This table reports the coefficients from long-differences regressions of firm growth from 2010 to 2018 on the contemporaneous changes in AI investments among U.S. public firms in tech sectors. Columns 1 and 2 consider firms in the information sector (2-digit NAICS = 51), and columns 3 and 4 consider firms in the professional and business services sector (2-digit NAICS = 54). Panel 1 considers the resume-based measure of the share of AI workers, while Panel 2 looks at the job-posting-based measure. The changes in the AI measures are standardized to mean zero and standard deviation of one within each sample. We consider two measures of firm growth: changes in log sales in odd columns and changes in log employment in even columns. Regressions are weighted by the number of Cognism resumes in 2010 in Panel 1 and the number of Burning Glass job postings in 2010 in Panel 2. All regressions include industry sector fixed effects and the baseline controls all measured as of 2010: firm-level characteristics (log sales, cash/assets, R&D/sales, log markup, and log number of jobs—resumes in Panel 1 and job postings in Panel 2), log industry wage, and characteristics of the commuting zones where the firms are located (the share of workers in IT-related occupations, the share of college-educated workers, log average wage, the share of foreign-born workers, the share of routine workers, the share of workers in finance and manufacturing industries, and the share of female workers). Standard errors are clustered at the 5-digit NAICS industry level and reported in parentheses. *, **, and *** denote statistical significance at the 10%, 5%, and 1% levels, respectively.

Panel 1: AI measure from resume data

	Information		Prof. & Business Svcs	
	Δ Log Sales	Δ Log Employment	Δ Log Sales	Δ Log Employment
	(1)	(2)	(3)	(4)
Δ Share AI Workers	0.495***	0.409***	0.035	0.140*
	(0.125)	(0.114)	(0.125)	(0.073)
Industry FE	Y	Y	Y	Y
Controls	Y	Y	Y	Y
Adj R-Squared	0.673	0.622	0.501	0.388
Obs	129	129	54	54

Panel 2: AI measure from job postings data

	Information		Prof. & Business Svcs	
	Δ Log Sales	Δ Log Employment	Δ Log Sales	Δ Log Employment
	(1)	(2)	(3)	(4)
Δ Share AI Workers	0.279***	0.286***	0.083	0.101*
	(0.036)	(0.042)	(0.085)	(0.050)
Industry FE	Y	Y	Y	Y
Controls	Y	Y	Y	Y
Adj R-Squared	0.384	0.388	0.092	0.395
Obs	119	119	49	49

Table A7. AI Investments and Firm Growth: Controlling for Detailed Industry Fixed Effects

This table reports the coefficients from long-differences regressions of firm growth from 2010 to 2018 on the contemporaneous firm-level changes in AI investments among U.S. public firms (in non-tech sectors), with detailed industry fixed effects. We consider three measures of growth: changes in log sales (columns 1–4), changes in log employment (columns 5–8), and changes in log market value (columns 9–12). The main independent variable is the growth in the share of AI workers from 2010 to 2018, standardized to mean zero and standard deviation of one. Panel 1 considers the resume-based measure of the share of AI workers, while Panel 2 looks at the job-posting-based measure. Regressions are weighted by the number of Cognism resumes in 2010 in Panel 1 and the number of Burning Glass job postings in 2010 in Panel 2. All specifications include the baseline controls all measured as of 2010: firm-level characteristics (log sales, cash/assets, R&D/sales, log markup, and log number of jobs—resumes in Panel 1 and job postings in Panel 2), log industry wage, and characteristics of the commuting zones where the firms are located (the share of workers in IT-related occupations, the share of college-educated workers, log average wage, the share of foreign-born workers, the share of routine workers, the share of workers in finance and manufacturing industries, and the share of female workers). Columns 1, 5, and 9 include the baseline specification with the industry sector fixed effects; columns 2, 6, and 10 control for 3-digit NAICS fixed effects; columns 3, 7, and 11 control for 4-digit NAICS fixed effects; columns 4, 8, and 12 control for 5-digit NAICS fixed effects. Standard errors are clustered at the 5-digit NAICS industry level and reported in parentheses. *, **, and *** denote statistical significance at the 10%, 5%, and 1% levels, respectively.

Panel 1: AI measure from resume data

	Δ Log Sales				Δ Log Employment				Δ Log Market Value			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
Δ Share AI Workers	0.203***	0.184***	0.194**	0.208**	0.219***	0.188**	0.193*	0.197*	0.224***	0.183**	0.195*	0.210**
	(0.060)	(0.066)	(0.090)	(0.090)	(0.077)	(0.082)	(0.107)	(0.104)	(0.077)	(0.081)	(0.102)	(0.100)
NAICS2 FE	Y	N	N	N	Y	N	N	N	Y	N	N	N
NAICS3 FE	N	Y	N	N	N	Y	N	N	N	Y	N	N
NAICS4 FE	N	N	Y	N	N	N	Y	N	N	N	Y	N
NAICS5 FE	N	N	N	Y	N	N	N	Y	N	N	N	Y
Controls	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y
Adj R-Squared	0.428	0.496	0.515	0.525	0.418	0.479	0.554	0.565	0.364	0.499	0.568	0.598
Observations	1,052	1,047	1,011	942	1,052	1,047	1,011	942	1,010	1,006	968	896

Panel 2: AI measure from job postings data

	Δ Log Sales				Δ Log Employment				Δ Log Market Value			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
Δ Share AI Workers	0.160***	0.141**	0.143*	0.156*	0.127**	0.101*	0.110	0.124	0.189***	0.112	0.111	0.134*
	(0.045)	(0.056)	(0.082)	(0.084)	(0.050)	(0.061)	(0.076)	(0.083)	(0.068)	(0.069)	(0.077)	(0.077)
NAICS2 FE	Y	N	N	N	Y	N	N	N	Y	N	N	N
NAICS3 FE	N	Y	N	N	N	Y	N	N	N	Y	N	N
NAICS4 FE	N	N	Y	N	N	N	Y	N	N	N	Y	N
NAICS5 FE	N	N	N	Y	N	N	N	Y	N	N	N	Y
Controls	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y
Adj R-Squared	0.452	0.532	0.578	0.615	0.526	0.624	0.654	0.670	0.461	0.570	0.647	0.686
Observations	935	927	886	827	935	927	886	827	903	896	852	790

Table A8. AI Investments (Including External AI Software) and Firm Growth

This table reports the coefficients from long-differences regressions of firm growth from 2010 to 2018 on the contemporaneous firm-level changes in AI investments among U.S. public firms (in non-tech sectors). The main independent variable is the growth in the share of AI-utilizing workers (including both workers with AI skills and workers with jobs referencing external AI software, such as IBM Watson) from 2010 to 2018, standardized to mean zero and standard deviation of one. We consider three measures of firm growth: changes in log sales (columns 1 and 2), changes in log employment (columns 3 and 4), and changes in log market value (columns 5 and 6). Regressions are weighted by the number of Cognism resumes in 2010. All specifications control for industry sector fixed effects. Columns 2, 4, and 6 also include the baseline controls all measured as of 2010: firm-level characteristics (log sales, cash/assets, R&D/sales, log markup, and log number of jobs—resumes in Panel 1 and job postings in Panel 2), log industry wage, and characteristics of the commuting zones where the firms are located (the share of workers in IT-related occupations, the share of college-educated workers, log average wage, the share of foreign-born workers, the share of routine workers, the share of workers in finance and manufacturing industries, and the share of female workers). Standard errors are clustered at the 5-digit NAICS industry level and reported in parentheses. *, **, and *** denote statistical significance at the 10%, 5%, and 1% levels, respectively.

	Δ Log Sales		Δ Log Employment		Δ Log Market Value	
	(1)	(2)	(3)	(4)	(5)	(6)
Δ Share AI Workers	0.242***	0.228***	0.289***	0.265***	0.285***	0.279***
	(0.056)	(0.050)	(0.081)	(0.062)	(0.075)	(0.063)
Industry FE	Y	Y	Y	Y	Y	Y
Controls	N	Y	N	Y	N	Y
Adj R-Squared	0.245	0.445	0.268	0.448	0.253	0.397
Observations	1,023	1,023	1,023	1,023	994	994

Table A9. AI Investments (Continuous AI Measure) and Firm Growth

This table reports the coefficients from long-differences regressions of growth of U.S. public firms (in non-tech sectors) from 2010 to 2018 on the contemporaneous changes in the average job-level continuous AI measure across all Burning Glass job postings of each firm. The independent variable is the change in the firm-level average continuous AI-relatedness measures from 2010 to 2018, standradized to mean zero and standard deviation of one. Panel 1 presents results from the continuous measure based on all skills (ω_j^{AI}), and Panel 2 shows results from the continuous measure based on narrow AI skills ($\omega_j^{NarrowAI}$). The continuous measure based on narrow AI skills removes the impact of general programming or statistics skills not specific to AI (e.g., Python or linear regression). See Section 3.1 for detailed descriptions of these measures. We consider three measures of firm growth: changes in log sales (columns 1 and 2), changes in log employment (columns 3 and 4), and changes in log market value (columns 5 and 6). Regressions are weighted by the number of Burning Glass job postings in 2010. All specifications control for industry sector fixed effects. Columns 2, 4, and 6 also include the baseline controls all measured as of 2010: firm-level characteristics (log sales, cash/assets, R&D/sales, log markup, log number of jobs—resumes in Panel 1 and job postings in Panel 2), log industry wage, as well as characteristics of the commuting zones where the firms are located (the share of workers in IT-related occupations, the share of college-educated workers, log average wage, the share of foreign-born workers, the share of routine workers, the share of workers in finance and manufacturing industries, and the share of female workers). Standard errors are clustered at the 5–digit NAICS industry level and reported in parentheses. *, **, and *** denote statistical significance at the 10%, 5%, and 1% levels, respectively.

Panel 1: Continuous All-skill AI Measure

	Δ Log Sales		Δ Log Employment		Δ Log Market Value	
	(1)	(2)	(3)	(4)	(5)	(6)
Δ Share AI Workers	0.143**	0.177***	0.144	0.130*	0.132	0.197**
	(0.069)	(0.053)	(0.093)	(0.069)	(0.095)	(0.076)
Industry FE	Y	Y	Y	Y	Y	Y
Controls	N	Y	N	Y	N	Y
Adj R-Squared	0.254	0.432	0.306	0.521	0.324	0.446
Observations	935	935	935	935	903	903

Panel 2: Continuous Narrow AI measure

	Δ Log Sales		Δ Log Employment		Δ Log Market Value	
	(1)	(2)	(3)	(4)	(5)	(6)
Δ Share AI Workers	0.147***	0.174***	0.154*	0.133***	0.139	0.202***
	(0.057)	(0.041)	(0.079)	(0.050)	(0.086)	(0.067)
Industry FE	Y	Y	Y	Y	Y	Y
Controls	N	Y	N	Y	N	Y
Adj R-Squared	0.271	0.443	0.317	0.525	0.335	0.459
Observations	935	935	935	935	903	903

Table A10. AI Investments (Alternative Cutoffs for AI Jobs in Burning Glass) and Firm Growth

This table reports the coefficients from long-differences regressions of firm growth from 2010 to 2018 on the contemporaneous changes in the share of AI job postings of U.S. public firms (in non-tech sectors). AI job postings are defined as job postings with continuous job-level measure ω_j^{AI} above 0.05 in Panel 1, and job postings with continuous job-level measure ω_j^{AI} above 0.15 in Panel 2. See Section 3.1 for the detailed description of the methodology. The independent variable is standardized to mean zero and standard deviation of one. We consider three measures of firm growth: changes in log sales (columns 1 and 2), changes in log employment (columns 3 and 4), and changes in log market value (columns 5 and 6). Regressions are weighted by the number of Burning Glass job postings in 2010. All regressions control for industry sector fixed effects. Columns 2, 4, and 6 also include the baseline controls all measured as of 2010: firm-level characteristics (log sales, cash/assets, R&D/sales, log markup, and log number of jobs—resumes in Panel 1 and job postings in Panel 2), log industry wage, and characteristics of the commuting zones where the firms are located (the share of workers in IT-related occupations, the share of college-educated workers, log average wage, the share of foreign-born workers, the share of routine workers, the share of workers in finance and manufacturing industries, and the share of female workers). Standard errors are clustered at the 5-digit NAICS industry level and reported in parentheses. *, **, and *** denote statistical significance at the 10%, 5%, and 1% levels, respectively.

Panel 1: Cutoff = 0.05

	Δ Log Sales		Δ Log Employment		Δ Log Market Value	
	(1)	(2)	(3)	(4)	(5)	(6)
Δ Share AI Workers	0.142**	0.168***	0.153*	0.138***	0.132	0.192***
	(0.058)	(0.042)	(0.080)	(0.049)	(0.087)	(0.067)
Industry FE	Y	Y	Y	Y	Y	Y
Controls	N	Y	N	Y	N	Y
Adj R-Squared	0.268	0.454	0.317	0.528	0.332	0.460
Observations	935	935	935	935	903	903

Panel 2: Cutoff = 0.15

	Δ Log Sales		Δ Log Employment		Δ Log Market Value	
	(1)	(2)	(3)	(4)	(5)	(6)
Δ Share AI Workers	0.146**	0.140**	0.154*	0.106*	0.149	0.176**
	(0.065)	(0.056)	(0.085)	(0.058)	(0.092)	(0.082)
Industry FE	Y	Y	Y	Y	Y	Y
Controls	N	Y	N	Y	N	Y
Adj R-Squared	0.261	0.437	0.311	0.520	0.334	0.453
Observations	935	935	935	935	903	903

Table A11. AI Investments and Firm Growth: Controlling for Industry- and Firm-level Pre-trends

This table reports the coefficients from long-differences regressions of firm growth from 2010 to 2018 on the contemporaneous firm-level changes in AI investments among U.S. public firms (in non-tech sectors). In this table, we control for past industry and firm growth, which helps address the concern that AI-investing firms might already be on higher growth trajectories prior to AI investments. We consider three measures of firm growth: changes in log sales (columns 1 and 2), changes in log employment (columns 3 and 4), and changes in log market value (columns 5 and 6). The main independent variable is the growth in the share of AI workers from 2010 to 2018, standardized to mean zero and standard deviation of one. Panel 1 considers the resume-based measure of the share of AI workers, while Panel 2 looks at the job-posting-based measure. Regressions are weighted by the number of Cognism resumes in 2010 in Panel 1 and the number of Burning Glass job postings in 2010 in Panel 2. All specifications control for industry sector fixed effects and include the baseline controls all measured as of 2010: firm-level characteristics (log sales, cash/assets, R&D/sales, log markup, and log number of jobs—resumes in Panel 1 and job postings in Panel 2), log industry wage, and characteristics of the commuting zones where the firms are located (the share of workers in IT-related occupations, the share of college-educated workers, log average wage, the share of foreign-born workers, the share of routine workers, the share of workers in finance and manufacturing industries, and the share of female workers). Additionally, columns 1, 3, and 5 control for industry-level growth in sales and employment from 2000 to 2008 (at the 5-digit NAICS level), and columns 2, 4, and 6 control for firm-level growth in sales and employment from 2000 to 2008. Standard errors are clustered at the 5-digit NAICS industry level and reported in parentheses. *, **, and *** denote statistical significance at the 10%, 5%, and 1% levels, respectively.

Panel 1: AI measure from resume data

	Δ Log Sales		Δ Log Employment		Δ Log Market Value	
	(1)	(2)	(3)	(4)	(5)	(6)
Δ Share AI Workers	0.196***	0.179***	0.212***	0.201***	0.210***	0.185***
	(0.057)	(0.046)	(0.074)	(0.065)	(0.069)	(0.060)
Industry FE	Y	Y	Y	Y	Y	Y
Controls	Y	Y	Y	Y	Y	Y
Industry pre-trend	Y	N	Y	N	Y	N
Firm pre-trend	N	Y	N	Y	N	Y
Observations	1,004	815	1,004	815	962	789

Panel 2: AI measure from job postings data

	Δ Log Sales		Δ Log Employment		Δ Log Market Value	
	(1)	(2)	(3)	(4)	(5)	(6)
Δ Share AI Workers	0.144***	0.132***	0.118**	0.158***	0.168***	0.149***
	(0.042)	(0.039)	(0.047)	(0.040)	(0.062)	(0.055)
Industry FE	Y	Y	Y	Y	Y	Y
Controls	Y	Y	Y	Y	Y	Y
Industry pre-trend	Y	N	Y	N	Y	N
Firm pre-trend	N	Y	N	Y	N	Y
Observations	897	738	897	738	865	718

Table A12. AI Investments and Firm Growth: Controlling for State FE and Tobin's q

This table reports the coefficients from long-differences regressions of firm growth from 2010 to 2018 on the contemporaneous firm-level changes in AI investments among U.S. public firms (in non-tech sectors), with additional controls for state FE and Tobin's q. We consider three measures of growth: changes in log sales (columns 1 and 2), changes in log employment (columns 3 and 4), and changes in log market value (columns 5 and 6). The main independent variable is the growth in the share of AI workers from 2010 to 2018, standardized to mean zero and standard deviation of one. Panel 1 considers the resume-based measure of the share of AI workers, while Panel 2 looks at the job-posting-based measure. Regressions are weighted by the number of Cognism resumes in 2010 in Panel 1 and the number of Burning Glass job postings in 2010 in Panel 2. All specifications control for industry sector fixed effects, state fixed effects, and the baseline controls all measured as of 2010: firm-level characteristics (log sales, cash/assets, R&D/sales, log markup, and log number of jobs—resumes in Panel 1 and job postings in Panel 2), log industry wage, and characteristics of the commuting zones where the firms are located (the share of workers in IT-related occupations, the share of college-educated workers, log average wage, the share of foreign-born workers, the share of routine workers, the share of workers in finance and manufacturing industries, and the share of female workers). Columns 2, 4, and 6 also control for Tobin's Q in 2010, defined as market value of assets divided by book value of assets. Standard errors are clustered at the 5-digit NAICS industry level and reported in parentheses. *, **, and *** denote statistical significance at the 10%, 5%, and 1% levels, respectively.

Panel 1: AI measure from resume data

	Δ Log Sales		Δ Log Employment		Δ Log Market Value	
	(1)	(2)	(3)	(4)	(5)	(6)
Δ Share AI Workers	0.178***	0.156***	0.181***	0.160***	0.196***	0.172***
	(0.059)	(0.043)	(0.069)	(0.053)	(0.068)	(0.051)
Tobin's Q 2010		0.223***		0.231***		0.246***
		(0.065)		(0.063)		(0.077)
Industry FE	Y	Y	Y	Y	Y	Y
Controls	Y	Y	Y	Y	Y	Y
State FE	Y	Y	Y	Y	Y	Y
NAICS3 FE	N	N	N	N	N	N
Adj R-Squared	0.492	0.551	0.511	0.563	0.425	0.474
Observations	1,046	1,004	1,046	1,004	1,001	1,000

Panel 2: AI measure from job postings data

	Δ Log Sales		Δ Log Employment		Δ Log Market Value	
	(1)	(2)	(3)	(4)	(5)	(6)
Δ Share AI Workers	0.139***	0.108***	0.155***	0.122***	0.143**	0.110**
	(0.050)	(0.036)	(0.053)	(0.038)	(0.065)	(0.049)
Tobin's Q 2010		0.188***		0.194***		0.214***
		(0.043)		(0.048)		(0.062)
Industry FE	Y	Y	Y	Y	Y	Y
Controls	Y	Y	Y	Y	Y	Y
State FE	Y	Y	Y	Y	Y	Y
NAICS3 FE	N	N	N	N	N	N
Adj R-Squared	0.547	0.581	0.681	0.701	0.568	0.593
Observations	927	896	927	896	893	892

Table A13. Past AI Investments and Future Firm Growth

This table reports the coefficients from a predictive regression of firm growth during the later part of our sample (2015–2020) on growth in AI investments during the earlier part of the sample (2010-2015) among U.S. public firms (in non-tech sectors). The dependent variables are changes in log sales in columns 1 and 2, and changes in log employment in columns 3 and 4. The main independent variable is the growth in the share of AI workers from 2010 to 2015, standardized to mean zero and standard deviation of one. Panel 1 considers the resume-based measure of the share of AI workers, while Panel 2 looks at the job-posting-based measure. Regressions are weighted by the number of Cognism resumes in 2010 in Panel 1 and the number of Burning Glass job postings in 2010 in Panel 2. All specifications control for industry sector fixed effects. Columns 2 and 4 also include the baseline controls all measured as of 2010: firm-level characteristics (log sales, cash/assets, R&D/sales, log markup, and log number of jobs—resumes in Panel 1 and job postings in Panel 2), log industry wage, and characteristics of the commuting zones where the firms are located (the share of workers in IT-related occupations, the share of college-educated workers, log average wage, the share of foreign-born workers, the share of routine workers, the share of workers in finance and manufacturing industries, and the share of female workers). Standard errors are clustered at the 5-digit NAICS industry level and reported in parentheses. *, **, and *** denote statistical significance at the 10%, 5%, and 1% levels, respectively.

Panel 1: AI measure from resume data

	Δ Log Sales		Δ Log Employment	
	(1)	(2)	(3)	(4)
Δ Share AI Workers	0.128**	0.084**	0.149*	0.088*
	(0.053)	(0.041)	(0.079)	(0.052)
Industry FE	Y	Y	Y	Y
Controls	N	Y	N	Y
Adj R-Squared	0.250	0.372	0.211	0.357
Observations	990	990	982	982

Panel 2: AI measure from job postings data

	Δ Log Sales		Δ Log Employment	
	(1)	(2)	(3)	(4)
Δ Share AI Workers	0.109**	0.116***	0.144**	0.151***
	(0.051)	(0.031)	(0.073)	(0.048)
Industry FE	Y	Y	Y	Y
Controls	N	Y	N	Y
Adj R-Squared	0.179	0.327	0.150	0.353
Observations	875	875	868	868

Table A14. AI Investments and Firm Growth: Controlling for Other Technologies

This table reports the coefficients from long-differences regressions of firm growth from 2010 to 2018 on the contemporaneous changes in AI investments among U.S. public firms (in non-tech sectors), controlling for investments in other (non-AI-specific) technologies measured using the Burning Glass data. Columns 1 and 2 control for the 2010-2018 change in the firm's share of non-AI IT jobs, columns 3 and 4 control for the change in the share of robot-related jobs, columns 5 and 6 control for the change in the share of non-AI data-related jobs, and columns 7 and 8 control for the changs in the share of non-AI data analysis jobs. An IT job is defined as a job for which at least 10% of the required skills are in the "Information Technology" skill cluster; a robot-related job is a job with a robot relatedness score (constructed with the same methodology as the AI-relatedness score but using the core skill of "Robotics") above 0.1; a data-related job is a job with at least 10% of required skills in data-related skill clusters; a data analysis job is a job with at least 10% of required skills in the "Analysis" skill cluster. All measures are standardized to mean zero and standard deviation of one. Panel 1 considers the resume-based measure of the share of AI workers, while Panel 2 looks at the job-posting-based measure. Regressions are weighted by the number of Cognism resumes in 2010 in Panel 1 and the number of Burning Glass job postings in 2010 in Panel 2. All regressions control for industry sector fixed effects and the baseline controls all measured as of 2010: firm-level characteristics (log sales, cash/assets, R&D/sales, log markup, and log number of jobs—resumes in Panel 1 and job postings in Panel 2), log industry wage, and characteristics of the commuting zones where the firms are located (the share of workers in IT-related occupations, the share of college-educated workers, log average wage, the share of foreign-born workers, the share of routine workers, the share of workers in finance and manufacturing industries, and the share of female workers). Standard errors are clustered at the 5-digit NAICS industry level and reported in parentheses. *, **, and *** denote statistical significance at the 10%, 5%, and 1% levels, respectively.

Panel 1: AI measure from resume data

	Δ Log Sales	Δ Log Employment	Δ Log Sales	Δ Log Employment	Δ Log Sales	Δ Log Employment	Δ Log Sales	Δ Log Employment
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Δ Share AI Workers	0.207***	0.217***	0.208***	0.219**	0.192***	0.203***	0.195***	0.207***
	(0.053)	(0.075)	(0.063)	(0.084)	(0.049)	(0.070)	(0.053)	(0.074)
Δ Share Non-AI IT Workers	0.138***	0.114**						
	(0.051)	(0.045)						
Δ Share Robot Workers			-0.016	-0.016				
			(0.032)	(0.040)				
Δ Share Non-AI Data Workers					0.136***	0.131***		
					(0.037)	(0.034)		
Δ Share Non-AI Data Analysis Workers							0.073**	0.068**
							(0.030)	(0.031)
Industry FE	Y	Y	Y	Y	Y	Y	Y	Y
Controls	Y	Y	Y	Y	Y	Y	Y	Y
Adj R-Squared	0.456	0.430	0.434	0.419	0.465	0.441	0.443	0.425
Observations	970	970	970	970	970	970	970	970

Panel 2: AI measure from job postings data

	Δ Log Sales	Δ Log Employment	Δ Log Sales	Δ Log Employment	Δ Log Sales	Δ Log Employment	Δ Log Sales	Δ Log Employment
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Δ Share AI Workers	0.158***	0.119**	0.154***	0.128***	0.160***	0.120***	0.165***	0.128***
	(0.039)	(0.047)	(0.038)	(0.047)	(0.039)	(0.045)	(0.040)	(0.047)
Δ Share Non-AI IT Workers	0.138***	0.178***						
	(0.052)	(0.057)						
Δ Share Robot Workers			0.071	0.015				
			(0.048)	(0.049)				
Δ Share Non-AI Data Workers					0.054	0.093		
					(0.064)	(0.072)		
Δ Share Non-AI Data Analysis Workers							0.019	0.022
							(0.050)	(0.052)
Industry FE	Y	Y	Y	Y	Y	Y	Y	Y
Controls	Y	Y	Y	Y	Y	Y	Y	Y
Adj R-Squared	0.459	0.540	0.445	0.525	0.443	0.528	0.441	0.525
Observations	935	935	935	935	935	935	935	935

Table A15. AI Investments and Firm Growth: IV Estimates Using Job Postings Data

This table estimates the relationship between AI investments and firm growth from 2010 to 2018 for U.S. public firms (in non-tech sectors), where firms' AI investments are instrumented with firm-level ex-ante exposure to AI-skilled graduates from AI-strong universities. The independent variable is the change in the share of AI jobs from 2010 to 2018 based on the job postings data. Regressions are weighted by the number of Burning Glass job postings in 2010. The independent variable and the instrument are standardized to mean zero and standard deviation of one. We consider changes in log sales in columns 1 to 4, log employment in columns 5 to 8, and log market value in columns 9 to 12. All specifications control for industry sector fixed effects and ex-ante exposure to universities that are strong in computer science research. Columns 2–4, 6–8, and 10–12 also include the baseline controls all measured as of 2010: firm-level characteristics (log sales, cash/assets, R&D/sales, log markup, and log number of jobs—resumes in Panel 1 and job postings in Panel 2), log industry wage, and characteristics of the commuting zones where the firms are located (the share of workers in IT-related occupations, the share of college-educated workers, log average wage, the share of foreign-born workers, the share of routine workers, the share of workers in finance and manufacturing industries, and the share of female workers). Columns 3, 4, 7, 8, 11, and 12 additionally control for firm-level changes in log sales and log employment from 2000 to 2008. Columns 4, 8, and 12 control for state fixed effects. Standard errors are clustered at the 5-digit NAICS industry level and reported in parentheses. The first-stage F-statistics of the instrument are reported for all specifications. *, **, and *** denote statistical significance at the 10%, 5%, and 1% levels, respectively.

	Δ Log Sales				Δ Log Employment				Δ Log Market Value			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
Δ Share AI Workers	0.154	0.247**	0.230***	0.093	0.440**	0.436***	0.234*	0.048	0.100	0.160	0.210*	0.119
	(0.099)	(0.103)	(0.078)	(0.097)	(0.207)	(0.140)	(0.122)	(0.138)	(0.164)	(0.152)	(0.114)	(0.138)
Industry FE	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y
CS Control	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y
Baseline Controls	N	Y	Y	Y	N	Y	Y	Y	N	Y	Y	Y
Control Pre-trend	N	N	Y	Y	N	N	Y	Y	N	N	Y	Y
State FE	N	N	N	Y	N	N	N	Y	N	N	N	Y
F Statistic	9.2	9.4	10.9	12.5	9.2	9.4	10.9	12.5	9.2	9.4	10.9	12.5
Observations	889	889	702	696	889	889	702	696	860	860	684	676

Table A16. AI Investments and Product Innovation: IV Estimates Using Resume Data

This table estimates the relationship between AI investments and product innovation from 2010 to 2018 for U.S. public firms (in non-tech sectors), where firms' AI investments are instrumented with firm-level ex-ante exposure to AI-skilled graduates from AI-strong universities. The independent variable is the change in the share of AI workers from 2010 to 2018 based on the resume data. The independent variable and the instrument are standardized to mean zero and standard deviation of one. Regressions are weighted by the number of Cognism resumes in 2010. We consider the change in log(1+number of trademarks) in columns 1 to 4, the change in log(1+number of product patents) in columns 5 to 8, and the change in the product mix in columns 9 to 12. Product patents are patents with over 50% of the claims being product claims, following the categorization in Ganglmair et al. (2021). The change in the product mix is measured as the sum of annual changes between from 2010 to 2018, where each annual change is the angle between the two word vectors indicating firms' product offerings in that year and the previous year, following Hoberg et al. (2014). All specifications control for industry sector fixed effects and ex-ante exposure to universities that are strong in computer science research. Columns 2–4, 6–8, and 10–12 also include the baseline controls all measured as of 2010: firm-level characteristics (log sales, cash/assets, R&D/sales, log markup, and log number of jobs—resumes in Panel 1 and job postings in Panel 2), log industry wage, and characteristics of the commuting zones where the firms are located (the share of workers in IT-related occupations, the share of college-educated workers, log average wage, the share of foreign-born workers, the share of routine workers, the share of workers in finance and manufacturing industries, and the share of female workers). Columns 3, 4, 7, 8, 11, and 12 additionally control for firm-level changes in log sales and log employment from 2000 to 2008. Columns 4, 8, and 12 add state fixed effects. Standard errors are clustered at the 5-digit NAICS industry level, and reported in parentheses. The first-stage F-statistics of the instrument are reported for all specifications. *, **, and *** denote statistical significance at the 10%, 5%, and 1% levels, respectively.

	Δ Log Number of Trademarks				Δ Log Number of Product Patents				Change in Product Mix			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
Δ Share AI Workers	0.081	0.367	0.435*	0.429**	0.175	0.317*	0.472***	0.466**	0.189	0.180	0.179	0.444**
	(0.267)	(0.351)	(0.261)	(0.212)	(0.174)	(0.170)	(0.159)	(0.200)	(0.272)	(0.228)	(0.205)	(0.205)
Industry FE	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y
CS Control	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y
Baseline Controls	N	Y	Y	Y	N	Y	Y	Y	N	Y	Y	Y
Control Pre-trend	N	N	Y	Y	N	N	Y	Y	N	N	Y	Y
State FE	N	N	N	Y	N	N	N	Y	N	N	N	Y
F Statistic	14.0	15.2	15.8	24.0	14.3	21.5	41.1	40.0	15.0	15.7	13.1	19.4
Observations	528	528	435	426	586	586	479	469	932	932	725	717

Table A17. AI Investments and Productivity of Early Adopters

This table reports the coefficients from long-differences regressions of changes in firm productivity from 2010 to 2018 on changes in AI investments by U.S. public firms (in non-tech sectors) from 2010 to 2014. We consider two measures of productivity: log sales per worker in columns 1 and 2 and revenue TFP in columns 3 and 4. Revenue TFP is the residual from regressing log revenue on log employment and log capital (constructed using the perpetual inventory method), with separate regressions for each industry sector. The main independent variable is growth in the share of AI workers from 2010 to 2014, calculated based on resumes in Panel 1 and based on job postings in Panel 2. All independent variables are standardized to mean zero and standard deviation of one. Regressions are weighted by the number of Cognism resumes in 2010 in Panel 1 and the number of Burning Glass job postings in 2010 in Panel 2. All specifications control for industry sector fixed effects. Columns 2 and 4 also include the baseline controls all measured as of 2010: firm-level characteristics (log sales, cash/assets, R&D/sales, log markup, and log number of jobs—resumes in Panel 1 and job postings in Panel 2), log industry wage, and characteristics of the commuting zones where the firms are located (the share of workers in IT-related occupations, the share of college-educated workers, log average wage, the share of foreign-born workers, the share of routine workers, the share of workers in finance and manufacturing industries, and the share of female workers). Standard errors are clustered at the 5-digit NAICS industry level and reported in parentheses. *, **, and *** denote statistical significance at the 10%, 5%, and 1% levels, respectively.

Panel 1: AI measure from resume data

	Δ Log Sales per Worker		Δ Revenue TFP	
	(1)	(2)	(3)	(4)
Δ Share AI Workers 2010-2014	-0.075	-0.046	-0.041	-0.018
	(0.065)	(0.042)	(0.053)	(0.039)
Industry FE	Y	Y	Y	Y
Controls	N	Y	N	Y
Adj R-Squared	0.225	0.383	0.211	0.344
Observations	1,033	1,033	961	961

Panel 2: AI measure from job postings data

	Δ Log Sales per Worker		Δ Revenue TFP	
	(1)	(2)	(3)	(4)
Δ Share AI Workers 2010-2014	0.010	0.045	0.026	0.045
	(0.070)	(0.042)	(0.055)	(0.037)
Industry FE	Y	Y	Y	Y
Controls	N	Y	N	Y
Adj R-Squared	0.368	0.557	0.287	0.426
Observations	929	929	869	869

Table A18. AI Investments and Industry Growth within a Balanced Panel of Firms

This table reports the coefficients from industry-level long-differences regressions of the changes in total industry sales and employment on contemporaneous changes in AI investments for a balanced panel of firms existing in both 2010 and 2018. Each observation is a 5-digit NAICS industry, and we exclude tech sectors. The independent variable is the change in the share of AI workers from 2010 to 2018, standardized to mean zero and standard deviation of one. Panel 1 considers the resume-based measure of the share of AI workers, while Panel 2 looks at the job-posting-based measure. Regressions are weighted by the total (industry-level) number of Cognism resumes in 2010 in Panel 1 and the total (industry-level) number of Burning Glass job postings in 2010 in Panel 2. The dependent variables are the changes, from 2010 to 2018, in log total sales (not including entrants and exits between 2010 and 2018) in columns 1 and 2 and in log total employment in columns 3 and 4. All measures are calculated using Compustat firms that exist at the beginning (2010) and the end (2018) of our sample. All specifications control for industry sector fixed effects. Regressions in columns 2 and 4 also include industry-level controls for log total employment, log total sales, and log average wage in 2010. Standard errors are robust against heteroskedasticity and reported in parentheses. *, **, and *** denote statistical significance at the 10%, 5%, and 1% levels, respectively.

Panel 1: AI measure from resume data

	Δ Sales		Δ Employment	
	(1)	(2)	(3)	(4)
Δ Share AI Workers	0.171***	0.188***	0.194***	0.211***
	(0.058)	(0.045)	(0.074)	(0.062)
Industry Sector FE	Y	Y	Y	Y
Controls	N	Y	N	Y
Observations	275	275	275	275

Panel 2: AI measure from job postings data

	Δ Sales		Δ Employment	
	(1)	(2)	(3)	(4)
Δ Share AI Workers	0.157**	0.138**	0.161	0.142
	(0.073)	(0.064)	(0.116)	(0.108)
Industry Sector FE	Y	Y	Y	Y
Controls	N	Y	N	Y
Observations	261	261	261	261