AI's Job Shakeup: Analyzing the Uneven Impact of AI Adoption on Labor Demand

Bin Gu, Mohammad Soltaniehha, and Xiupeng Wang[1]

March, 2026

[Preliminary, please do not circulate.]

Abstract: This paper studies how direct AI adoption, rather than AI skill hiring alone, reshapes firm-level labor demand. We construct a novel measure of operational AI adoption from 10-K filings using natural language processing, large language models, and a neural-network classifier, combine it with Webb's (2019) occupation-level AI exposure score, and link it to Lightcast job postings for Compustat firms from 2010 to 2022. Using a long difference design over one to five year horizons, we find that job postings in high exposure occupations decline over time following AI adoption, reaching -0.14 percent at the two-year horizon and -0.30 percent at the five-year horizon, with the richer specification reaching -0.32 percent at five years. Low-exposure occupations remain close to zero throughout. A falsification test confirms such a pattern is specific to operational AI deployment rather than broader AI engagement. Heterogeneity analysis shows that the negative hiring response is concentrated in non-IT firms, is substantially stronger in manufacturing than in services, and is more pronounced among firms that adopt externally developed AI than among firms that build AI in-house. We also find that AI adoption is positively associated with labor productivity in the full sample and in both large and small firms. Taken together, the results suggest that AI adoption raises efficiency while selectively reducing demand in the occupations most exposed to automation. The paper contributes a new text-based measure of operational AI adoption, distinguishes buy versus build pathways, and documents heterogeneity in labor-demand adjustment across occupations and firm types.

Key words: AI adoption, Labor Demand, Labor Productivity

---

[1] Gu: Boston University, bgu@bu.edu; Soltaniehha, Boston University, msoltani@bu.edu; Wang, Northeastern University, xiupeng.j.wang@gmail.com.
The authors thank Wang Jin, Daniel Rock, and Michael Webb for sharing their data.

1. Introduction

Artificial intelligence (AI) is widely regarded as one of the most consequential technological developments of the modern era and a leading candidate for a general-purpose technology — one that transforms production, organizational processes, and labor demand across much of the economy (Brynjolfsson et al., 2021; Eloundou et al., 2023). Unlike earlier waves of automation concentrated in routine manual or clerical tasks, modern AI can perform or assist with a wider range of cognitive activities, including prediction, language processing, and decision support (Brynjolfsson et al., 2018). This broader scope has intensified interest in how AI will affect hiring within firms: which occupations will see declining demand, which will remain insulated, and how quickly adjustment will unfold. A growing body of research has begun to investigate these questions from the perspective of firm performance (Alekseeva et al., 2021; Babina et al., 2024) and labor markets more broadly (Acemoglu et al., 2022; Babina et al., 2022; Bessen, 2018). Yet despite this progress, we still know relatively little about how firms' operational use of AI — as distinct from their hiring of AI talent — reshapes labor demand.

Much of the existing literature infers firm-level AI activity from AI-related hiring. Acemoglu et al. (2022) use the share of job postings requiring AI skills as a proxy for AI exposure. Babina et al. (2024) track inflows of workers with AI skills, while Alekseeva et al. (2021) distinguish between AI talent deployed in IT roles and AI talent deployed in managerial roles. These approaches are informative, but they primarily capture firms that build AI capabilities internally. They are less suited to identifying firms that adopt AI tools developed elsewhere — a distinction that matters increasingly as AI matures and as the economy specializes. A pioneering AI firm is likely to maintain its own AI talent pool to develop proprietary applications. But as AI technologies have become more commoditized, many firms

now purchase and deploy AI developed by others without building substantial in-house human capital. Evidence from AI assistants in customer support, for example, shows that adopting firms can realize substantial productivity changes even when the underlying technology is vendor-supplied (Brynjolfsson et al., 2025). This literature suggests that build and buy channels should be distinguished rather than conflated.

This paper introduces a new disclosure-based measure of firm-level AI adoption designed to help separate those channels. We use text mining and large-language-model classification to identify paragraphs in annual 10-K filings that describe operational AI deployment, rather than the development or marketing of AI products. Specifically, we create a hybrid labeled dataset that combines synthetic and manually validated paragraphs, train a neural-network classifier on sentence-transformer embeddings to identify AI-related text, and then use GPT-5.2 to classify those paragraphs into operational AI adoption, AI development, or general mentions. This pipeline processes more than 30 million paragraphs in 10-K filings of Compustat firms over 2010 to 2022, identifying approximately 92,000 AI-related paragraphs and isolating disclosures that describe firms' use of AI in internal operations. The resulting measure — the cumulative count of operational AI paragraphs (CumAI) — captures the breadth of disclosed operational deployment over time. Because the Ops/Dev split distinguishes operational adoption from AI development in firms' disclosures, it allows us to compare adopter-oriented and developer-oriented firms within a common framework.

We then link this firm-level measure to occupation-level AI exposure scores from (Webb, 2019) and near-universe online vacancy data from Lightcast. This design allows us to examine whether changes in hiring are concentrated in occupations whose task content is more exposed to AI-related substitution. Our empirical approach uses long-difference specifications over one- to

five-year horizons, a standard approach in the technology-and-labor literature for studying medium-run adjustment while reducing sensitivity to short-run noise (Acemoglu & Restrepo, 2020; Brynjolfsson & Hitt, 2003; Tambe, 2014). Rather than asking whether AI adoption changes hiring on average, we ask whether adoption is associated with differential changes in postings for more- and less-exposed occupations, and whether those relationships vary across sectors and firm types.

Our results point to a clear and selective pattern. At the one-year horizon, we find little evidence that AI adoption affects postings in highly exposed occupations. The relationship becomes increasingly negative over longer horizons, reaching approximately −0.30 to −0.33 by the five-year horizon, while postings in less-exposed occupations remain close to zero. A falsification test that replaces the operational AI measure with total AI mentions, including development and general disclosures, produces no significant effects on hiring in exposed occupations, confirming that the baseline results are driven specifically by operational deployment rather than broader AI engagement.

The heterogeneity analysis shows that this negative hiring response is concentrated in non-IT firms, is substantially stronger in manufacturing than in services, and is more pronounced among firms that emphasize operational adoption rather than AI development. We also find patterns consistent with AI adoption supporting labor productivity gains, suggesting that the reduction in hiring for exposed occupations occurs alongside improved firm performance rather than simple contraction.

The paper makes three contributions. First, it develops a novel firm-level measure of operational AI adoption from regulatory disclosures, complementing hiring-based proxies used in prior work. Second, it links firm-level AI adoption to occupation-level exposure within a

unified framework, showing that the labor-demand response is concentrated where task-based measures would predict it should be. Third, it provides evidence that AI-related labor adjustment is heterogeneous across firms and unfolds gradually over time. Taken together, the findings suggest that AI's labor-market effects are neither immediate nor uniform: they depend on how AI is measured, which occupations are examined, what type of firm is adopting the technology, and the horizon over which adjustment is observed.

The paper is organized as follows: Section 2 reviews literature on measuring firm-level AI adoption, the task-based occupational exposure framework, and labor market adjustments. Section 3 details our data sources and describes a novel firm-level AI adoption measure from 10-K filings. Section 4 introduces the empirical methodology. Section 5 presents the main empirical findings, covering baseline results and extensive heterogeneity analysis. Section 6 concludes with a summary, policy implications, and study limitations.

2. Literature Review

2.1. Measuring Firm-Level AI Adoption and Labor Demand

A central issue in the emerging literature is how to measure firms' engagement with AI. Much of the existing empirical work relies on labor-market proxies, especially AI-related hiring, to infer firm-level adoption. (Acemoglu et al., 2022) classify establishments as AI-exposed based on the compatibility of their existing occupational task structure with AI capabilities, and show that exposed establishments subsequently reduce non-AI hiring and alter the skill composition of their remaining vacancies. (Babina et al., 2024) track the accumulation of AI talent through worker resumes and job postings, and find that firms investing more in AI-skilled workers

experience faster growth in sales, employment, and market valuation — an effect that operates primarily through increased product innovation rather than operational cost savings. Alekseeva et al., (2024) further demonstrate that the organizational location of AI adoption matters: firms intensifying AI adoption significantly increase their demand for managerial roles and shift the required skill profile of those roles toward cognitive and interpersonal competencies such as collaboration, creativity, and data analysis, while reducing demand for routine administrative skills. Although they have established that AI-related hiring is informative about capability building within firms, but it does not fully capture operational use of AI, particularly when firms deploy vendor-supplied tools without developing the technology in-house.

This distinction between internal capability building and operational deployment is increasingly important. Firms may adopt AI through software vendors, cloud-based tools, and packaged applications without making substantial AI-specific hires. Evidence from generative AI in customer support illustrates this point: Brynjolfsson, Li, and Raymond (2023) show that firms can realize meaningful productivity gains from AI tools even when the underlying technology is externally supplied rather than internally developed. This pattern implies that the literature benefits from measures that distinguish firms that build AI from firms that primarily adopt it.

The broader technology literature also suggests caution in interpreting short-run labor-market responses. Bessen (2018) argues that automation does not mechanically reduce employment because cost reductions may expand output demand and sustain labor demand in affected industries. At the same time, work in information systems and productivity economics emphasizes that the impact of new general-purpose technologies often unfolds slowly because firms must accumulate complementary investments in organization, business processes, and

worker capabilities. These arguments imply that both the measurement of AI activity and the time horizon of analysis are likely to shape empirical conclusions.

Our paper builds on this literature by introducing a disclosure-based measure of AI adoption derived from firms' 10-K filings. Relative to hiring-based proxies, this approach is better suited to capturing firms that deploy AI in operations without necessarily hiring specialized AI talent — a distinction that hiring flows alone cannot identify. Our measure also allows us to distinguish between firms that adopt AI as a productive input and firms engaged in AI development, enabling us to examine whether labor-demand responses differ across these two orientations. This distinction has been difficult to study using existing firm-level datasets, yet it is central to understanding how AI reshapes hiring patterns within incumbent firms.

2.2. Occupation-Level Exposure and the Task-Based Framework

A second strand of the literature asks which jobs are most exposed to AI. Here, the dominant approach is task-based: occupations are treated as bundles of tasks, and technology affects labor demand by substituting for some tasks while complementing others. Autor et al., (2003) established this framework in the context of computerization, and Acemoglu and Restrepo (2019, 2020) extend it by distinguishing among displacement, productivity, and reinstatement effects. In their framework, employment declines when automation replaces workers in existing tasks, but employment may be supported or even expanded when productivity gains raise output demand or when new labor-intensive tasks emerge alongside the technology. This framework is especially useful in the AI context because recent AI systems reach well beyond routine manual work and increasingly affect cognitive and non-routine tasks that were once seen as relatively protected (Brynjolfsson et al., 2018).

Several recent studies have developed occupation-level measures of AI exposure. Webb (2019) maps AI-related patents to ONET task descriptions and constructs a continuous exposure score that links the direction of technical change to the task content of occupations. (Felten et al., 2021; Felten et al., 2018) instead connect advances in AI applications — such as image recognition, language processing, and prediction — to occupational abilities in ONET, and construct occupation-, industry-, and geography-level exposure measures. Eloundou et al., (2023) develop an exposure measure tailored to large language models by combining human and GPT-based assessments of the share of occupational tasks that can be accelerated by a ChatGPT-like system. Although these measures differ in construction, they generally suggest that AI exposure is concentrated in cognitive work and does not align neatly with earlier notions of routine-task intensity.

This feature differentiates AI from previous automation waves. Industrial robots, for example, were concentrated in manufacturing and produced displacement effects largely in manual and routine occupations (Acemoglu and Restrepo, 2020). By contrast, recent work on AI and large language models suggests substantial exposure in professional, administrative, and scientific occupations. Webb (2019) shows that AI exposure is relatively high among high-skill occupations, while Felten, Raj, and Seamans (2021) similarly indicate that AI exposure is not confined to traditional routine work. Eloundou et al. (2023) further show that large language models have broad potential reach across the labor market, with especially strong exposure in many higher-wage occupations. Taken together, this literature suggests that AI may reshape labor demand in ways that are broader and more heterogeneous than earlier forms of automation.

This occupation-level heterogeneity is central to our empirical strategy. We combine a firm-level measure of AI adoption with Webb's occupation-level exposure scores and compare

firms' postings in high- and low-exposure occupations. Doing so allows us to move beyond average employment effects and test whether the relationship between AI adoption and labor demand is concentrated precisely where task-based exposure would predict it should be. In that sense, our design links the firm-level adoption literature to the occupation-level exposure literature within a unified empirical framework.

2.3. Heterogeneity Across Sectors and Firm Types

A recurring lesson in the literature is that the labor-market consequences of AI are unlikely to be uniform across firms. Differences in sector, organizational structure, and mode of adoption may all shape whether AI primarily augments labor or substitutes for it. Yet these margins remain underexplored in firm-level evidence.

Sectoral heterogeneity is especially important. Manufacturing and services differ sharply in their task composition, production environment, and adoption margins, which may lead AI to operate differently across sectors. In contrast to earlier automation — which, as discussed above, was concentrated in manufacturing and manual occupations — recent evidence on generative AI has largely focused on service environments: Brynjolfsson, Li, and Raymond (2023) study AI adoption in customer support and find productivity gains that are broadly augmentative rather than substitutive, at least in that context. Whether this pattern extends beyond service settings remains an open question.

Heterogeneity across firm types is equally important. Firms in the IT sector may combine AI adoption with continued demand for engineers, data scientists, and related technical staff, while non-IT firms may be more likely to use AI to automate existing workflows. Firm size may also matter: using Census data, (Brynjolfsson et al., 2023) find that IT investment benefits larger

firms disproportionately by enabling them to replicate operations across establishments, with effects on sales that outpace effects on employment — a pattern that may carry over to AI adoption. These distinctions suggest that the labor-demand consequences of AI could vary systematically with firm size and sector, though they remain less well documented for AI specifically.

Perhaps the most consequential but least directly measured dimension of heterogeneity is the distinction between firms that develop AI and firms that adopt it. Developer firms may experience offsetting labor demand because building, refining, and integrating AI systems requires engineers, researchers, and implementation specialists. Adopter firms that use external AI tools may instead realize labor-saving benefits with far fewer complementary hiring needs. The existing literature has touched parts of this distinction — through AI hiring, workforce composition, or exposure measures — but has not cleanly separated labor-demand responses for developer-oriented and adopter-oriented firms. Our operational-versus-development classification provides a disclosure-based way to examine this margin. While it is not a perfect measure, it offers a tractable way to test whether the labor effects of AI differ depending on whether firms are creating AI technologies or primarily deploying them in operations.

2.4. AI, Productivity, and Adjustment Dynamics

A final strand of the literature examines the productivity effects of AI at the task and worker level. This evidence is important because it clarifies the mechanisms through which AI may reshape labor demand, even though productivity gains do not translate mechanically into employment outcomes.

Recent experimental and quasi-experimental studies consistently show that AI can raise productivity in a range of white-collar settings. (Noy & Zhang, 2023) report that ChatGPT improves both speed and quality in realistic writing tasks and compresses the productivity distribution by benefiting lower-ability workers more. (Peng et al., 2023) show that access to GitHub Copilot allows software developers to complete coding tasks around 56% faster. (Dell'Acqua et al., 2023) find that GPT-4 substantially improves consultant performance on tasks within the model's capability frontier, while documenting negative effects when tasks require judgment beyond it. Across settings, the common pattern is that AI often augments worker performance, but the gains depend on worker skill, task complexity, and the fit between the technology and the task.

These findings matter for firm-level labor demand, but they do not resolve it. In the task-based framework, higher productivity can either reduce labor required per unit of output or raise demand by lowering costs and expanding output. Moreover, adoption often requires complementary investments in organization and worker training, which can delay observable gains and spread adjustment over multiple years (Brynjolfsson et al., 2021; Brynjolfsson & Hitt, 2003; Tambe, 2014). As a result, short-run hiring responses may understate the longer-run effects of AI adoption.

Our paper contributes to this discussion by examining firm-level hiring over one- to five-year horizons. This medium-run perspective is important because it is better aligned with the gradual way firms reorganize work around new technologies. The results we report — little evidence of immediate change, followed by increasingly negative hiring responses in more exposed occupations — are consistent with an adjustment process in which the labor-demand effects of AI emerge over time rather than instantaneously. In this way, the paper connects the

task-level productivity literature to the firm-level labor-demand literature and highlights the importance of considering both exposure and adjustment dynamics when evaluating the employment consequences of AI.

3. Data

3.1. Firm-Level AI Adoption Metrics

3.1.1. Curating a Hybrid Dataset

To analyze the extent of AI adoption at the firm level, we downloaded 10-K forms for S&P 500 companies directly using the SEC API. The raw data, containing text from these filings, underwent a thorough preprocessing pipeline. Initially, the content was extracted from the HTML structure of each document, isolating the text within the body tags. This text was converted to lowercase to maintain consistency and facilitate further processing. We normalized the text by eliminating messy characters and restoring any corrupted Unicode characters to ensure uniformity. Additional steps involved removing extraneous whitespace, new line breaks, and non-alphabetic characters, preserving only letters and standard punctuation (Bird et al., 2009). The resulting clean text was then segmented into manageable "text chunks", each approximately 100-150 words long, with a preference for splitting at natural sentence boundaries. Only chunks containing at least 20 English words were retained to ensure relevance in classification tasks.

Given the lack of an existing labeled dataset to identify AI-related content within the 10-K filings, we devised an innovative approach to create a hybrid dataset. As illustrated in

Figure 1, our LLM-assist & semi-supervised approach consists four major steps: (1) building a synthetic model; (2) Initial classification of 10-K paragraphs using the synthetic model; (3) GPT reclassification and manual review to establish a "ground truth" set of AI-related paragraphs; and (4) Constructing the final hybrid training dataset by combining synthetic and confirmed real-world paragraphs.

In detail, we used OpenAI's GPT-4 model to generate 1,000 "AI" and 1,000 "generic" synthetic text chunks, each tailored to specific GICS sectors and sub-industries (Schick and Schütze 2021). These paragraphs were transformed into 384-dimensional embeddings using the sentence transformer “all-MiniLM-L6-v2” model (Reimers and Gurevych 2019). We then trained an initial binary classifier, utilizing a neural network with two layers (64 and 32 nodes) and dropout layers. Although not expected to excel on real-world data, this model was instrumental in labeling the approximately 30.8 million paragraphs from the 10-K filings, identifying 92,056 as potentially discussing AI.

We manually reviewed a random sample of 100 AI and 100 generic paragraphs to assess these results, revealing a 33% accuracy rate for AI-related content and 100% for generic content. GPT-4 had a 93% agreement with human labels in the AI category and 100% in the generic category. Recognizing the model's tendency to generate false positives in the AI class, we used GPT-4 to reclassify the 10,102 identified paragraphs. This results in 2,129 paragraphs being confirmed as AI-related, which serves as the “ground truth” for the rest of our analysis.

To construct the final training dataset, we created an AI class comprised of 2,632 instances, which included 503 synthetic AI paragraphs initially generated by GPT-4 and the 2,129 GPT-4 classified AI paragraphs extracted from the 10-K filings. For the generic class, we also compiled 2,632 instances, consisting of 503 synthetic generic paragraphs from GPT-4, 1,129

generic paragraphs identified by GPT-4 from the 10,102 initial predictions, and an additional 1,000 generic paragraphs from the 10-K filings that our initial model classified as non-AI with high confidence. The 1,129 generic paragraphs are the first model's false positives, and including them in the final training dataset is crucial as they are more difficult data points to predict accurately. Combining synthetic and real-world data, this hybrid dataset provides the foundation for training a more robust classifier capable of more accurately identifying AI-related content in 10-K filings.

3.1.2. Building an AI Classifier Model

The final classifier was developed using a balanced dataset comprising 2,632 AI-related and 2,632 generic text instances. Each paragraph was transformed into a 384-dimensional vector representation using the "all-MiniLM-L6-v2" sentence transformer, capturing the semantic nuances of the text. We opted for a larger fully connected neural network with this enriched dataset to enhance the model's classification capability. The network architecture began with an initial dense layer containing 256 nodes followed by another dense layer with 8 nodes, utilizing the ReLU activation function and 50% dropout layers. The final layer employed a single node with a sigmoid activation function, outputting a probability score for binary classification, distinguishing between AI and generic text.

The model was compiled with the binary cross-entropy loss function, optimized using the Adam optimizer, and evaluated based on accuracy. During inference, a threshold of 0.5 was applied to the model's output to classify paragraphs. This process yielded 92,056 AI paragraphs out of the 30,885,282 total paragraphs processed, accounting for approximately 0.3% of the dataset. This final classifier thus enables a more reliable identification and quantification of AI-related content within the vast corpus of 10-K filings.

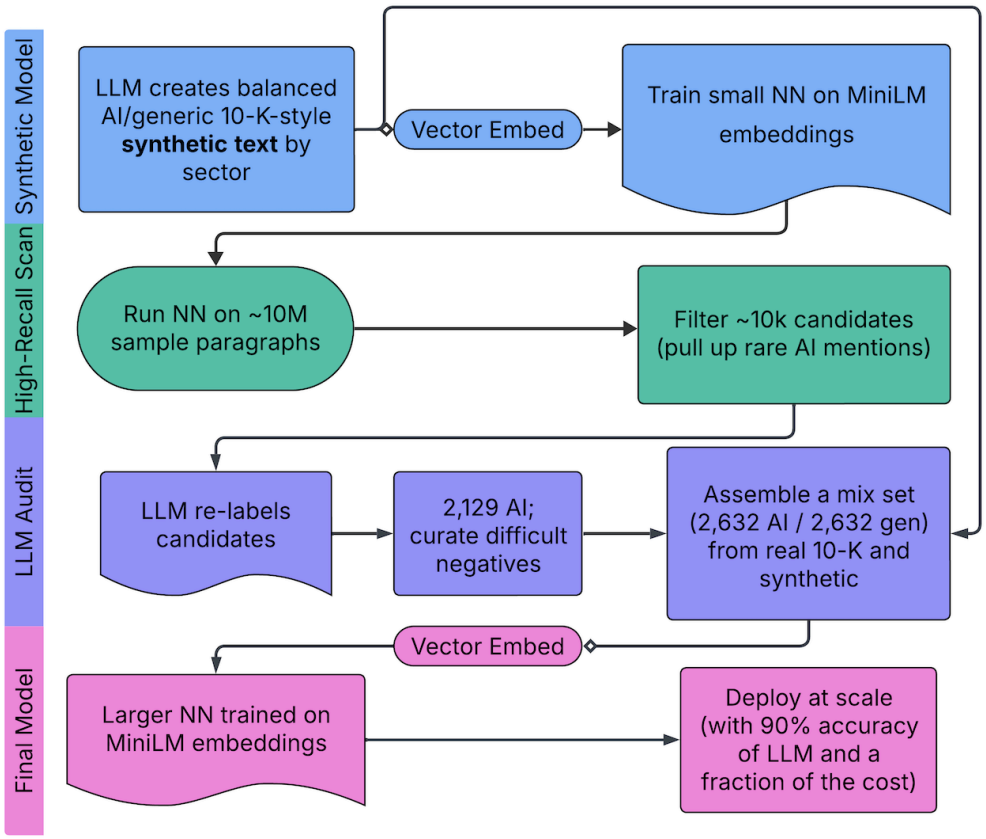


Figure 1. An LLM-assist & semi-supervised approach of measuring AI from 10-Ks

3.1.3. Constructing the AI Adoption Metric

To gain deeper insights into how firms engage with AI, we categorized the identified AI-related paragraphs into three categories, Ops/Dev/None, to introduce a metric based on the number of “Ops” paragraphs within a 10K filing. The "Ops" category captures paragraphs discussing how companies use AI to enhance their internal processes, efficiency, customer engagement, human resources, and productivity. It covers operational efficiency, business intelligence, customer support, HR automation, predictive maintenance, and workflow optimization. The "Dev" category includes paragraphs describing companies' efforts to develop, market, or sell AI technologies as products or services. This encompasses AI products, services, platforms,

research, innovations, patents, and publications related to AI. Paragraphs that did not fit these categories were classified as "None."

The classification was conducted in an unsupervised fashion, using OpenAI's GPT-5.2 model, guided by a main prompt and a system prompt that defined the categories. The LLM analyzed each paragraph's content to determine the most appropriate category, ensuring that each paragraph was assigned to only one category based on its context and specific mentions of AI usage. Of the 92,056 AI-related paragraphs, 7,748 were classified as "Ops," 16,294 as "Dev," and 68,014 as "None." For example, a paragraph discussing a newly developed AI-powered customer service platform was categorized as "Dev," while one describing AI-driven predictive maintenance was classified under "Ops." If the paragraph mentioned AI in a more general or unrelated context, it was classified as "None."

This categorization approach allows us to differentiate between firms' application of AI in operational processes, their development of AI technologies, and general mentions of AI, providing a nuanced understanding of AI's impact on a firm’s activities and potential implications for employment decision-making.

3.2. Occupation Level Automation Scores

To classify occupations based on their inherent exposure to AI-driven automation, we utilize the task-based automation scores developed by Webb (2019). This metric, which maps advancements in artificial intelligence to specific occupational task structures, is effectively designed to differentiate between roles characterized by varying degrees of AI-related substitution risk (Acemoglu et al., 2022). In our empirical framework, we identify high-impact

occupations as those exceeding the median Webb exposure score, while those falling below this threshold are classified as low-impact.

Webb's (2019) occupation-level AI exposure score measures the extent to which an occupation's task content overlaps with the applications described in AI patents. The measure is constructed by extracting verb–noun pairs from AI patent titles and from ONET task descriptions, identifying semantic overlap between the two, and then aggregating these task-level overlaps to the occupation level using ONET task weights. Accordingly, the Webb score captures occupational exposure to patented AI capabilities rather than a direct estimate of realized job displacement. This feature makes it well suited to our setting, as it allows us to test whether firms' operational AI adoption reduces labor demand more strongly in occupations whose task bundles are more closely aligned with historically patented AI applications.

This measure also differs conceptually from other prominent AI exposure indices. Suitability for Machine Learning (SML) index (Brynjolfsson et al., 2018) captures the technical potential for occupational tasks to be performed by machine learning, while Felten et al. (2021) measures links AI progress to the human abilities used in occupations. Eloundou et al. (2023), in turn, focus specifically on the exposure of occupations to large language models. Relative to these alternatives, Webb's measure is better aligned with our question and sample period: because it is anchored in historically patented AI applications, it more directly captures exposure to AI capabilities that were plausibly available for diffusion during 2010–2022, rather than broader technological potential or the more recent capabilities of generative AI.

3.3. Job Posting Data from Lightcast (Formerly Known as Burning Glass Technology)

The Lightcast data covers the near-universe of online job vacancy postings from over 40 thousand distinct online job platforms in the U.S. It is a popular data source for understanding firm-level labor demand and has been widely adopted by scholars (Sarah H. Bana et al., 2024). This study identifies Compustat firms and their job postings from 2010 to 2022. For each firm in each month, we identify the number of job postings in occupations above and below the median Webb exposure score, and we then aggregate those counts to the firm-year level for the Compustat firms in our sample for use in the empirical analysis.

3.4. Compustat Data

For other covariates included in this study, such as total employment and total revenue, we obtained Compustat data from Wharton Research Data Services. Because Compustat firms do not share common firm identifiers with Lightcast data, we utilize the Lightcast-to-Compustat crosswalk from Bai et al. (2021), who undertook a multi-step approach to merge the two databases following a combination of name and address fuzzy matching that was adopted in Campello et al. (2023). Our final sample consists of the Compustat firms that can be linked across Lightcast, Compustat, and our AI adoption measures.

3.5. Summary Statistics

Table 1 presents summary statistics for the linked sample of 18,358 firm-year observations. Panel A describes the AI adoption measures. About 61 percent of firm-year observations contain

at least one AI-related paragraph in the 10-K filing, but only about 11 percent report any operational AI use, reflecting the fact that the majority of AI mentions are general or development-oriented rather than describing deployment in internal operations. Conditional on the full sample, the mean number of operational AI paragraphs is 0.244 and the mean total AI paragraph count is 3.15, out of an average 10-K length of approximately 1,271 paragraphs. The operational AI share of total 10-K text is correspondingly small (0.023 percent), indicating that even among firms that discuss AI, operational deployment remains a modest share of disclosure content.

Table 1. Summary Statistics

Variable	N	Mean
Panel A: AI Adoption Measures		
AI Mention (=1 if any AI paragraph)	18,358	0.609
AI Operational Use (=1 if ops>0)	18,358	0.107
AI Operation Paragraphs (ops)	18,358	0.244
Total AI Paragraphs	18,358	3.150
Total 10-K Paragraphs	18,358	1,271.4
AI Ops Share (ops/total paragraphs)	18,358	0.000234
Panel A: Labor Demand (Job Postings)		
Total Job Postings	18,358	1,953.3
High AI Exposure Jobs (Webb)	18,358	925.4
Low AI Exposure Jobs (Webb)	18,358	927.4
Panel C: Firm Characteristics		
Total Assets ($M)	18,352	26,799.7
Employees (thousands)	18,246	21.500
Revenue ($M)	18,351	8,260.0
Capital Expenditures ($M)	16,068	500.4

Panel B reports labor demand as measured by Lightcast job postings. The average firm posts approximately 1,953 vacancies per year, split roughly evenly between high-AI-exposure occupations (925 postings) and low-AI-exposure occupations (927 postings) based on the median Webb score. This near-equal split ensures that neither exposure group dominates the sample mechanically. Panel C summarizes firm characteristics from Compustat. The average firm in the sample has total assets of approximately $26.8 billion, about 21,500 employees, and annual revenue of approximately $8.3 billion, reflecting the large-firm composition of the Compustat–Lightcast linked sample.

In addition, Figure 2 illustrates the time-series evolution of AI-related disclosures in our sample. Panel A plots the share of firms with any AI-related paragraph in their 10-K filings and the share with any operational AI paragraph, which presents the evolution of the extensive margin. Both series rise over time. The share of firms with any AI mention increases from about 50 percent in 2010 to roughly 75 percent in 2022, while the share with operational AI rises from under 10 percent to around 20 percent.

Panel B presents the intensive margin for the subset of firms that ever adopted operational AI, showing the average annual number of total AI paragraphs and operational AI paragraphs per firm. While the average number of total AI paragraphs roughly doubled from about 4 in 2010 to around 9 by 2022, with especially rapid growth after 2017, the average number of operational AI paragraphs remained much lower and stayed below 2 for most of the period.

The large and persistent gaps between the two series in both panels highlight the importance of distinguishing general AI discussion from actual operational deployment. Collectively, these panels indicate that while AI-related disclosures have surged over the last decade, this growth primarily involves general or development-focused discussion rather than

operational implementation. The low and only gradual rise in operational paragraph frequency suggests that despite widespread AI awareness, most firms are still in the preliminary phases of incorporating AI into their internal workflows.

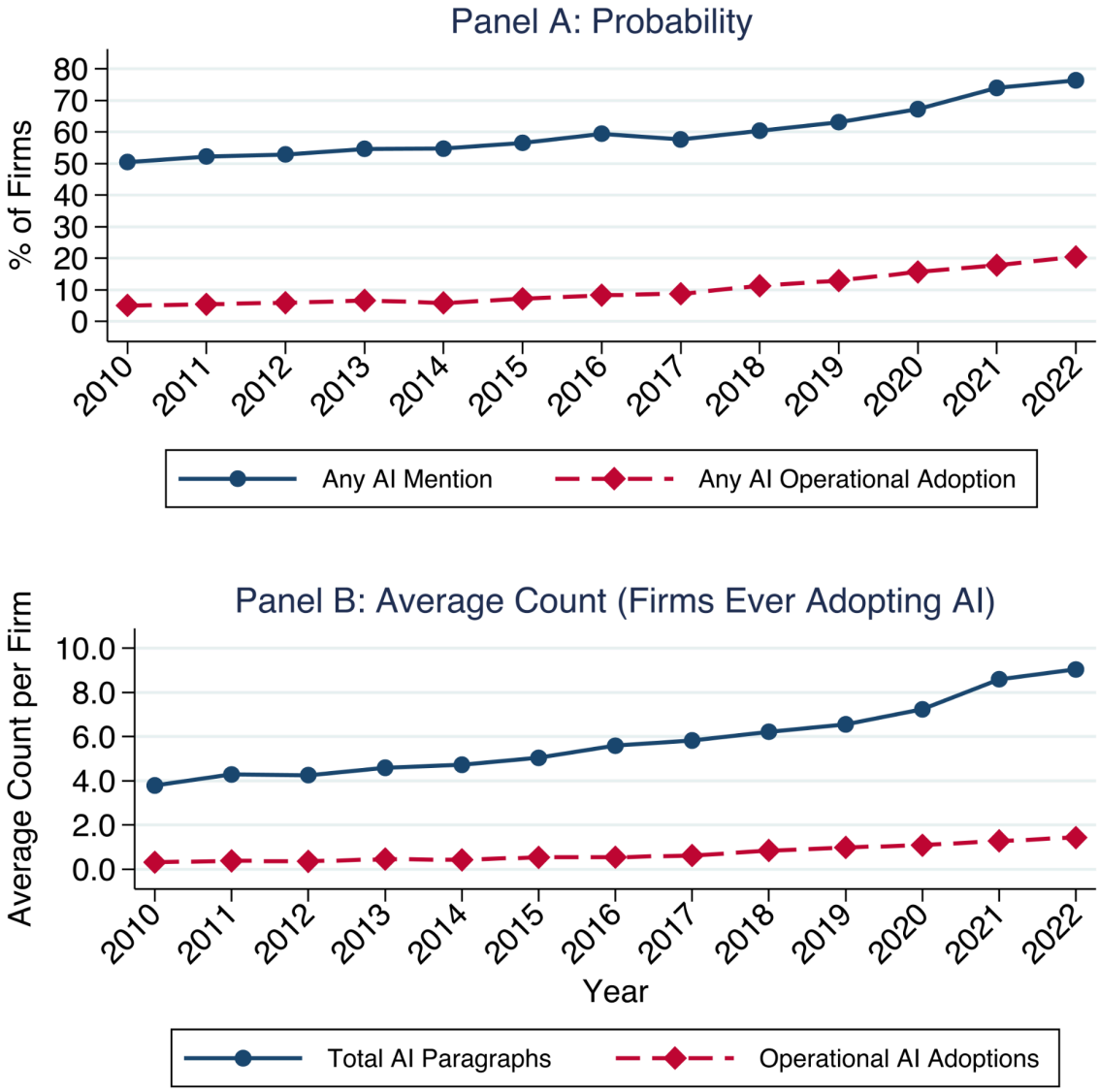


Figure 2. AI Mentions and Adoptions Over Time

4. Empirical Method

To study the relationship between AI adoption and firm-level labor demand, we employ a long difference specification. The long difference estimator is widely used in the labor and technology economics literature to identify medium-run effects of technology on employment (Acemoglu & Restrepo, 2020; Brynjolfsson & Hitt, 2003; Tambe & Hitt, 2012). By comparing changes in outcomes and treatments across multi-year intervals, the approach is less sensitive to short-run measurement noise and better matches the gradual way in which hiring responds to new technology. Specifically, we estimate the following long difference model:

$$\Delta log(Jobs_{i,t}) = \alpha + \beta_1 \Delta log(cumAI_{i,t}) + \gamma' \Delta X_{i,t} + \varepsilon_{i,t}$$

Here, $\Delta log(Jobs_{i,t})$ denotes the k-year long difference in the logarithm of job postings by firm i at the end of interval t, defined as $log(Jobs_{i,t}) - log(Jobs_{i,t-k})$. This study investigates job postings in occupations with high automation scores from Webb as well as those with low automation scores. The main regressor is the corresponding change in $log(cumAI_{i,t})$, where $cumAI_{i,t}$ is the cumulative count of operational AI paragraphs disclosed in the firm's 10-K filings. Measuring AI adoption cumulatively is useful because it captures the stock of disclosed operational deployment rather than one-off mentions. $\Delta X_{i,t}$ is a vector of first-differenced covariates that include changes in the firm's employment and revenue, controlling for time-varying firm characteristics that may affect both AI adoption decisions and labor demand. Accordingly, the coefficient on $\Delta log(cumAI_{i,t})$ summarizes the medium-run association between job postings and disclosed AI adoption, conditional on observed firm characteristics and common time shocks.

5. Empirical Findings

5.1. Baseline Results

We next examine how changes in Cumulative AI adoption are related to changes in job postings across occupations with different exposure to AI automation. Table 2 reports long-difference estimates for the full sample across k-year horizons (k = 1 to 5 years). All columns include firm and year fixed effects together with employment and revenue controls, and columns (2) and (4) add the additional controls reported in the table.

As presented in column (1), for high-impact occupations (those classified as most vulnerable to AI displacement according to Webb's automation scores), we observe a clear negative pattern that strengthens with the horizon length: while the one-year effect is small and insignificant, the effect grows substantially more negative at longer horizons. By two-year difference, the coefficient reaches −0.144 ($p<0.05$), meaning one percent increase in AI operation paragraphs between the two year difference is associated with 0.14 percent decrease in the demand of jobs that are in high-impact occupations. The decline of high-impact jobs continues to grow, and after a five-year difference, it reaches 0.3 percent ($p<0.01$). These estimates imply that increases in disclosed AI adoption are associated with progressively larger reductions in postings for Webb-exposed jobs over the years.

The addition of further controls in column (2) does not materially alter these patterns, though in some cases the magnitudes increase slightly, suggesting that estimated effects are not driven by omitted firm characteristics. Specifically, at a two-year difference, the magnitude of demand decline for high-impact jobs reaches 0.18 percent ($p<0.05$) with one percent increase in

AI operation paragraphs with additional controls, and continues rising to 0.32 percent ($p<0.01$) at the three-year difference and beyond.

The temporal trajectory of these effects is economically important. The absence of significant one-year effects combined with growing negative effects over two to five years suggests that labor market adjustment to AI adoption is gradual. Firms may not immediately reduce hiring in exposed occupations, but the cumulative effect of sustained AI investment manifests increasingly over the medium-to-long run. This pattern is consistent with search frictions, contractual constraints, or organizational adjustment costs that delay the reallocation of labor.

Table 2. Baseline Results with Full Sample

	Highly Impacted Occupations		Low Impacted Occupations	
DV: AI adoption	(1)	(2)	(3)	(4)
1 year difference	-0.0609	-0.0281	0.0360	0.0901
	(0.0526)	(0.0610)	(0.0542)	(0.0628)
2 year difference	-0.144**	-0.179**	-0.0269	-0.0383
	(0.0686)	(0.0790)	(0.0713)	(0.0826)
3 year difference	-0.196**	-0.270***	-0.0779	-0.0975
	(0.0773)	(0.0887)	(0.0812)	(0.0937)
4 year difference	-0.242***	-0.322***	-0.133	-0.137
	(0.0856)	(0.0972)	(0.0909)	(0.104)
5 year difference	-0.296***	-0.326***	-0.162	-0.129
	(0.0938)	(0.105)	(0.101)	(0.114)
Log Employment	Y	Y	Y	Y
Log Revenue	Y	Y	Y	Y
Additional Controls		Y		Y

Standard errors in parentheses. *** $p<0.01$, ** $p<0.05$, * $p<0.1$. All specifications include firm- and year-fixed effects.

In contrast, employment in low-impact occupations shows little systematic response. Across all horizons and control specifications (columns 3 and 4), coefficients remain small and statistically insignificant, hovering around zero or slightly negative. At a five-year difference, the low-impact coefficients are both insignificant, compared to the highly significant −0.296 and −0.326 for high-impact occupations. This occupational asymmetry strongly supports the hypothesis that AI adoption selectively reduces labor demand in occupations most exposed to AI automation, rather than producing broad-based employment contractions.

5.2. Falsification Test with Total AI Mentions

To further evaluate the argument that the observed contraction in labor demand for exposed occupations is driven by operational AI deployment rather than generalized AI engagement, this section performs a falsification exercise. Table 3 presents the results from specifications of replacing the operational AI metric with a count of all AI-related disclosures in a firm's 10-K filings, including development-oriented and general mentions. If our baseline results merely reflected a spurious correlation between any form of AI activity and hiring trajectories, this broader measure should yield comparable negative elasticities across the long-difference horizons.

The results from this alternative specification diverge sharply from the baseline. As reported in Table 3, none of the estimated coefficients for high-impact occupations achieve statistical significance at the five-percent level across any of the k-year horizons or control configurations. The point estimates are economically small and fluctuate around zero. The one year coefficient is marginally positive and statistically significant only at the 10 percent level in the specification without additional controls, while the estimates at longer horizons remain

statistically indistinguishable from zero. A similar null pattern is observed for low-impact occupations, where coefficients remain small and insignificant throughout. This stands in stark contrast to the baseline results presented in Table 2, where operational AI adoption produces displacement effects that intensify and gain significance as the adjustment horizon extends.

Table 3. Falsification Test with Total AI Mentions

	Highly Impacted Occupations		Low Impacted Occupations	
DV: Total AI Mentions	(1)	(2)	(3)	(4)
1 year difference	0.0712*	0.0662	0.0443	0.0533
	(0.0367)	(0.0417)	(0.0355)	(0.0410)
2 year difference	0.0009	-0.0460	-0.0080	-0.0101
	(0.0576)	(0.0624)	(0.0571)	(0.0631)
3 year difference	-0.0191	-0.1050	0.0039	-0.0148
	(0.0744)	(0.0798)	(0.0747)	(0.0812)
4 year difference	-0.0243	-0.0954	-0.0147	-0.0276
	(0.0886)	(0.0950)	(0.0890)	(0.0956)
5 year difference	-0.0491	-0.1030	-0.0193	-0.0552
	(0.0985)	(0.104)	(0.103)	(0.110)
Log Employment	Y	Y	Y	Y
Log Revenue	Y	Y	Y	Y
Additional Controls		Y		Y

Standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1. All specifications include firm- and year-fixed effects.

These null results reinforce the internal validity of our analysis through two distinct falsification channels. First, they demonstrate that general AI-related disclosures, whether pertaining to product innovation, competitive strategy, or industry trend, lack predictive power regarding subsequent labor demand shifts in exposed occupations. The displacement signal is

uniquely isolated to disclosures characterizing the operational integration of AI within the firm. Second, this divergence confirms the value of our Ops/Dev/None classification pipeline; if these categories were interchangeable proxies for a same underlying concept, the total-mention metric would replicate the baseline findings. The observed attenuation confirms that operational AI adoption represents a distinct and economically consequential dimension of firm activity with specific implications for medium-run labor market adjustment.

5.3. Extensive Margin

While the baseline specification characterizes AI adoption on the intensive margin through the logarithmic transformation of cumulative operational disclosures, Table 4 shifts the focus to the extensive margin by substituting cumulative AI operation counts with a binary indicator. This indicator equals one if a firm reports any operational AI integration in its 10-K filings and zero otherwise, testing whether to adopt AI, rather than the depth of disclosed deployment, drives the observed contraction in labor demand.

The extensive margin estimates are consistently larger in absolute value than the baseline intensive margin estimates. For high impact occupations, the coefficient in column (1) is -0.10 at the one year horizon, compared with -0.06 in the baseline, although it is not statistically significant. By the two year horizon, the elasticity reaches -0.23 ($p < 0.01$), exceeding the baseline estimate of -0.14. The negative effect continues to grow over time, reaching -0.35 at four years and -0.39 at five years, both significant at the 1 percent level. Including the additional controls in column (2) further strengthens the results: the five year coefficient is -0.48 ($p < 0.01$), which is much greater than the corresponding baseline estimate of -0.33. In contrast, the

estimates for low impact occupations remain small and statistically indistinguishable from zero across all horizons in both specifications.

Table 4. Extensive Margin Results

	Highly Impacted Occupations		Low Impacted Occupations	
DV: AI Adoption Binary	(1)	(2)	(3)	(4)
1 year difference	-0.1010	-0.0864	0.0236	0.0917
	(0.0631)	(0.0723)	(0.0697)	(0.0810)
2 year difference	-0.2310***	-0.2600**	-0.0977	-0.0356
	(0.0890)	(0.105)	(0.103)	(0.120)
3 year difference	-0.2990***	-0.3520***	-0.1710	-0.1250
	(0.111)	(0.132)	(0.128)	(0.154)
4 year difference	-0.3450***	-0.4050***	-0.2130	-0.2070
	(0.129)	(0.155)	(0.154)	(0.187)
5 year difference	-0.3920***	-0.4770***	-0.2270	-0.2910
	(0.144)	(0.169)	(0.182)	(0.221)
Log Employment	Y	Y	Y	Y
Log Revenue	Y	Y	Y	Y
Additional Controls		Y		Y

Standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1. All specifications include firm- and year-fixed effects.

The larger extensive margin estimates suggest that the discrete shift into operational AI deployment is more predictive of subsequent hiring adjustment than variation in the volume of AI related disclosures. This pattern is consistent with a threshold view of technology adoption: once firms move into operational deployment, labor demand adjusts more substantially, while the marginal effect of additional disclosed AI activity appears comparatively limited. The result also provides a useful robustness check for the baseline specification, since the binary indicator relies

on a simpler and more transparent source of variation than the continuous log(CumAI) measure, reducing concerns that the baseline findings are driven by idiosyncratic variation in disclosure counts.

5.4. Result Heterogeneity

5.4.1. IT versus non-IT sectors

Figure 3 displays dynamic coefficients separately for IT and non-IT sectors. Interestingly, firms in the IT sector exhibit no significant negative employment effect of AI adoption in high-impact occupations across any horizon from one-year to five-year differences. The coefficients remain close to zero throughout, and confidence intervals consistently cross zero. This pattern holds for both high-impact and low-impact occupations in the IT sector, suggesting that AI adoption in technology firms does not translate into measurable job posting reductions within this sample.

Conversely, non-IT firms display the negative displacement pattern identified in the baseline results. For high-impact occupations, the coefficient starts near zero at the one-year difference and declines to approximately 0.4 percent by five-year difference, with confidence intervals excluding zero by three-year difference and thereafter. Low-impact occupations in non-IT sectors show no significant effects, mirroring the baseline pattern.

Taken together, these results indicate that the labor demand effects of AI adoption differ sharply across sectors: in IT firms, adoption appears to be accompanied by offsetting demand for complementary technical labor, whereas in non IT firms it is more closely associated with substitution away from workers in Webb exposed occupations. The sectoral distinction illuminates the heterogeneous labor market consequences of AI. This pattern suggests that AI

adoption is more strongly associated with labor-saving adjustment in non-IT firms than in IT firms.

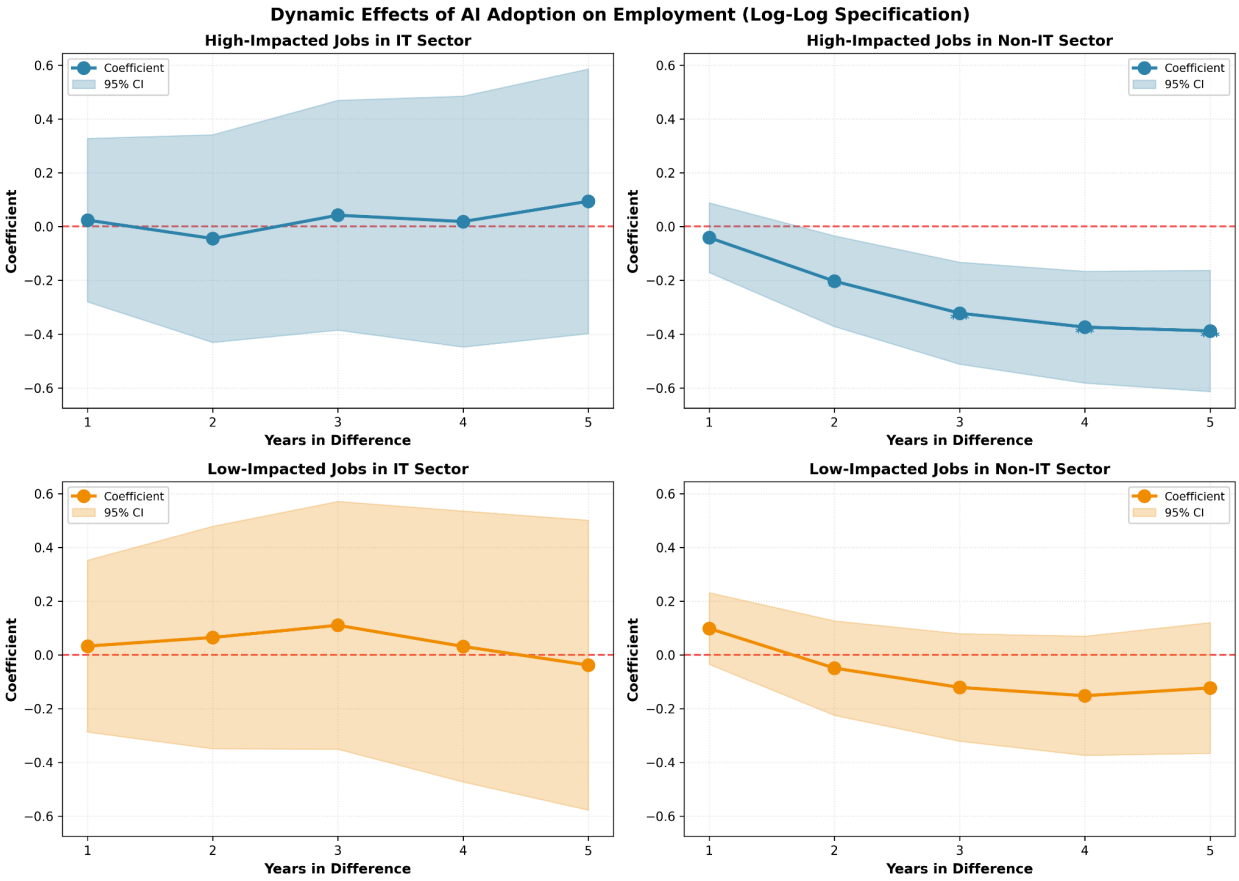


Figure 3. Effect Heterogeneity for IT and Non-IT Sectors

5.4.2. Service versus Manufacturing Sectors

Figure 4 compares dynamic effects across service and manufacturing sectors, revealing substantial heterogeneity. In the service sector, high impact occupations experience modest negative effects that grow over time: the coefficient is close to zero at the one year horizon, declines to about -0.25 by the five year horizon, and becomes statistically significant from the three year horizon onward. Low impact occupations in services show no significant effects, with

coefficients remaining small and confidence intervals consistently including zero. By contrast, the manufacturing sector exhibits much larger and more persistent negative effects. High impact occupations in manufacturing show substantial displacement, with coefficients around -0.20 at the one year horizon and falling to roughly -0.70 to -0.80 by the five year horizon, with confidence intervals excluding zero at all horizons. Low impact occupations in manufacturing also experience significant declines at longer horizons, with coefficients falling from near zero to about -0.60 by five years and becoming statistically significant beginning at the three year horizon.

The magnitude and persistence of the manufacturing results stand in clear contrast to the more muted pattern in services. One interpretation is that manufacturing tasks are more closely aligned with the types of production activities that AI can substitute for directly, generating faster and larger reductions in labor demand. In services, by contrast, AI may more often complement labor or be used in ways that do not immediately reduce hiring, which is consistent with the smaller estimated effects.

More broadly, these results suggest that the labor market consequences of AI adoption depend importantly on the context in which the technology is deployed. Manufacturing's steep decline in job postings reflects the alignment between AI capabilities and the production tasks that dominate that sector. The weaker response in the service sector suggests that technical complementarities with labor remain important in service-intensive industries, or that firms in those sectors are using AI to expand service capacity rather than reduce headcount. The heterogeneity across sectors implies that aggregate labor market estimates that pool manufacturing and services may substantially understate the local displacement effects concentrated in manufacturing.

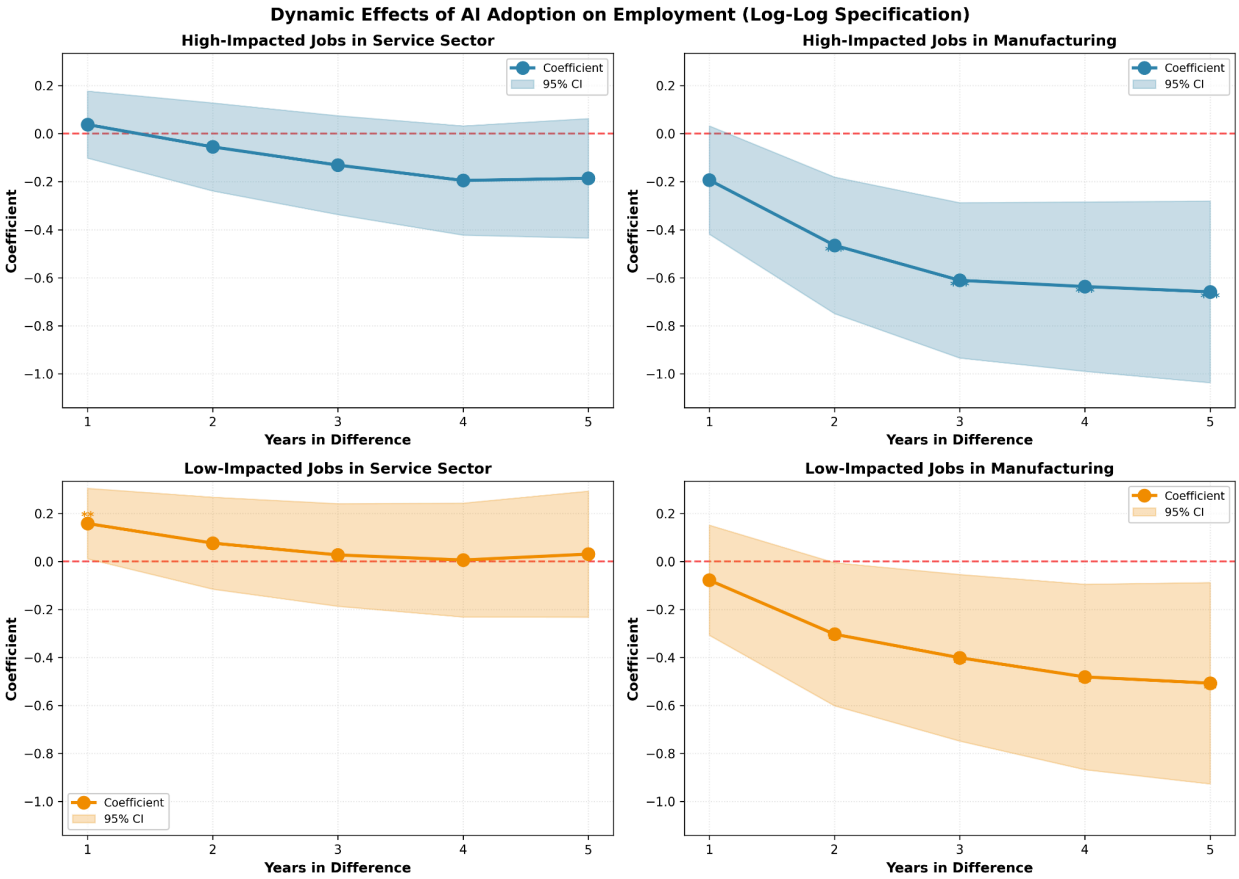


Figure 4. Effect Heterogeneity for Service and Manufacturing Industries

5.4.3. AI Developing Firms vs Non-Developing

Figure 5 distinguishes between firms that develop AI technologies and those that adopt or use AI developed elsewhere. For high impact occupations, AI developing firms exhibit only a modest negative trend: the coefficient is slightly positive at the one year horizon and declines to about -0.30 by the five year horizon, with confidence intervals often including zero and the effect reaching only borderline significance at five years. By contrast, non developing firms show much stronger displacement effects. Their coefficients are close to zero at the one year horizon

but fall to about -0.50 by the five year horizon, with confidence intervals excluding zero from the three year horizon onward, indicating robust effects at longer horizons.

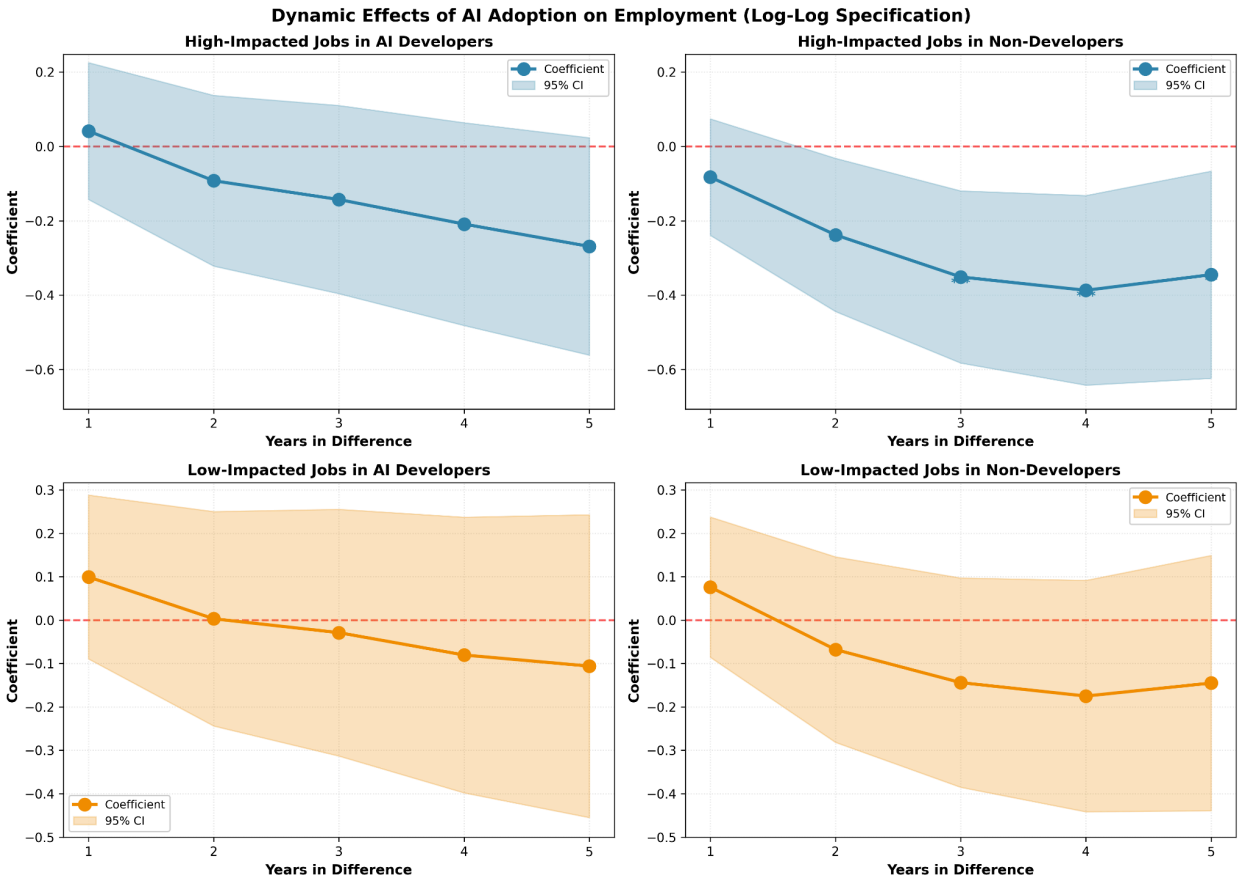


Figure 5. Effect Heterogeneity for AI Developers and Non-Developers

This pattern aligns with economic theory on innovation and labor demand. Firms that develop AI technologies in-house may face stronger incentives to expand employment of skilled engineers and researchers to build and refine those technologies, even as AI itself reduces demand for routine workers. The internal innovation process generates complementary hiring that partially offsets the displacement effect of the technology itself. Non-developing firms, lacking these internal innovation dynamics, are adopting already-developed AI tools designed to automate existing tasks, resulting in net labor reduction with no offsetting research employment

expansion. Low-impact occupations show no significant effects in either group, remaining small and insignificant across all horizons.

This finding implies that the identity and capabilities of the firm matter substantially for labor market outcomes. The AI job shakeup is sharper outside the technology development sector, where adoption represents a straightforward substitution of capital for labor. Inside AI-developing firms, innovation-driven hiring potentially mitigates the displacement effect to some extent. These dynamics suggest that the economy-wide labor market impact of AI will depend on the relative growth of AI development and AI adoption activities, and on whether the skilled workforce expansion in AI development can accommodate workers displaced from routine tasks in non-developing firms.

5.4.4. Large vs Small Firms

Figure 6 presents dynamic effects stratified by firm size, using a $10 billion revenue threshold to distinguish mega firms and relatively smaller firms. Both groups show significant negative employment effects for high impact occupations at medium to longer horizons. Among large firms, the coefficient is slightly positive at the one year horizon and remains close to zero through the two year horizon, before becoming more negative over time and reaching about -0.40 by the five year horizon, with confidence intervals excluding zero from the four year horizon onward. Small firms, by contrast, display a steadier negative pattern beginning at the one year horizon, with coefficients gradually declining to about -0.50 by five years and becoming statistically significant by the three year and four year horizons. For low impact occupations, the estimates remain small and mostly statistically insignificant in both size categories.

The difference in timing across firm size is economically informative. Large firms appear to adjust labor demand more gradually following AI adoption, with the effects concentrated at

medium to longer horizons. One possible explanation is that larger firms face greater adjustment frictions, including more complex organizational structures, established internal labor markets, or other costs of workforce reallocation. Smaller firms, by contrast, may be able to adjust more quickly, consistent with greater organizational flexibility and lower costs of changing hiring plans.

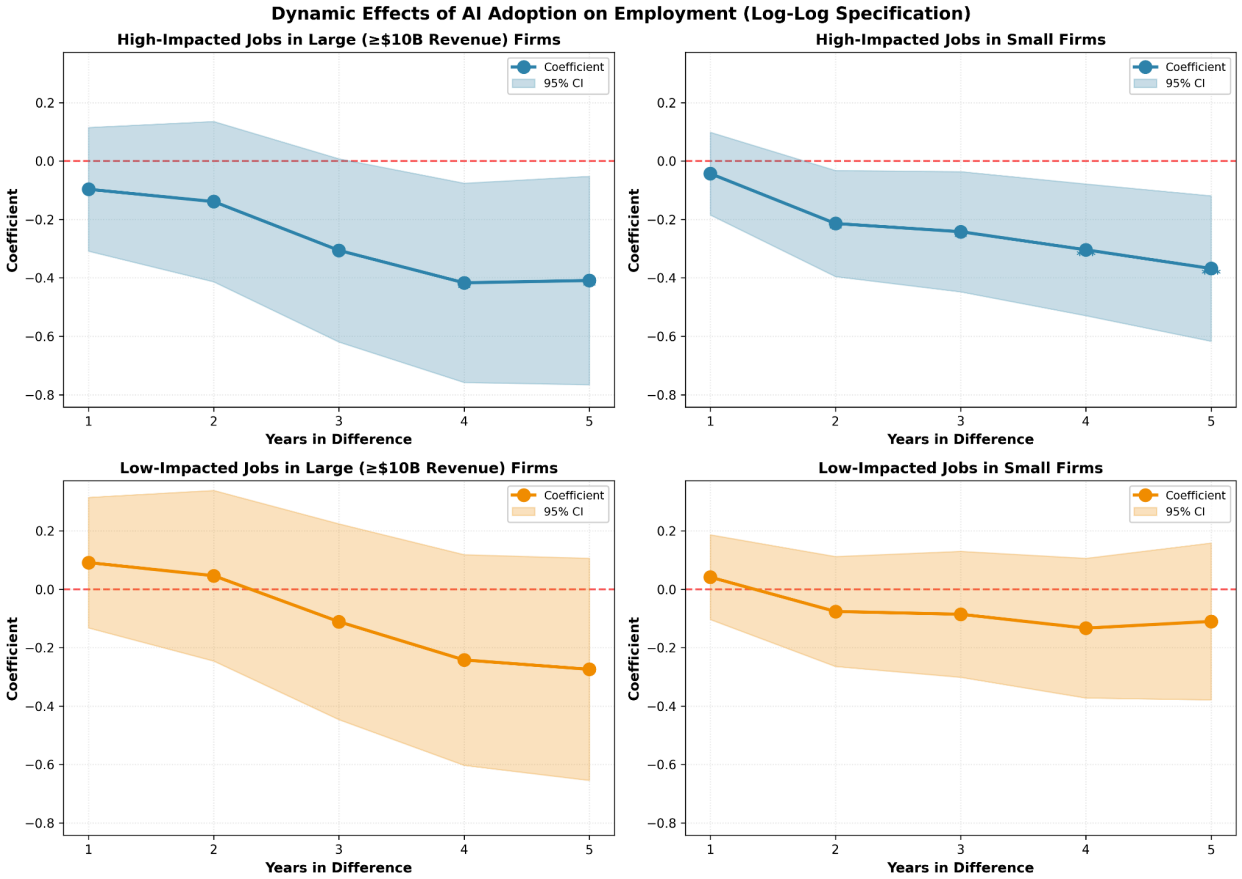


Figure 6. Effect Heterogeneity Based on Firm Size

Both firm size groups ultimately converge to substantial medium to long run displacement in high impact occupations, with coefficients reaching about -0.40 by the five year horizon. The main difference lies in the timing rather than the eventual magnitude of the effect. These results suggest that firm size shapes the pace, but not necessarily the ultimate extent, of

labor demand adjustment following AI adoption. Large firms may absorb more of the short run impact through existing employment relationships and internal reallocation, but over time sustained AI adoption is associated with displacement of a magnitude similar to that observed among smaller firms.

5.5. AI and Labor Productivity

We then examine the relationship between AI adoption and labor productivity, measured as revenue per employee, in the same log log long difference framework used for the hiring analysis. The task based framework implies that a labor saving technology may reduce labor demand in automatable tasks while increasing output per worker through productivity gains. The productivity analysis therefore complements the displacement results above. Absent such productivity gains, a decline in hiring following AI adoption would be more difficult to distinguish from firm contraction rather than efficiency enhancing technological change.

Figure 7 presents that the estimated elasticity of labor productivity with respect to CumAI is positive at every horizon. The coefficient rises from about 0.05 at the one year horizon to 0.08 at the two year horizon, where it first becomes statistically significant at the 1 percent level, and continues to increase to about 0.13 at the four year horizon before easing slightly to 0.12 at five years. The 95 percent confidence interval excludes zero from the two year horizon onward. These estimates imply that a 1 percent increase in cumulative operational AI paragraphs is associated with a 0.08 to 0.13 percent increase in revenue per employee over horizons of two to five years.

Read alongside the baseline hiring results in Table 2, these estimates indicate that firms with greater increases in operational AI adoption tend to experience rising labor productivity at

the same time that their postings decline in the occupations most exposed to automation. This combination is consistent with AI operating as a labor-saving technology: it reduces the labor input required in certain task domains while raising output per remaining worker.

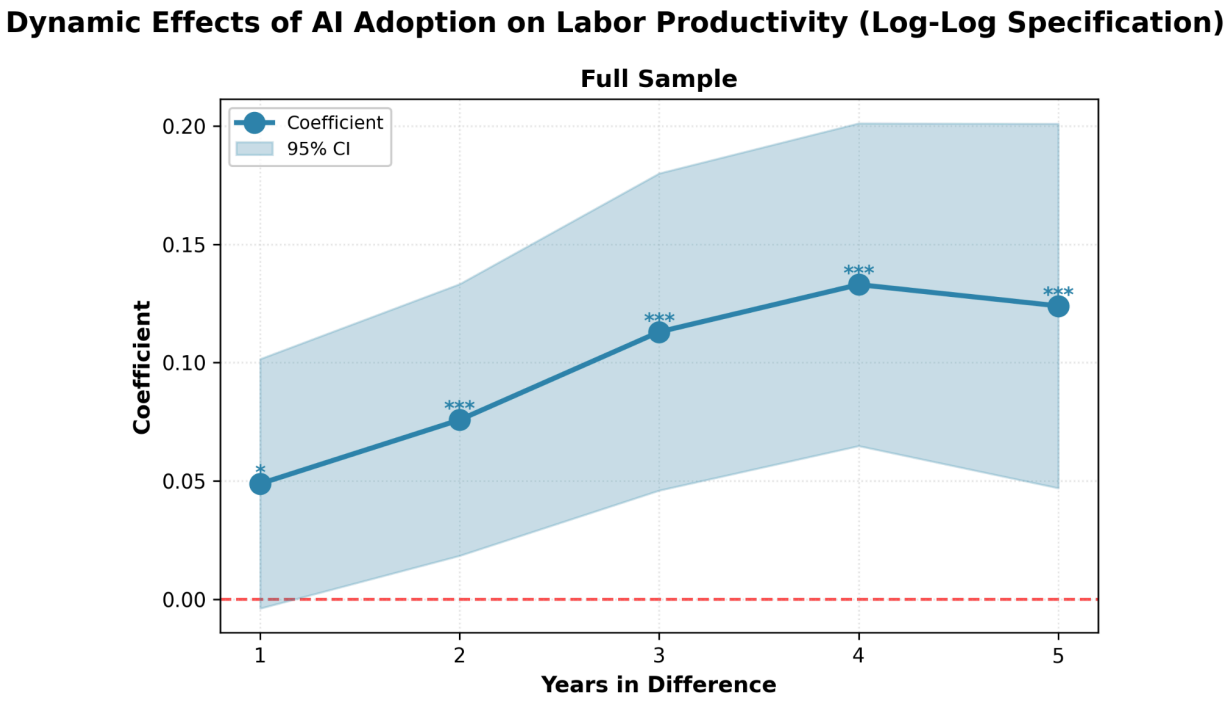


Figure 7. AI adoption and Labor Productivity

5.5.1. Heterogeneity by Firm Size

Figure 8 decomposes the productivity analysis by firm size using a $10 billion revenue threshold. The positive association between AI adoption and labor productivity appears in both groups, and the overall patterns are broadly similar. Among large firms, the coefficient rises from about 0.03 at the one year horizon to roughly 0.20 at the five year horizon. The estimates become

statistically distinguishable from zero by the three year horizon, and the upward pattern does not level off within the five year window.

Small firms display a similar qualitative pattern. The coefficient increases from about 0.05 at the one year horizon to roughly 0.13 at the three year and four year horizons, before easing slightly to about 0.11 at five years. The confidence intervals are somewhat wider than those for large firms, indicating greater heterogeneity, but the substantial overlap between the two sets of estimates suggests that the data do not support a sharp size based difference in the magnitude of the productivity effects.

Such similarity across firm size groups is perhaps not surprising, given that the Compustat sample is tilted toward relatively large publicly traded firms. As a result, even the smaller firm group is not representative of small firms in the broader economy.

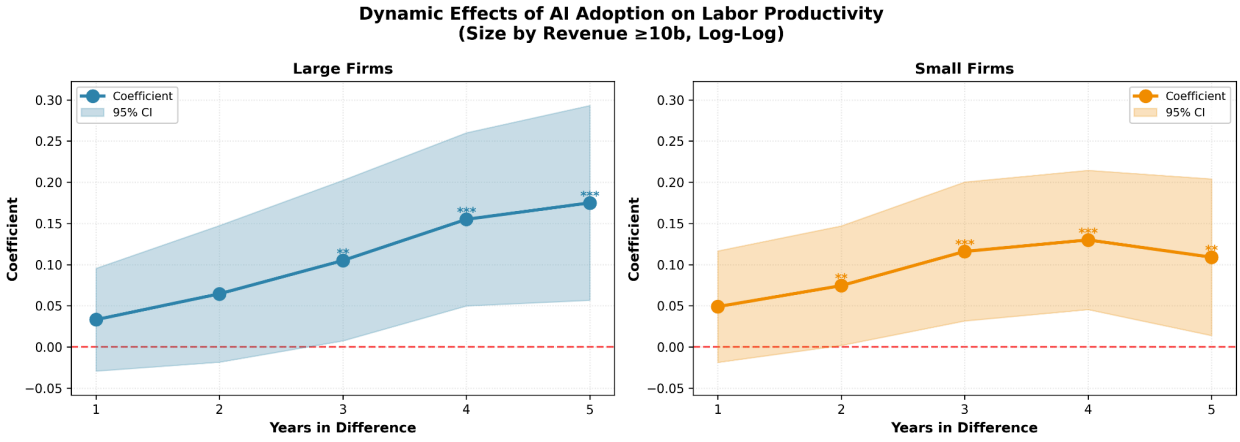


Figure 8. AI adoption and Labor Productivity by Firm Size

Read together with the hiring results in Section 5.4.4, these estimates suggest a broadly similar pattern across both size groups. The timing of adjustment differs modestly by firm size: large firms appear somewhat slower to reduce hiring in exposed occupations and somewhat slower to realize measurable productivity gains. One possible interpretation is that organizational

frictions in larger firms, such as more complex implementation processes, longer planning cycles, or internal labor market considerations, may slow both workforce adjustment and the realization of productivity benefits. Overall, however, any differences across firm size seem to relate more to timing than to the direction of the effect.

6. Summary and Discussion

This paper examines how firm-level AI adoption reshapes labor demand across occupations with different exposure to AI. Using a new disclosure-based measure of operational AI adoption constructed from firms' 10-K filings, we link cumulative operational AI adoption to occupation-level exposure scores from Webb (2019) and firm-level vacancy data from Lightcast. The empirical design uses long-difference specifications over one- to five-year horizons to study medium-run adjustment in hiring. Across specifications, the main pattern is consistent: AI adoption is associated with progressively larger declines in job postings for occupations with high AI exposure, while postings in less-exposed occupations remain broadly unchanged. The estimates are small and statistically insignificant at the one-year horizon, but they become increasingly negative over longer horizons, indicating that the labor-demand effects of AI emerge gradually rather than immediately. Falsification tests suggest that the decline in labor demand for exposed occupations is not driven by broader AI engagements.

The heterogeneity results further sharpen this conclusion. The negative hiring response is concentrated in non-IT firms, in manufacturing, and among firms whose disclosures are more consistent with operational adoption than with AI development. By contrast, the relationship is weak or absent in IT firms and more muted in service industries. This pattern suggests that AI does not affect all firms through a common channel. In firms that develop AI, continued demand

for engineers, technical specialists, and implementation-related capabilities may offset part of the labor-saving effect. In non-developing firms, however, AI appears more likely to be deployed as an input into existing workflows, generating a clearer substitution effect in occupations whose task content is more exposed to AI. The distinction between building AI and deploying AI therefore matters not only for measurement, but also for understanding the direction and magnitude of labor-demand adjustment.

The results also suggest that the employment consequences of AI should be interpreted together with productivity. The evidence indicates that AI adoption is positively associated with labor productivity even as hiring declines in more exposed occupations. Taken together, these findings are consistent with AI functioning as a labor-saving technology in selected task domains while improving firm performance. This does not imply a uniform reduction in labor demand across the firm. Rather, the results point to selective adjustment: firms appear to reduce hiring where AI can substitute for a meaningful share of task content, while demand in less-exposed occupations is comparatively stable. In this sense, the findings align with the broader task-based view of technological change, in which the effect of a new technology depends on the composition of work it can replace, complement, or leave untouched.

Methodologically, the paper develops a scalable approach to measuring firm level technology adoption from regulatory filings. By combining neural network classification of sentence transformer embeddings with large language model reclassification, the pipeline processes more than 30 million 10 K paragraphs and isolates the small subset describing operational AI deployment, a signal that conventional keyword searches or hiring based proxies would likely miss. More broadly, this text mining framework is not specific to AI and could be adapted to measure other forms of technology adoption disclosed in corporate filings.

More broadly, the paper contributes to the emerging literature by showing that how researchers measure AI adoption materially affects the conclusions they draw. Much of the existing firm-level literature relies on AI-related hiring as an indicator of AI activity. That approach is informative for identifying firms that build internal AI capability, but it is less well suited to capturing firms that adopt externally developed tools without substantial AI hiring. Our disclosure-based measure complements that literature by identifying operational AI deployment directly from regulatory filings and by distinguishing operational adoption from AI development. This distinction is especially important in a period when AI technologies are increasingly available through external vendors, cloud platforms, and packaged enterprise applications. The results suggest that the labor-market consequences of AI are likely to be understated or mischaracterized when build and buy channels are not separated.

At the same time, the study has several limitations. First, the estimates should be interpreted as conditional associations rather than definitive causal effects. Although the long-difference design, firm fixed effects, year fixed effects, and time-varying controls help reduce several sources of bias, omitted shocks correlated with both AI adoption and hiring may still affect the estimates. Second, the disclosure-based measure necessarily captures reported AI use rather than the full underlying intensity of adoption, and disclosure practices may vary across firms for reasons unrelated to actual implementation. Third, the analysis is limited to publicly traded firms that can be linked across Compustat, Lightcast, and 10-K data, so the findings may not generalize to smaller private firms or to sectors where online vacancies are less informative about labor demand. Finally, the analysis focuses on vacancies rather than realized employment, which makes it most informative about hiring adjustment rather than total workforce change. These limitations point to several directions for future work, including stronger identification

strategies, richer measures of within-firm occupational change, and broader evidence on how AI adoption affects wages, job design, and worker reallocation over time.

References

Acemoglu, D., Autor, D., Hazell, J., & Restrepo, P. (2022). Artificial Intelligence and Jobs: Evidence from Online Vacancies. *Journal of Labor Economics*, *40*(April 2022).

Acemoglu, D., & Restrepo, P. (2020). Robots and Jobs: Evidence from US Labor Markets. *Journal of Political Economy*, *128*(6). https://doi.org/10.3386/w23285

Alekseeva, L., Azar, J., Gine, M., & Samila, S. (2024). *AI Adoption and the Demand for Managerial Expertise* (SSRN Scholarly Paper No. 4986379). Social Science Research Network. https://doi.org/10.2139/ssrn.4986379

Alekseeva, L., Azar, J., Giné, M., Samila, S., & Taska, B. (2021). The demand for AI skills in the labor market. *Labour Economics*, *71*, 102002. https://doi.org/10.1016/j.labeco.2021.102002

Autor, D. H., Levy, F., & Murnane, R. J. (2003). The Skill Content of Recent Technological Change: An Empirical Exploration. *The Quarterly Journal of Economics*, *118*(4), 1279–1333. https://doi.org/10.1162/003355303322552801

Babina, T., Fedyk, A., He, A., & Hodson, J. (2024). Artificial intelligence, firm growth, and product innovation. *Journal of Financial Economics*, *151*, 103745. https://doi.org/10.1016/j.jfineco.2023.103745

Babina, T., Fedyk, A., He, A. X., & Hodson, J. (2022). *Firm Investments in Artificial Intelligence Technologies and Changes in Workforce Composition* (SSRN Scholarly Paper No. 4060233). https://doi.org/10.2139/ssrn.4060233

Bai, J. (Jianqiu), Brynjolfsson, E., Jin, W., Steffen, S., & Wan, C. (2021). *Digital Resilience: How Work-From-Home Feasibility Affects Firm Performance* (Working Paper No. 28588). National Bureau of Economic Research. https://doi.org/10.3386/w28588

Bessen, J. (2018). Artificial Intelligence and Jobs: The Role of Demand. In *The Economics of Artificial Intelligence: An Agenda* (pp. 291–307). University of Chicago Press. https://www.nber.org/books-and-chapters/economics-artificial-intelligence-agenda/artificial-intelligence-and-jobs-role-demand

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

Brynjolfsson, E., Jin, W., & Wang, X. (2023). *Information Technology, Firm Size, and Industrial Concentration* (Working Paper No. 31065). National Bureau of Economic Research. https://doi.org/10.3386/w31065

Brynjolfsson, E., Li, D., & Raymond, L. (2025). Generative AI at Work. *The Quarterly Journal of Economics*, *140*(2), 889–942. https://doi.org/10.1093/qje/qjae044

Brynjolfsson, E., Mitchell, T., & Rock, D. (2018). What Can Machines Learn and What Does It Mean for Occupations and the Economy? *AEA Papers and Proceedings*, *108*(May), 43–47.

Brynjolfsson, E., Rock, D., & Syverson, C. (2021). The Productivity J-Curve: How Intangibles Complement General Purpose Technologies. *American Economic Journal: Macroeconomics*, *13*(1), 333–372. https://doi.org/10.2139/ssrn.3346739

Campello, M., Gao, J., & Xu, Q. (2023). Personal Taxes and Firm Skill Hiring: Evidence from 27 Million Job Postings. *Management Science*. https://doi.org/10.1287/mnsc.2022.01211

Dell'Acqua, F., McFowland III, E., Mollick, E. R., Lifshitz-Assaf, H., Kellogg, K., Rajendran, S., Krayer, L., Candelon, F., & Lakhani, K. R. (2023). *Navigating the Jagged Technological Frontier: Field Experimental Evidence of the Effects of AI on Knowledge Worker Productivity and Quality* (SSRN Scholarly Paper No. 4573321). Social Science Research Network. https://doi.org/10.2139/ssrn.4573321

Eloundou, T., Manning, S., Mishkin, P., & Rock, D. (2023). *GPTs are GPTs: An Early Look at the Labor Market Impact Potential of Large Language Models* (arXiv:2303.10130). arXiv. http://arxiv.org/abs/2303.10130

Felten, E., Raj, M., & Seamans, R. (2021). Occupational, industry, and geographic exposure to artificial intelligence: A novel dataset and its potential uses. *Strategic Management Journal*, *42*(12), 2195–2217. https://doi.org/10.1002/smj.3286

Felten, E. W., Raj, M., & Seamans, R. (2018). A Method to Link Advances in Artificial Intelligence to Occupational Abilities. *AEA Papers and Proceedings*, *108*, 54–57. https://doi.org/10.1257/pandp.20181021

Noy, S., & Zhang, W. (2023). Experimental evidence on the productivity effects of generative artificial intelligence. *Science*, *381*(6654), 187–192. https://doi.org/10.1126/science.adh2586

Peng, S., Kalliamvakou, E., Cihon, P., & Demirer, M. (2023, February 13). *The Impact of AI on Developer Productivity: Evidence from GitHub Copilot*. arXiv.Org. https://arxiv.org/abs/2302.06590v1

Sarah H. Bana, Erik Brynjolfsson, Wang Jin, Sebastian Steffen, & Xiupeng Wang. (2024). Human Capital Acquisition in Response to Data Breaches. *MIS Quarterly*, (Forthcoming). https://misq.umn.edu/human-capital-acquisition-in-response-to-data-breaches.html

Tambe, P. (2014). Big Data Investment, Skills, and Firm Value. *Management Science*, *60*(6), 1452–1469. https://doi.org/10.1287/mnsc.2014.1899

Tambe, P., & Hitt, L. M. (2012). The Productivity of Information Technology Investments: New Evidence from IT Labor Data. *Information Systems Research*, *23*(July 2015), 599–617. https://doi.org/10.1287/isre.1110.0398

Webb, M. (2019). *The Impact of Artificial Intelligence on the Labor Market* (Issue 5388).