The Labor Impact of Generative AI on Firm Values*

Andrea L. Eisfeldt† Gregor Schubert‡ Bledi Taska§ Miao Ben Zhang¶

March 2024

Abstract

How do recent advances in Generative AI affect firm value? We construct the first measure of *firms' workforce exposures* to Generative AI and show that an "Artificial-Minus-Human" (AMH) portfolio that is long high-exposure firms and short low-exposure firms earned daily returns of 0.44% in the two weeks following the release of ChatGPT. The labor-exposure effect is more pronounced for firms with greater data assets and is distinct from the effect of firms' product exposures to Generative AI. Highly-exposed workforces can be either substituted for or complemented by Generative AI technologies. We measure whether the exposed tasks are core or supplemental to assess relative substitutability. Examining firms' labor demand and profitability following the release of ChatGPT supports a labor-technology substitution channel for the increases in exposed firms' values that we document.

Keywords: Artificial Intelligence, Generative AI, Chat GPT, Corporate Valuations, Equity Returns, Technological Change, Labor Market Effects of Technology

*We thank Tania Babina, Thomas Chaney, Anastassia Fedyk, Chad Jones, Selale Tuzel, Junbo Wang, Paolo Zaffaroni, and participants in the UCLA Anderson Finance Brown Bag, the UCSD Finance Seminar, NYU Stern Finance Seminar, KDD Finance Day, Q Group 2023 Meeting, the 2023 "Artificial Intelligence and the Economy" conference, Carey Finance Conference, TCU Finance Conference, the Chicago Booth Empirical Finance Conference, the OECD, the ITAM Finance Workshop, and the NBER Big Data and Securities Markets 2023 Meeting for helpful feedback. All remaining errors are our own.

†UCLA Anderson School of Management and NBER, Email: andrea.eisfeldt@anderson.ucla.edu

‡UCLA Anderson School of Management, Email: gregor.schubert@anderson.ucla.edu.

§SkyHive, Email: bt540@nyu.edu.

¶USC Marshall School of Business, Email: Miao.Zhang@marshall.usc.edu

Recent advances in Generative Artificial Intelligence are widely seen as a major technological breakthrough. We construct the first *firm-level* measure of the exposure of U.S. publicly traded firms' workforces to being disrupted by Generative AI. Using this measure, we show that the release of ChatGPT resulted in a significant and substantial change in the relative valuation of US firms. Firms whose workforces are highly exposed to Generative AI increased in market value by almost 5% relative to firms with a low exposure in the two weeks following the release of ChatGPT on November 30th, 2022. The divergence in value reflects the heterogeneous potential for Generative AI to execute the tasks currently accomplished by firms' workforces, and the resulting effects on firms' expected future free cash flows. It is not driven by firms' product-market exposure to Generative AI.

The release of ChatGPT can improve the expected future cash flows of firms with a high share of exposed workforces due to productivity improvements either from substituting labor with technology to save costs, or from complementing labor with Generative AI tools. We show that the labor substitution channel drives our results, and develop a new method for assessing an occupation's potential for substitution vs. complementarity. In particular, examining job postings and wages before and after the ChatGPT release, we show that the substitution effects are concentrated in occupations that are most exposed to Generative AI through their *core* tasks such as proofreaders and web developers, while occupations most exposed through *supplemental* tasks such as financial managers and pharmacy aids experience much smaller substitution effects. Our use of core vs. supplemental task exposure to study substitution vs. complementarity effects of technology on occupations can help to understand which occupations may ultimately benefit from Generative AI enhancements. However, our results for Generative AI's impact on firm value and future profitability illustrate that the main mechanism is the substitution channel—the observed increase in firm value for firms whose workforces are more exposed to Generative AI is driven by occupations whose core tasks are more easily executed by Generative AI technologies.

Generative AI is a general-purpose technology and is changing (and expected to change) the way work is conducted across a broad array of products and industries. Relative to earlier artificial intelligence models, Generative AI models can digest more complex inputs, and can produce human-like output, making Generative AI models (of which ChatGPT is an important example) more versatile and scalable than prior innovations in AI and machine learning. In the past, technology shocks diffused over long periods of time, but the large change in the accessibility of AI tools and the massive amount of analysis and attention generated by the release of ChatGPT allows us to study its impact almost in real-time.[1]

[1]A large body of literature examined the diffusion of technologies in the history and showed that the diffusion process is remarkably slow, particularly in the initial periods after the technologies are available,

Our study, which includes effects on the full cross-section of firm value, wages, job postings, and firm profits, thus sheds light on the quantitative effects of an important technology shock on key corporate and economic outcomes.

While recent work has documented variation in occupational exposure to Generative AI (see Eloundou et al. (2023)), without a connection to firm value one cannot say whether higher exposure is value increasing or decreasing for incumbent firms. If Generative AI does increase efficiency for existing firms, how much value do firms vs. workers accrue? By documenting the increase in the value of firms with higher workforce exposure to Generative AI, while wages and job postings decline for exposed workers, we show that firm owners benefitted relatively more than labor from the technology shock. This is an important finding regarding the economic impact especially in the context of the recent literature on the decline in the share of value added accruing to labor. Our study is the first to link occupational exposure to Generative AI, and firms' workforce exposures, to labor market outcomes including wages and job postings. Moreover, we link these labor outcomes to firm profits, providing additional support for a labor-substitution effect driving changes in firm value.

We build our measure of firms' exposure to Generative AI starting at the level of the tasks that workers in U.S. firms perform. From the O*NET database, we obtain information on the 19,265 tasks that constitute the activities performed in 923 occupations in the U.S. We use a large language model to classify each task's exposure to being done more effectively using ChatGPT based on descriptions of the tasks. We aggregate this task-level exposure to average exposures of occupations and map occupations to publicly traded firms using data from Revelio Labs, which provides firms' occupational shares based on millions of individual public profiles. Our firm-level measure thus captures the potential for the tasks currently performed by labor at those firms to be done more efficiently by using Generative AI.

Our approach yields intuitive exposures that reflect the distinctive features of Generative AI compared to prior technologies. Unlike computerization that mainly disrupted routine-task jobs (e.g., Autor et al. (2003) and Acemoglu and Autor (2011)) and robots that mainly disrupted manual-task jobs (Acemoglu and Restrepo (2020)), occupations exposed to Generative AI are cognitive-task jobs, such as legal, financial, mathematical and administrative professions.[2] Consistent with Generative AI being able to assist with high-level cognitive

constituting an S-curve for technology diffusion (see many examples in Greenwood (1999) and Manuelli and Seshadri (2014)). ChatGPT reached 1 million users in only 5 days after its release, according to OpenAI's CEO Sam Altman's Tweet on December 4, 2022. It reached 100 million users in only two months after its release, a milestone that took the World Wide Web 7 years, WhatsApp 3.5 years, Instagram 2.5 years, and Twitter 5 years to reach (see this link).

[2]Our finding is consistent with Webb (2019) who studies broadly-defined AI on the labor market using patent texts and shows that AI tends to disrupt non-routine cognitive-analytical jobs.

tasks, we show that occupations with higher wages also have higher exposure to Generative AI.[3] Overall, the Generative AI exposure of the tasks in all white-collar occupations is 40% vs. 9% for blue-collar and service occupations. Our firm-level measure also shows that Generative AI is indeed a general-purpose technology, which varies both across and within a broad array of industries, with 3-digit NAICS industry effects explaining only 31% of the variation in the firm-level exposure.

We derive our first main results from an event study that documents variation in the returns to firms with different labor-force exposures to Generative AI following the release of ChatGPT on November 30th, 2022. Since we study firms' relative returns following a major event, our study can be thought of as looking at the effect of the release of ChatGPT on differences in differences in the levels of firms' valuations. Sorting firms into five value-weighted portfolios based on their Generative AI exposures, we show that firms in the highest-exposure quintile, labeled the "Artificial" portfolio, earned 44 basis points higher daily returns than firms in the lowest-exposure quintile, labeled the "Human" portfolio, during the two weeks following the release of ChatGPT.[4] This finding is robust to controlling for firms' exposure to the market factor and the Fama-French 5 factors, and also to controlling for firm characteristics that have been shown to predict returns. Figure 1 plots the cumulative abnormal returns during the time before and after the event for the zero-investment portfolio that goes long the artificial stocks and short the human stocks, which we denote as the "Artificial Minus Human" portfolio (AMH). The AMH portfolio returns do not show an obvious pre-trend before the event window and also no reversal after the event window, supporting the validity of our treatment effect estimation.[5]

While the release of ChatGPT is also expected to affect firm values through effects on products and services,[6] we conduct three separate tests to show that our findings based on firms' labor exposure to Generative AI are distinct from any product-exposure channel. First, we show that the AMH returns during the event window hold when sorting portfolios within the industry. We show this using both the NAICS 3-digit industry classification and the 10K text-based FIC 50 industry classification by Hoberg and Phillips (2016). Noting that only 31% of the firm-level variation in labor exposure to Generative AI is explained by 3-

[3] Our result is consistent with recent findings by Kogan et al. (2019), who find that technological advances impact workers at the higher end of the wage distribution. On the other hand, other studies (Krusell et al. (2000) and Eisfeldt et al. (2023)) document substitutability between low-skilled labor and capital but complementarity between high-skilled labor and capital.

[4] Our choice of the two-week event window is guided by the intensity of public attention to the ChatGPT release event on Twitter.

[5] See Appendix Figure IA.1 for a version of the figure with a longer pre-period.

[6] For example, the stock value of chip maker NVIDIA more than doubled in early 2023 as the firm is a key chip supplier for training Generative AI models.

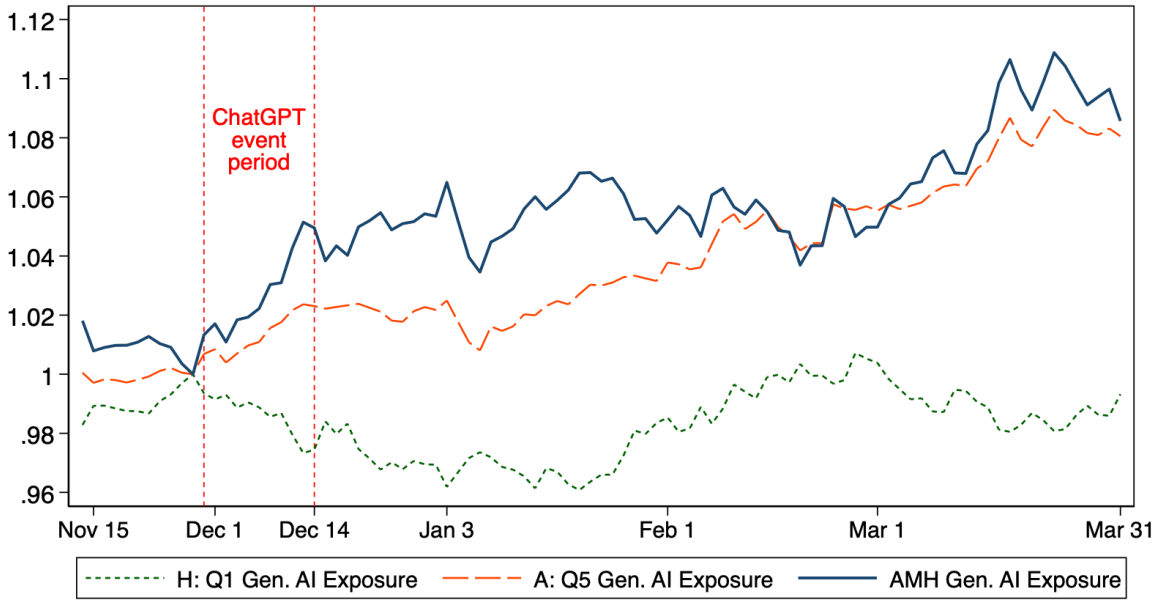


Figure 1. Cumulative Abnormal Returns by Generative AI Exposure. The figure plots the cumulative abnormal returns (CARs) of value-weighted quintile portfolios sorted by firms' labor-based Generative AI exposure. The graph shows the CARs of the lowest-exposure quintile portfolio, "Human" (H), the highest-exposure quintile portfolio, "Artificial" (A), and the zero investment portfolio that longs A and shorts H, "Artificial-minus-Human" (AMH). Market-adjusted daily abnormal returns are cumulated from November 29, 2022, the day before the release of ChatGPT, and are based on factor exposures computed over the 6-month period preceding the period shown in the graph. Daily stock returns are from Yahoo Finance. The dashed vertical lines indicate the "ChatGPT event period" from November 30, 2022, to December 14, 2022. See details of the definition of firms' Generative AI exposure in Section I and the construction of the portfolios, the calculation of portfolios' CARs, and the determination of the ChatGPT event period in Section II. GPT-4 was released on March 14, 2023.

digit industry effects, our results are, not surprisingly, robust to these within-industry sorts. Second, the AMH returns remain the same if we exclude firms from the tech sector whose products are most likely to be directly related to the Generative AI technology.[7] Third, we use four proxies for firms' product exposure to Generative AI: a classification of AI-related business models based on firm annual reports; the count of AI-related keywords based on texts of firms' annual reports; a Goldman Sachs classification of AI beneficiaries; and a measure based on firms' share of AI-skilled workers adopted from Babina et al. (2024). We then run a standard event study test by regressing stock-level cumulative abnormal returns (CARs) on our labor-based Generative AI exposure measure while controlling for the proxies for product exposures. While firms' product exposures separately predict CARs, the effect

[7]Following Acemoglu et al. (2022), we identify the tech sector as the NAICS 51 "Information" and NAICS 54 "Professional, Scientific, and Technical Services" and exclude firms from these two sectors.

of our labor-based Generative AI exposure remains significant, confirming the distinction from a product-exposure channel.

We show that data is an important complement to Generative AI. This may help to understand the benefits of AI improvements for incumbent firms with existing data corpuses. Examples highlighting the importance of data for using Generative AI can be seen in applications training customer service chatbots, automating workflows, improving predictions and analytics, and many others (Caserta et al. (2023)). Realizing the value associated with this technology at scale therefore likely requires a baseline level of data management capabilities and access to relevant input data. We thus expect the return effect of Generative AI exposure to be particularly pronounced among firms with readily available data. We find strong confirmation for this proposition. In particular, we construct two measures of firms' data assets following the prior literature (Begenau et al. (2018), Farboodi et al. (2019), Eeckhout and Veldkamp (2022), Farboodi and Veldkamp (2023), and Abis and Veldkamp (2023)), and we show that Generative AI exposure boosts firm value significantly more if the firm has greater data assets. This finding also explains why firms' potential benefits from Generative AI are not expected to be entirely competed away by new entrants, as data assets tend to be highly specific to the incumbent firms.

Our second set of main results documents the mechanism by which the labor channel generates our results for the impact of Generative AI on firm value. Importantly, we ask whether the effect of higher exposure to Generative AI reflects a greater substitution of, or complementarity to, firms' labor inputs. Ex ante, both channels could increase firms' future cash flows and boost their current market value. That is, firms whose labor force can be substituted for with cheaper Generative AI-based capital could save costs and generate higher future cash flows. On the other hand, if the technology complements the firm's workers and increases their productivity, the firm may also experience an increase in future profitability (Krusell et al. (2000) and Eisfeldt et al. (2022)).

To distinguish between these two channels, we refine our measure of occupational exposure. Specifically, we redefine exposure depending on whether the occupational tasks that are exposed to Generative AI are *core* or *supplemental* according to O*NET. Our key hypothesis is that an occupation is more likely to be substituted by Generative AI if its *core* tasks are more exposed to the productivity improvements enabled by the technology, while if an occupation's *supplemental* tasks are more exposed there is more opportunity for complementarity. Intuitively, core tasks represent the most fundamental duties an employee in that occupation is expected to perform. If Generative AI can more efficiently complete the core tasks at a much lower cost, then it is more likely that the technology can displace the occupation as a whole. In contrast, supplemental tasks are additional tasks or duties asso-

ciated with the core tasks but which are not themselves considered critical or central to the occupation. If Generative AI can help workers more efficiently complete their supplemental tasks, this can can free up time and effort for the worker to focus more on the core tasks, potentially increasing the worker's productivity without making them replaceable. To the best of our knowledge, this is the first study that distinguishes between core and supplemental task exposures within occupations.

We conduct two sets of tests to investigate the mechanism for the labor-channel impact on firm value. Our first set of tests investigates the impact on occupation-level labor demand and wage rates as a result of higher overall Generative AI exposure. Next, we show that, holding total exposure constant, it is the occupations that derive their Generative AI exposure from core tasks that experience a decline in labor demand. We construct numbers of job postings for each occupation-month from January 2019 to August 2023 from granular Lightcast job posting data. We also obtain individual-level hourly wage rates from January 2019 to October 2023 from the Census Current Population Survey data.

Our study contributes three key new findings regarding the labor market outcomes for different occupations before and after the release of ChatGPT. First, occupations with higher Generative AI exposure unconditionally experienced reduced labor demand and a lower relative wage rate after the release of ChatGPT: A one-standard-deviation increase in Generative AI exposure is associated with an 8% decline in job postings and a 0.6% decline in the hourly wage rate. Second, the magnitudes of the effects of Generative AI exposure on job posting and hourly wage rates are *66% and 97% greater*, respectively, if the occupations' Generative AI exposure is derived entirely from their core tasks, i.e., more substitutable according to our approach. Likewise, consistent with greater potential for complementarities when supplemental tasks generate an occupation's Generative AI exposure, the share of an occupation's exposure stemming from supplemental as opposed to core tasks' exposure significantly weakens the above associations. Third, the above findings remain when we examine *within-firm* occupational demand using granular job posting data, reinforcing our interpretation that the occupation-level results are driven by firms reorganizing their operations after the technology shock instead of, for instance, heterogeneous industry dynamics. In summary, these findings show that firms adjust their labor demand in response to the release of ChatGPT in a way that, on average, suggests a substitution effect in the initial months, but with significant heterogeneity depending on the kinds of tasks within an occupation that are affected.

Our second sets of tests examine the technology-labor substitution channel for explaining the cross-section of the labor demand of *firms* with high and low Generative AI exposure. This channel suggests that firms with a more exposed workforce might have a greater capacity to reduce reliance on the exposed occupations and thus decrease costs, improve future

profitability, and boost firm value. To test this channel, we construct an analogous measure of firms' share of Generative AI exposure that derives from its workers' supplemental tasks. We show that firms with higher Generative AI exposure unconditionally reduce job postings more for the occupations that are highly exposed to Generative AI. These firms experience greater increases in profitability in quarterly earnings after the release of ChatGPT. Reinforcing the importance of the role of heterogeneity in the substitution effects, our estimates show that when firms' Generative AI exposure is derived entirely from core tasks of their workforce, the effects of Generative AI exposure on labor demand, and profitability, are more than twice as large as the unconditional average effects. The impact of core-task Generative AI exposure on cumulative abnormal returns is also 75% greater than the impact of overall exposure. Our results for firms' labor demand, profitability, and market value strongly support that a labor-substitution channel is driving Generative AI's impact on firm value.

Our study contributes to the literature on disruptive technologies' impact on firm hiring and firm valuation.[8] Papanikolaou (2011) and Kogan and Papanikolaou (2014) study the effects of investment-specific technological changes on asset prices. Zhang (2019) studies firms' exposure to routine-biased automation. Babina et al. (2024) and Babina et al. (2022) are key early contributions studying the effects of AI on firm growth, compensation, and workforce composition.[9] See also Webb (2019) for the impact of AI on the workforce, and Acemoglu et al. (2022) for evidence of the effects of firm exposure to AI on hiring and skill demand over the 2010-2018 period. Kelly et al. (2021) study firms' exposure to disruptive technological shocks using patent textual data, and Kogan et al. (2019) assess worker displacement from technological change over a very long sample. These two studies offer important insights into investors' and firms' responses to technological shocks using a long panel containing several innovation waves.

Our study focuses on measuring firms' exposure to Generative AI and assessing investors' reaction to the technology shock upon its arrival. We show that the release of ChatGPT in November of 2022 is an observable, large technology shock that created a substantive impact on firm valuation. As the stock market is forward-looking, the information contained in market prices can potentially inform firms and employees about where the technology is likely to be most disruptive. Indeed, right as we released the first draft of this study, IBM, the company ranked #1 in our exposure to Generative AI measure among the largest U.S. firms, announced it would halt hiring of 7,800 jobs that could be replaced by AI.[10]. Timely

[8] See Greenwood et al. (1997) for an early contribution on the long-run impacts of investment-specific technological change.

[9] These studies using job-posting data up to 2018 shed light on the contribution of AI to firm outcomes even before the advances offered by more recent Generative AI models.

[10] We released our first draft to the SSRN on May 9, 2023, while Bloomberg reported the IBM announce-

assessment of the market's expectations of Generative AI's impact on firms can also help policymakers to effectively evaluate regulatory policies in response to the arrival of the new technology.

Our study also contributes to a large body of literature examining technologies' heterogeneous effects on workers with different skills or tasks (see, for examples, Krusell et al. (2000), Autor et al. (2003), Acemoglu and Autor (2011), Acemoglu and Restrepo (2018), Eisfeldt and Papanikolaou (2013), Tuzel and Zhang (2021), Kogan et al. (2023), among others). A common empirical approach measures an occupation's substitutability by a technology based on whether the technology can more efficiently complete its tasks at a lower cost. We adopt this approach in our study but make a novel contribution by distinguishing an "occupation's substitutability" from a "task's substitutability" and a "workforce's substitutability". Importantly, we highlight that an occupation can be substituted or complemented by the new technology depending on whether the new technology can more efficiently complete its core or supplemental tasks. Our empirical results analyzing job postings and wage rates show that recognizing this distinction substantially improves the traditional approach in identifying technology substitution, as our core-task-based Generative AI exposure shows significantly stronger results than the overall task-based Generative AI exposure.

The paper proceeds as follows: Section I describes our data and measure of firms' exposures to Generative AI. Section II presents the impact of the ChatGPT release on the market value of firms with different Generative AI exposures. Section III supports a labor-technology substitution channel by proposing and testing a novel methodology to identify workers' substitutability by Generative AI. Section IV concludes.

I. Data, Measurement, and Stylized Facts

We measure a firm's labor exposure to Generative AI in three steps, starting with task-level exposures, and then aggregating to the occupation- and firm-level exposures, respectively. The Occupational Information Network (O*NET) database describes 19,265 tasks that make up the 923 occupations of U.S.. workers. Each occupation executes a subset of these tasks. The task and occupational exposure measurement follows Eloundou et al. (2023). To aggregate the exposures from the occupation level to the firm level, we use the Revelio Lab database to measure firm-level occupational employment shares. A summary of our three-step procedure, which we detail further below, is:

1. *Task-level exposure:* We use Open AI's GPT 3.5 Turbo model to assign each of the

ment on May 1, 2023. See https://www.bloomberg.com/news/articles/2023-05-01/ibm-to-pause-hiring-for-back-office-jobs-that-ai-could-kill

19,265 tasks in the O*NET an exposure to Generative AI by evaluating whether each task can be more efficiently completed by having access to the capabilities of LLM-based tools like ChatGPT.

2. *Occupation-level exposure:* Aggregate from tasks to occupation exposures by averaging the task-level exposures within each of the 923 occupations.[11]

3. *Firm-level exposure:* Compute the firm-level exposures by weighted averaging the occupation-level exposures using each firm's occupational employment share from the Revelio Lab database as the weight.

A. Measuring task exposure to Generative AI

Occupational task data We consider an occupation to be a portfolio of tasks to be done. From the O*NET V27.2 database, we obtain the task statement for each task provided by practitioners or occupational experts.[12] A task statement is usually one sentence, and an occupation consists of 22 tasks on average. We code each of the 19,265 tasks as being exposed to Generative AI technologies or not using the task statement.

Task scoring We use Open AI's GPT 3.5 Turbo model and the following algorithm to score each task $T's$ exposure X^T to Generative AI following the approach suggested and validated by Eloundou et al. (2023). This approach categorizes each task into one of the following three categories $X^T \in \{1, 0.5, 0\}$ in terms of its Generative AI exposure:

- *Direct Exposure ($X^T = 1$)* if using ChatGPT reduces the time required to complete the task by at least half.

- *Plus-Overlay Exposure ($X^T = 0.5$)* if using ChatGPT would not reduce the time required to complete the task by at least half, but additional software could be developed on top of the existing capabilities of ChatGPT or related LLMs that could reduce the time it takes to complete the task with equal quality by at least half.

- *No Exposure ($X^T = 0$)* if using ChatGPT does not reduce the time required to complete the task by half while maintaining equivalent quality, or using ChatGPT reduces the quality of the task's output.

[11] We begin by equally weighting tasks within an occupation. In Section III.B, we exploit the weights and distinguish between core and supplemental tasks to shed light on the substitution and complementarity of occupation to the Generative AI technology.

[12] This data can be accessed via the O*NET website at https://www.onetonline.org.

Our classification uses the "few-shot prompting" technique and takes three steps. First, we create a system prompt that explains the classification job and describes each category. Second, we create two examples of user-assist prompts that showcase the expected output, where the user prompt asks ChatGPT to classify an example task and to explain its reasoning, while the assist prompt provides the example answers. Third, for each of the 19,265 tasks, T, we feed the GPT model with the system prompt, the two examples of user-assist prompts, and a new user prompt that includes the text of task T's statement and occupation title. The model produces answers similar to the assist prompt in the examples, including the classification and a short explanation, allowing us to audit whether GPT actually understands the prompt as intended and interprets the task correctly. The Internet Appendix C details our prompts and the classification procedure.

This approach using GPT instead of human labeling has several important advantages. First, the GPT model is less subject to individual idiosyncrasies compared to human labeling because it leverages a vast amount of prior information that may be difficult for individual humans to master (Gilardi et al. (2023)).[13] Second, the GPT approach permits rapid scaling of the method to categorize the complete set of 19,265 task statements, which will likely be very time-consuming for human labeling (Frey and Osborne (2017)). Third, the approach also provides explanations for labeling, allowing for an auditing capability often unavailable in instances of human labeling.

Of the 19,265 tasks, 14% were categorized as directly exposed to ChatGPT, $X^T = 1$. Table I provides examples of tasks in this category and explanations. For instance, "adjust sales scripts" for telemarketers and "write supporting codes for web applications" for web developers. We give these tasks an exposure score $X^T = 1$ as they are directly exposed to Generative AI. Another 22% were categorized as exposed to ChatGPT if appropriate software or applications based on the technology were developed, $X^T = 0.5$. For instance, "review financial transactions" for food service managers and "identify (trading) opportunities" for financial services sales agents. Following Eloundou et al. (2023), we give a one-half exposure score for these tasks since their exposure to Generative AI is not as direct and requires further developments of software and applications.[14] The rest of the tasks are categorized as not exposed, $X^T = 0$. For instance, "connect heating or AC equipment" for installers and "mentor new faculty" for postsecondary business teachers.

[13] We further check the model consistency and show in Appendix C that GPT reliably provides classifications that are highly consistent across different runs. We also check the confidence of the classification by requiring GPT to return a confidence score for its prediction, which shows "high confidence" in most cases. The GPT model has been increasingly used to classify content in recent academic studies, e.g., Hansen et al. (2023) and Lopez-Lira and Tang (2023).

[14] The results are similar if we use alternative discounts for the exposure score of tasks in the E2 category.

— Insert Table I about here —

B. *Measuring occupation exposure to Generative AI*

Scoring occupations' exposure to Generative AI We next aggregate tasks' exposures to Generative AI to the occupation level. For each 8-digit Standard Occupational Classification (SOC) occupation from the O*NET, we calculate the share of the total number of tasks for each occupation that have either a direct or "plus-overlay" exposure to Generative AI. Our measure of occupation-level exposure X^O is the sum of task-level exposures X^T for $T \in [0 : 19,265]$ within each occupation $O \in [1 : 923]$ divided by the total number of tasks in occupation O, expressed in percentage. That is,

$$X^O = \frac{\sum_{T \in O} X^T}{\sum_{T \in O} \mathbb{1}}. \tag{1}$$

We then aggregate the 8-digit O*Net occupation codes to 6-digit SOC codes using equal-weighted averages of each 6-digit code's 8-digit sub-codes to match the occupation-level exposure measure to firms' occupational employment data. Generally speaking, this measure captures the percentage of an occupation's tasks that can be more efficiently completed using ChatGPT and similar tools.

C. *Occupation-level stylized facts*

In this section, we present summary statistics describing occupational exposures. We note that 17% of occupations have zero exposure. The average exposure for white-collar occupations is 40%, while it is 9% for blue-collar and service occupations.[15] We also show that Generative AI exposures are higher for occupational skill sets that are more cognitive, and that more exposed occupations tend to be at the higher end of the occupational wage distribution. The latter facts show how the Generative AI substitution effect is likely to operate on a very different part of the labor force relative to prior automation waves.

Panel A of Table II shows that the mean and median Generative AI exposure of occupations, X^O, are 23% and 18%, respectively, with a standard deviation of 21%. The 10^{th} and 90^{th} percentiles of occupational exposure are 0% and 53%. The Internet Appendix Table IA.1 lists the 20 occupations with the highest and lowest Generative AI exposure

[15] This calculation is based on the commonly used classification of major SOC occupation groups 11-29 (Management, Business, Science, and Arts Occupations) and 41-43 (Sales and Office Occupations) as "white collar", and all other major occupation groups as blue-collar or service occupations. The reported percentages are 2022 employment-weighted averages of 6-digit SOC exposures within each group. For details on this aggregation, see https://www.bls.gov/soc/soc_2010_class_and_coding_structure.pdf.

scores.[16] Occupations such as "telemarketers", "proofreaders and copy markers", and "computer programmers" have the highest Generative AI exposure scores. This is intuitive as recent technological advances in Generative AI feature natural language-based conversations, translating texts between languages and styles, and generating functioning code based on high-level descriptions of a programming task. On the other hand, occupations requiring more physical manual tasks, such as "shampooers", "installers and repairs", and "stonemasons" seem to not be exposed to Generative AI.

— Insert Table II about here —

A large body of literature has explored how prior technologies such as computers, automation, and robots have affected different occupations over the past decades (e.g., Autor et al. (2003), Acemoglu and Autor (2011), Acemoglu and Restrepo (2020), among many others). Computerization have been shown to replace primarily routine-task jobs, which are jobs that perform repetitive and codifiable tasks such as production line assemblers and record keepers (Autor et al. (2003) and Acemoglu and Autor (2011)). Robots have been shown to replace primarily programmable manual-task jobs such as machinists, material handlers, and welders (Acemoglu and Restrepo (2020)) and Webb (2019)).

To examine whether occupations exposed to Generative AI are distinct from routine and manual jobs, we follow Acemoglu and Autor (2011) and construct scores for what level of non-routine-analytical, non-routine-interpersonal, routine-cognitive, routine-manual, non-routine-interpersonal, and non-routine-manual skills are required for each occupation. For any given occupation, the skill-measure scores essentially measure how important each of the six skills is for executing its tasks. Unlike our measure of exposure to Generative AI, which is truly task-based, Acemoglu and Autor (2011) use O*NET's importance metrics for several pre-defined skills within occupations to construct the scores for these six skill measures for each occupation. See the Internet Appendix C for the detailed construction of the scores for these six skill measures.[17]

Next, we study whether occupations' exposures to Generative AI are distinct from previously documented occupational exposure to computerization and robots. Specifically, we

[16] Note that the lowest score category only shows an alphabetically sorted subset of a larger set of occupations with zero Generative AI exposure. Overall, 17% of occupations have zero exposure.

[17] For each occupation, the O*NET database not only provides the textual statements of each task the occupation performs, which is the data that we used to construct our Generative AI exposure measure, the database also provides a numerical "importance" score for a large number of pre-defined skills for each occupation. Acemoglu and Autor (2011) select certain pre-defined skills and aggregate them to measure an occupation's skill requirement in each of the six skill dimensions mentioned above. For instance, the measure of non-routine-analytical skill for an occupation is constructed as the average of the standardized importance score of three detailed O*NET skills, *analyzing data/information, thinking creatively,* and *interpreting information from others.*

regress an occupation's Generative AI exposure measure on its six occupational skill scores (Skill_S^O) in the following cross-sectional regression:

$$X^O = \alpha + \sum_S \beta_S * \text{Skill}_S^O + \varepsilon^O. \tag{2}$$

Figure 2 shows the results. We show that occupations with higher Generative AI exposure are more likely to involve non-routine cognitive analytical skills and routine cognitive skills, and less likely to involve other, non-cognitive, skills. Hence, distinct from computerization which disrupts routine jobs, and robots which disrupt manual jobs, Generative AI tends to mostly affect cognitive jobs, in particular those requiring non-routine analytical skills.[18] Hence, one would expect that firms with more cognitive jobs in their occupational portfolio will be more disrupted by Generative AI.

We show in Figure 3 that major occupation groups with higher exposures to Generative AI also tend to have higher wages. This raises the question of whether automation, which affected lower-wage workers, was more of an exception than a rule. Important recent work by Kogan et al. (2023) studies the longer-term impact of breakthrough patents on workers' wages and finds that breakthrough technologies do, in fact, impact high-wage workers more negatively.

— Insert Figure 2 about here —

— Insert Figure 3 about here —

D. Measuring firms' exposures to Generative AI

To measure a firm's exposure to Generative AI, we obtain data on firms' occupational employment from Revelio Labs, which collects information on job titles and employers from LinkedIn and other resume profiles and constructs occupation-by-firm employment counts.[19] For each Compustat firm, we use its employment counts at the 6-digit SOC occupation level as of March 2022, the latest month in our Revelio data. A firm's Generative AI exposure is the weighted average of its occupations' Generative AI exposure, X^O. That is, each firm f in Compustat's labor force exposure to Generative AI is computed as:

$$X^f = \sum_{O \in f} EmpShare_{f,O} \times X^O, \tag{3}$$

[18] Using a similar regression setting, Webb (2019) shows that broadly-defined AI technologies before the recent advances in Generative AI also tended to disrupt cognitive non-routine analytical jobs, but not routine cognitive jobs.

[19] See, for example, Li et al. (2022) and Chen et al. (2023) for more descriptions of the Revelio Labs data.

where $EmpShare_{f,O} = \frac{emp_{f,O}}{emp_f}$ is the employment share of occupation O in firm f. This procedure generates the Generative AI exposures for 2,518 publicly traded firms in 2022. Intuitively, our exposure measure captures the fraction of labor tasks in the firm that can be more efficiently completed using Generative AI.

E. Firm-level stylized facts

Panel B of Table II reports the summary statistics. The mean and median of firms' Generative AI exposures are both 35%, with a standard deviation of 8%. The 10^{th} and 90^{th} percentiles of firm-level exposures to Generative AI are 27% and 44%, respectively, implying that labor exposure to Generative AI is a broad phenomenon.

Table III lists the 15 firms among the top 100 largest U.S. firms by market capitalization with the highest and lowest exposure to Generative AI, respectively.[20] Although many IT firms, such as IBM and Intuit, not surprisingly have a large fraction of employees exposed to Generative AI, we also find manufacturing firms, such as 3M, and administrative conglomerates, such as S&P Global, in the high-exposure list The large U.S. firms ranked at the bottom of the exposure distribution include restaurants, such as Starbucks and McDonald's, retail firms, such as Target and Walmart, transportation firms, such as UPS, and manufacturing firms, such as Tesla, suggesting that these firms' activities require more manual tasks and have a smaller fraction of employees exposed to Generative AI.

— Insert Table III about here —

We next formally investigate the industry component of firms' Generative AI exposures. Figure 4 shows that the variation in firm-level Generative AI exposure has a substantial within-industry component where the industry is classified at the NAICS 3-digit level. While firms in industries such as "data processing, hosting, and related services" and "professional, scientific, and technical services" have high average exposures to Generative AI and firms in industries such as "clothing and clothing accessories stores" and "food service and drinking places" have low average exposures, there is considerable variation of firms' exposures within each industry. A variance decomposition shows that NAICS 3-digit industry differences explain only 31% of the firm-level variation in exposure to Generative AI. The rich within-industry variation in firms' exposures to Generative AI suggests that our labor-based measure captures firms' exposures to Generative AI that can be distinct from their products' exposures to Generative AI. We explore labor vs. product exposure effects in-depth in the next section.

[20] See Internet Appendix Table IA.3 for the full list of the exposures for the 100 largest firms in 2022.

In the Internet Appendix Table IA.4, we also show that our measure of firms' exposures to Generative AI is not explained by other firm characteristics that have been shown to predict stock returns in the cross-section, such as size, Tobin's Q, ROA, labor Intensity, and asset tangibility. Cross-sectional regressions of firms' Generative AI exposures on each of these variables yield adjusted R^2s between 0.6% and 10.7%.

— Insert Figure 4 about here —

II. Generative AI Exposures and Firm value

A. Empirical design

We identify the impact of Generative AI on firm value by studying the relative returns, or changes in market value, of firms with high vs. low Generative AI exposures immediately following the release of ChatGPT. We sort portfolios into quintiles based on stocks' labor exposures to Generative AI and compare the returns of the top "Artificial" quintile to the bottom "Human" quintile. We show that an AMH portfolio that is long the highest exposure firms and short the lowest exposure firms earns 0.45% daily in the two weeks following the release of ChatGPT (0.44% if we control for market exposures). We argue, and provide supporting evidence, that these changes in value are consistent with changes in the expected future cash flows of firms based on labor cost savings.

By exploring changes in firms' market value, i.e., stock returns, our empirical design uses a first-difference method to control for time-invariant firm-level heterogeneity. By studying the alphas of the portfolios from asset pricing factor models, our setting controls for firms' exposures to time-varying priced risk factors. Finally, because we study relative returns, our study highlights the differences in differences of firm value following the release of ChatGPT across firms whose labor forces are more vs. less exposed to Generative AI. To rule out alternative channels, we corroborate our portfolio sorting results with standard event study regressions of stocks' cumulative abnormal returns (CAR) on firms' Generative AI exposures while controlling for other firm characteristics.

We start by defining our event window. We choose the two weeks from November 30, 2022 to December 14, 2022 (the release date and subsequent ten trading days) as our event window, and we label these two weeks as the "**ChatGPT event period**." In choosing the event window, we make a tradeoff. On the one hand, we want to have a window that is long enough for investors to digest the full cross-section of labor implications on firm value. On the other hand, we want the window to be short enough to prevent our estimated stock returns from being contaminated by other related events. Figure 5 plots the daily mentions

of "ChatGPT" or "GPT" on Twitter around the ChatGPT public release date on November 30, 2023. The figure shows that the daily mentions reached over 100K in just a few days after the release of ChatGPT and remained at 50K two weeks after the release. This pattern of attention is consistent with the release of ChatGPT immediately garnering massive public attention, but also with continued discussion and information processing over a longer time period.[21] In the Internet Appendix we show that our results are robust to using a shorter, one-week, event window.[22]

— Insert Figure 5 about here —

B. Realized returns and Generative AI exposures

Portfolio sorts To measure the effect of firms' workforce exposures to Generative AI on changes in value during the ChatGPT event period, we form five portfolios based on our firm-level Generative AI exposure measure. Specifically, at the beginning of our sample period on November 29, 2022, we sort stocks traded on the NYSE exchange into five quintiles based on their Generative AI exposures. We use these NYSE breakpoints to assign non-NYSE stocks into the quintile bins. We compute the value-weighted daily returns of each portfolio as the average daily returns of stocks in the portfolio weighted by their previous day's market capitalization.[23] We refer to the high-minus-low portfolio as the "Artificial Minus Human" (AMH) portfolio, which represents the zero-net-investment portfolio that goes long high-exposure (A) stocks and shorts the low-exposure (H) stocks.

Panel A of Table IV shows the realized excess returns, i.e., the raw daily returns minus the daily risk-free rate, of the quintile portfolios sorted by firms' Generative AI exposures and also the long-short AMH portfolio during the ChatGPT event period. The AMH portfolio yields positive daily returns of 0.45% (t-statistic = 3.53) on average during the two weeks after ChatGPT's release.

One concern might be that our results are driven by firms' differential exposures to risk factors. We show that this is not the case. In Panel B of Table IV, we report the alpha of each portfolio after controlling for the market factor (i.e., the CAPM model). The market-adjusted alpha of the AMH portfolio shows very similar results with a point estimate of 0.44% per day (t-statistic = 4.70), suggesting that our main finding is not driven by firms'

[21] ChatGPT amassed 100 million users in just two months after its release, a milestone which took the World Wide Web 7 years, WhatsApp 3.5 years, Instagram 2.5 years, and Twitter 5 years to reach. See https://economictimes.indiatimes.com/news/new-updates/chatgpt-witnesses-massive-rise-chatbot-gains-100-million-users-in-two-months/articleshow/98428443.cms?from=mdr.

[22] See Appendix Table IA.5.

[23] See Appendix Section C for more details on the portfolio construction.

heterogeneous exposures to the market. In Panel C, we further control for the Fama-French 5-factor model (Fama and French (2015)) and again find similar results. The 5-factor adjusted alpha of the AMH portfolio is 0.35% per day (t-statistic = 3.85).

A related issue arises if the differential stock returns between the high and low-Generative AI exposure portfolios are driven by different trends in the stock returns of the two portfolios ex-ante. We show that this is also not the case. In the Internet Appendix Table IA.6, we conduct a placebo test and examine the excess returns and alphas during periods immediately before and after our ChatGPT event period. We do not find significant differences between the high and low-Generative AI exposure portfolios. This absence of statistically significant differences in returns outside of the ChatGPT event period supports our main findings during the ChatGPT event period being due to the impact of the release of ChatGPT.

— Insert Table IV about here —

Time series of cumulative abnormal returns We further visualize the differential impact of the ChatGPT release on high and low-Generative AI exposure portfolios by plotting the difference in cumulative abnormal returns (CAR) between portfolios in Figure 1. To do so, we first compute the cumulative abnormal returns of the top and bottom quintile portfolios in Table IV using November 29, 2022 as the reference date.[24] This figure represents the holding period returns of the long-short AMH portfolio at any given date in the two weeks before and three months after the ChatGPT event period. It is reassuring that we do not find a strong trend in the AMH portfolio either immediately before or after the ChatGPT event period, suggesting that the market actively reassessed the impact of Generative AI on firm value via our labor channel during the two weeks after the release of ChatGPT when the Twitter mentions of ChatGPT were exceptionally high. In Appendix Figure IA.1, we extend the pre-period out to 3 months before the release and also observe no systematic pre-trend in the AMH returns. Moreover, the cumulative abnormal returns of the AMH portfolio remain high and do not reverse after the event period, suggesting that the revaluation is not short-lived. In fact, the AMH portfolio seems to have experienced another period of outsized returns in early March leading up to the release of GPT 4, in line with the anticipation of the release and functionality of this more capable model representing an additional increase in the expected productivity impact of Generative AI on exposed firms.[25]

[24] Following the literature, we compute each portfolio's daily abnormal return as the portfolio's daily return minus the product of each portfolio's market beta (computed over the 6 months before the time window shown in the figure) and the daily market return, and we compute the CAR for each portfolio by accumulating its daily abnormal returns from November 29, 2022.

[25] While the time series of returns suggests that the revaluation impact for the AMH portfolio may have been similar in the run-up to the GPT 4 release, we cannot repeat our event study analysis in that case, as the

Event study regressions controlling for firm characteristics We show that our finding of increased value for firms with greater labor exposures to Generative AI during the ChatGPT event period is not due to the exposure measure being correlated with other firm characteristics. We show that our results are robust to controlling for the impact of standard firm characteristics used to explain the cross-section of firm value on our portfolio returns during the ChatGPT event period. Specifically, we run a standard event study regression of firms' cumulative abnormal returns in excess of market returns during the ChatGPT event period on their labor-based Generative AI exposure, and also firm characteristics, in the following specification:

$$\text{CAR}_i = \beta \times \text{GenAI Exp}_i + \gamma \times \text{Char}_i + \varepsilon_i, \tag{4}$$

where CAR_i is firm i's cumulative abnormal return during the ChatGPT event period (from November 30, 2022 to December 14, 2022), $GenAI\ Exp_i$ is the firm's labor-based Generative AI exposure, and $Char_i$ is firm characteristics, including size, Tobin's Q, ROA, labor Intensity, and asset tangibility. We are interested in β, which measures the impact of the ChatGPT release on the value of firms with a higher labor-based Generative AI exposure relative to other firms while allowing a differential impact on firms with particular characteristics. We utilize weighted least squares, weighting each stock by its market capitalization as of November 29, 2022, to be consistent with the value-weighted portfolio sorts.

Table V shows the results. Column (1) provides the benchmark result without controlling for firm characteristics, which shows that firms with higher Generative AI exposures experienced a greater boost in firm value during the ChatGPT event period, consistent with our portfolio-sorting results. The magnitude of the estimate suggests that a 1 standard deviation increase in the firm-level Generative AI exposure is associated with 1.7 pp higher cumulative abnormal returns over the event period, which is comparable to the effect size estimated using the portfolio sorts.[26] Columns (2)-(8) show that the effect of our Generative AI exposure remains highly robust with coefficients similar to the benchmark estimate after controlling for the effects of each firm characteristic separately or the effects of all firm characteristics jointly.[27]

existence of the model was widely known in advance, and only the uncertainty about the exact capabilities and precise release date for the more advanced model was resolved during the run-up.

[26]The returns for a 1-standard-deviation change are computed as $0.078 \times 21.6 \approx 1.7$. Note that the market cap weighted exposure scores of the firms in the highest and lowest quintile portfolios are 44.5% and 26.1%, respectively, so the firm-level CAR[-1,10] event study estimate in Table V would predict a daily return difference of $(.445 - .261) \times 21.6/12 = 0.33$ pp for the portfolios, which is not far from the 0.44 pp market-adjusted long-short returns estimated in Table IV.

[27]The Internet Appendix Table IA.8 shows that the results are also robust to controlling for NAICS 3-digit industry fixed effects and using alternative models for computing abnormal returns.

— Insert Table V about here —

C. Product exposure vs. labor exposure to Generative AI

A key concern is that labor market effects may be confounded by firms' product market exposures to Generative AI. The stock-market value of chip-maker NVIDIA more than doubled from the beginning of 2023 to September 19, 2023, which has been widely interpreted as being due to the company being a key supplier of the chips used in data centers and for training Generative AI models. We perform several tests to show that our labor-exposure results are very robust to controlling for firms' product-market exposures to Generative AI. To be clear, we do not claim that the ChatGPT release does not impact firms' products and services. Instead, we provide strong support for the idea that the differential returns we document for firms with different labor force exposures to Generative AI are not driven by the product exposure channel.

C.1. Importance of within-industry heterogeneity

Our first robustness check controlling for the product exposure channel examines the within-industry variation of firms' Generative AI exposure in our portfolio sorting test. This test takes the industry classification as a categorization of the product market. Hence, to the degree that product markets align with detailed industry definitions, a significant industry-neutral AMH portfolio return would suggest that our main findings in Section II.B are not driven by firms' product exposures to Generative AI.

We have shown earlier in Figure 4 that firms' exposures to Generative AI vary substantially within-industry. Indeed, only 31% of firm-level variation in labor force exposure to Generative AI can be explained by 3-digit industry fixed effects. Here, we sort firms into portfolios based on their Generative AI exposures within industries. We consider two industry classifications. The first one is the NAICS 3-digit industry classification, which includes 46 categories with at least ten firms in our data. Hoberg and Phillips (2016) show that the rich texts of firms' product descriptions in their 10K filings can more precisely separate product markets. We thus also adopt their 10K-based FIC50 industry classification as an alternative measure of industries. Given the more limited number of stocks within industries, we form tercile portfolios (instead of quintile portfolios) within each industry and then value-weight stock returns in the industry-neutral tercile portfolios.

Table VI shows the results. Panels A1-A3 show the market-adjusted alphas for the tercile portfolios sorted by Generative AI exposures across all firms unconditionally, within NAICS 3-digit industries, and within FIC50 industries, respectively. Using Panel A1 as the

unconditional benchmark for the tercile portfolios, Panels A2 and A3 show that the within-industry AMH portfolios using both industry definitions also have large and statistically significant market-adjusted alphas during the ChatGPT event period. The magnitudes of the within-industry alphas are also very similar to the alpha of the unconditionally sorted AMH portfolio in Panel A1. Panels B1-B3 show similar results when we examine the FF5-factor-adjusted alphas instead of the market-adjusted alphas.

In summary, if industry classifications capture the segmentation of firms' products and services, then the within-industry findings suggest that the impact of Generative AI on firm value, as we described in our main findings in Section II.B is not driven by firms' products' differential exposures to Generative AI.

— Insert Table VI about here —

C.2. Results excluding the tech sector

Our results are also robust to excluding the tech sector altogether. We follow Acemoglu et al. (2022) and identify firms as potential tech firms if they are in the following two sectors: information sector (NAICS 51) and professional, scientific, and technical services sector (NAICS 54).[28] In Panel D of Table IV, we report our baseline quintile sorting results excluding firms from these two sectors. We document a strikingly similar result to our main results when we exclude the tech sectors. In particular, the FF 5-factor-adjusted alpha for the AMH portfolio using non-tech firms is 0.38% per day during the ChatGPT event period as compared to 0.35% per day using all firms. These results, along with our within-industry results, confirm that the labor impact of Generative AI is broad, and that Generative AI is a general-purpose technology.

C.3. Horseraces between labor vs. product exposures

Next, we run horserace tests between our labor exposure to Generative AI and direct measures of product market exposures. By constructing these measures, we aim to directly identify firms that can benefit from an AI boom either because they are selling related products (e.g., cloud computing hardware and graphics processing units) or because they are using existing versions of AI-related technologies as direct inputs into their products and services. Our product exposure measures may be of independent interest since we show they are related to firm returns; however, importantly, our labor exposure effects remain even when we control for these product market effects.

[28] The 2023 U.S. Census Business Trends and Outlook Survey shows that these two sectors have the highest use of AI in their products. See https://www.census.gov/library/stories/2023/11/businesses-use-ai.html.

Our first measure of firms' product exposure to Generative AI follows Hoberg and Phillips (2016) and many other prior studies and infers information about a firm's products from the business description section of a firm's most recent 10-K annual report. We then use GPT 3.5 Turbo to assess whether firms' business products involve enabling or scaling new Generative AI technologies or benefit from a direct incorporation of the new Generative AI capabilities as a functionality of their products.[29] We provide details in the Internet Appendix C. This procedure generates a dummy variable that equals 1 if the answer is yes–the firm's annual report suggests product market exposure to Generative AI, and 0 otherwise. We label this measure *GPT10K Product Exposure*.

Our second measure counts the number of occurrences of AI-related keywords in the firm's business section in its 10-K report. This keyword-count approach is consistent with the approach used in prior studies (e.g., Webb (2019) and Babina et al. (2024)), and it generates a continuous measure instead of the binary value obtained in our first product exposure measure. We label this AI keyword-based measure *Count10K Product Exposure*.

Our third measure adopts the list of firms recently classified as "near-term beneficiaries of AI" by the investment bank Goldman Sachs.[30] This measure identifies stocks with business models that are directly exposed to the development of AI and which, as a result, are expected to experience an immediate increase in earnings. The list includes makers of semiconductors and related equipment, and large technology companies with extensive cloud computing infrastructure or business models that are likely to benefit from incorporating AI capabilities into their products. This measure is explicitly focused on *ex post* identifying stocks with high recent returns due to their earnings potential increasing as a result of the AI boom, and it may thus account for firms' product exposure to AI beyond firms' discussion in past annual reports. We label this measure from Goldman Sachs *GS Product Exposure*.

Finally, our fourth measure uses the share of workers at a firm that has AI skills on their resumes constructed by Babina et al. (2024), who show that this measure effectively predicts AI-related product innovations and R&D spending during the pre-ChatGPT wave of AI advancements. We use a firm's most recently available AI skill share in the data from Babina et al. (2024) to proxy for its prior investment in using AI-related tools pre-ChatGPT and its susceptibility to adopt new AI tools in their products post-ChatGPT.[31] We label this

[29] Conceptually, this product-exposure channel differs from our labor-exposure channel as the former channel does not directly imply an impact of Generative AI on the firm's operation and production process.

[30] See Goldman Sachs US Equity Views, August 21, 2023, "*The (AI) trade after the trade: Identifying potential long-term EPS beneficiaries of AI adoption*" available from https://research.gs.com/. For a publicly accessible write-up of the results of the study, see, for instance, https://markets.businessinsider.com/news/stocks/goldman-sachs-reveals-long-term-ai-portfolio-here-are-the-50-stocks-to-monitor-1032573922.

[31] We thank Babina et al. (2024) for making the data publicly available. The last year of available skill share

measure from Babina et al. (2024) *BFHH Product Exposure*.

How good are these four proxies for firms' product exposure to Generative AI in capturing the potential for firms' adopting the recent AI in their products and services after the release of ChatGPT? To answer this question, we use data from the U.S. Census Business Trends and Outlook Survey, which reports data on the share of firms using "Artificial Intelligence (AI) in producing goods and services" for a number of industries after August 2023.[32] We thus aggregate our four proxies of product exposure to Generative AI to the industry level and compare our proxies with the intensity of industries' actual use of AI in their products after the ChatGPT release. Internet Appendix Table IA.10 shows that all four proxies correlate strongly with the Census survey results at the industry level. Hence, we argue that these four proxies capture firms' product exposure to Generative AI reasonably well. Moreover, they capture different dimensions of product exposure, as evidenced by the fact that they are only modestly correlated with one another - see Appendix Table IA.9.

We investigate whether our finding of increased value for companies with greater labor-based Generative AI exposure is confounded by our measure's relation with firms' product exposure to Generative AI. Specifically, we run an event study regression of firms' cumulative abnormal returns during the ChatGPT event period on their labor-based Generative AI exposure and also various proxies for firms' product exposure to Generative AI:

$$\text{CAR}_i = \beta \times \text{GenAI Exp}_i + \gamma \times \text{Product Exp}_i + FEs + \varepsilon_i, \tag{5}$$

where CAR_i is firm i's cumulative abnormal return during the ChatGPT event period, $GenAI\ Exp_i$ is the firm's labor-based Generative AI exposure, and $Product\ Exp_i$ is the proxy for the firm's product exposure to Generative AI. We are interested in β, which measures the impact of the ChatGPT release on the value of firms with a higher labor-based Generative AI exposure relative to other firms while allowing for a differential impact on firms with a particular product exposure to Generative AI. Similar to our estimation in equation (4), we weigh each stock by its market capitalization as of November 29, 2022. Finally, we include NAICS 3-digit industry fixed effects so that the horseraces are run within industries, further controlling for any remaining product market effects.

Table VII reports the results of the cross-sectional regression in equation (5). Column (1) reiterates the within-industry importance of our effects documented in Section II.C.1 by

data is 2021 for most firms in the sample. While their measure also derives from firms' labor heterogeneity, their approach uses keywords to account for firms' employees' AI skills, and their measure aims to capture firms' AI investment. In contrast, our labor-based Generative AI exposure accounts for firms' employees whose tasks can be more efficiently completed by using Generative AI, and thus the firms' *labor exposure* to Generative AI.

[32] See details about the survey and data at https://www.census.gov/hfp/btos/about.

showing that the coefficients of our labor-based Generative AI exposure are similar when controlling for NAICS 3-digit industry fixed effects. Columns (2)-(6) show the horserace results by controlling for firms' product exposure to Generative AI. Two observations stand out: First, three out of four proxies for firms' product exposure positively boost firm value during the ChatGPT event period. This result supports findings in the prior literature that AI technologies significantly improve firms' product functionality and innovations (Babina et al. (2024)).

Second, Table VII shows that firms' labor exposure to Generative AI is robustly related to increases in firm value during the ChatGPT event period even after controlling for the proxies for firms' product exposure to Generative AI individually in Columns (3)-(5) or jointly in Column (6). The economic magnitude of the CAR associated with the labor exposure to Generative AI is only modestly attenuated, suggesting a 1.4 pp higher event period CAR for firms with a 1 SD higher Generative AI exposure, even when controlling for all product market exposure proxies jointly.

— Insert Table VII about here —

D. The role of firms' data assets

A salient feature of Generative AI is its superior capability to learn from and work with firms' data to upgrade their operations and save labor costs, such as by training customer service chatbots, automating workflows, improving predictions and analytics, and performing many other applications (Caserta et al. (2023)). For instance, to adopt Generative AI tools and improve the chatbots for internal customer service knowledge management, firms need internal data from past interactions and the technology to upgrade existing labor-intensive customer service systems.

In this section, we test whether our stock return results are more pronounced among firms relying more heavily on data assets, which can better incorporate Generative AI technologies. Measuring firms' data assets is challenging (Veldkamp (2023)).[33] We construct two proxies for firms' reliance on data assets. Our first measure uses the text of the business description in a firm's 10-K annual report. We again apply the GPT model to assign a score of 0 to 3 to indicate whether the firms' business descriptions provide no, little, moderate, or high evidence of data that would be valuable for Generative AI-based analysis. The model is instructed to base its assessment on mentions of six data-related topics such as data

[33]See seminal works on the measurement of firms' data assets from Begenau et al. (2018), Farboodi et al. (2019), Eeckhout and Veldkamp (2022), Farboodi and Veldkamp (2023), Abis and Veldkamp (2023), among others.

collection, data utilization, data infrastructure & management, and data regulation and privacy (see more details in the Internet Appendix C). The overall score assigned by the model is our first measure of the firm's reliance on data assets. We label this proxy *10K Data Assets*.

Our second measure follows Abis and Veldkamp (2023) and calculates firms' share of labor skills in data management. In particular, we first measure the likelihood that an occupation is a "data management" position by analyzing whether the requirements listed in each occupation's job postings in recent years correspond to data management skills classified by Abis and Veldkamp (2023). Job postings are classified as suggesting a data management-intensive role if at least three of these skills are mentioned. For each occupation, we compute the share of data management-intensive job postings as a proxy for the propensity of that occupation to be a data management position.[34] Then, we aggregate the data management intensity from the occupation level to the firm level using the firm's occupational employment shares from the Revelio Labs database, similar to our calculation of firms' Generative AI exposure (see more details in the Internet Appendix C). The intuition for this measure is that whether a firm has many workers responsible for managing data is a good proxy for whether it has a lot of valuable data to be managed–and for the firm's ability to handle data-based technologies in general. We label this proxy *AV Data Assets*.[35]

Equipped with the two proxies for firms' reliance on data assets, we test whether our stock return results are more pronounced among firms with a greater reliance on their data assets. In particular, we regress firms' cumulative abnormal returns on the interaction between their Generative AI exposure and the proxies, one at a time. Table VIII shows the results. The positive coefficients of the interaction terms suggest that firms with greater access to data assets have significantly higher returns to Generative AI exposure during the ChatGPT event period. This finding supports our conjecture that firm value increases disproportionately accrue to firms with labor exposure to Generative AI *and* significant data assets, allowing them to better adopt Generative AI technologies and upgrade their labor operations. We also note that data assets can be firm-specific, making existing incumbent firms' potential gains from adopting Generative AI larger, and less likely to be competed away by new entrants. Our finding that the stock returns to Generative AI exposure during the ChatGPT event period are greater among firms with greater reliance on data assets is consistent with this view.

— Insert Table VIII about here —

[34] Occupations' data management intensity is 9% correlated with their Generative AI exposure.

[35] See Appendix Table IA.9 for the raw correlation of the data asset measures with the firm-level Generative AI exposure, and the product market exposure measures.

III. Generative AI: Labor Complement or Substitute?

In this section, we shed light on the mechanism for our findings on firms' market value by examining labor-market and firm profitability outcomes following the release of ChatGPT. First, we show that, overall, Generative AI seems to be a substitute for labor more exposed to its technology. At the occupation level, wages and job postings decline, and we show that job postings also decline at the firm level for firms with more exposed workforces. Within firms, we show that it is the more exposed occupations for which the declines in job postings occur. Next, we examine whether there may be any highly exposed occupations that may experience Generative AI as more of a complement than a substitute. We develop a novel way to address this using O*NET's classification of tasks as *core* or *supplemental*. We hypothesize that occupations whose core tasks are more exposed to Generative AI are more likely to experience labor substitution, while workers in occupations whose supplemental tasks are more exposed may be more able to leverage complementarities. In support of this, we find that the negative effects on wages and job postings are much stronger for occupations whose high exposures stem from core-task exposure vs. supplemental task exposure. Finally, we link our findings to firm profitability. We show that firms with higher labor exposures to Generative AI experience larger increases in profitability following the release of ChatGPT, and that the increase in profitability is stronger for firms whose labor force exposure to Generative AI stems from core-task exposure.

Detecting changes in the labor market and real corporate outcomes shortly after the arrival of new technologies is challenging: A large body of literature shows that the diffusion of major technologies in history is usually slow due to various frictions, including limited attention of businesses to the new technologies (Greenwood (1999)).[36] Previously, we showed that the release of ChatGPT garnered massive immediate public attention. Here, we document that the release of ChatGPT also attracted the attention of *firms*. In Figure 6, we measure firms' mentions of Generative AI in their quarterly earnings calls. Panel A shows that the share of firms mentioning Generative AI in their quarterly earnings calls rose substantially from less than 5% before ChatGPT's release to 27% in the first quarter after the release.[37] More importantly, Panel B shows that firms with a *higher* Generative AI exposure see a *stronger* increase in the likelihood of mentioning Generative AI after ChatGPT's release (see the Internet Appendix C for details on this estimation). These findings suggest that

[36]The diffusion of new technologies is particularly slow in the initial periods after the technologies are available, constituting an S-curve for technology diffusion (see many examples in Greenwood (1999) and Manuelli and Seshadri (2014)). As a result, many seminal studies on the impact of major technology shocks were highly retrospective rather than conducted soon after the arrival of the technologies.

[37]In contrast, the share of firms mentioning other related topics, such as engineering, does not increase. See the Internet Appendix Figure IA.3 for this result.

the release of ChatGPT likely garnered the attention of firms to evaluate how Generative AI can be applied to affect their labor, operation, and profits. We test these real implications in this section.

— Insert Figure 6 about here —

A. Hypothesis development

A large body of literature has studied the complementarity and substitution effects of technologies on workers in the context of automation (Autor et al. (2003), Autor et al. (2006), Frey and Osborne (2017), and Acemoglu and Autor (2011)), robots (Acemoglu and Restrepo (2020)), artificial intelligence (Webb (2019), Agrawal et al. (2019), and Babina et al. (2022)), and disruptive technologies in general (Krusell et al. (2000), Kogan et al. (2019) and Bloom et al. (2021)). While some technologies can complement labor in that they improve the productivity of the workers, other technologies can substitute labor in that they erase the necessity for humans to work on those jobs due to the low cost of using more capital-intensive technologies. The same technology can also have both effects at the same time, depending on the nature of the jobs under consideration. For instance, in influential work by Autor et al. (2003) studying the labor impact of computer-driven automation, the production function is specified as $Q = (L_R + C)^{1-\beta} L_N^{\beta}$, where computer capital C substitutes for routine-task occupations, L_R, and complements non-routine-task occupations, L_N. See also Acemoglu et al. (2022) for a richer production function conceptualizing the heterogeneous effects of artificial intelligence on different types of labor.

We leverage these ideas from the prior literature to make two key points. First, we argue that both the labor complementing and labor substituting mechanisms can boost the value of firms with a higher Generative AI exposure when ChatGPT becomes available. If ChatGPT complements the occupations with a high Generative AI exposure, then, firms with a high share of such occupations, i.e., firms with a high Generative AI exposure, can experience a boost in their labor productivity, which increases their future profits and thus positively impacts their current firm value. If ChatGPT substitutes for occupations with a high Generative AI exposure, then firms with a high Generative AI exposure can potentially also save labor costs by adopting the new technology, which again increases their future profits and boosts their firm value. Note the above argument relies on increases in future cash flows not being fully appropriated by workers or competed away in the product market. This assumption is consistent with studies documenting that firms may possess some monopsony power in the labor market (e.g., Berger et al. (2022), Yeh et al. (2022), and Seegmiller (2021)) and that measures of product market competition in the U.S. have been low and declining

(e.g., Gutiérrez and Philippon (2017a) and Akcigit and Ates (2021)).

On the other hand, the two channels predict different wage and employment trajectories for the exposed occupations, with the labor substituting channel predicting lower labor demand and lower wages for the occupations with a higher Generative AI exposure and the labor complementing channel predicting the opposite effect for both labor demand and wages. Hence, while both channels are consistent with our findings on firm value, we can examine the labor market impact to test which channel drives the increase in market value for firms with a higher Generative AI exposure during the ChatGPT event period.

Our second point is that there is an important distinction between substitution (or complementarity) at the task vs. the occupation level. We argue that even if many tasks in an occupation can be executed by Generative AI, if the *core* tasks require a human, that occupation may be less likely to see a substitution effect. To study this, we propose a novel method to discern whether an occupation with a higher Generative AI exposure is more likely to be substituted or complemented by Generative AI based on whether its exposed tasks are "core" or "supplemental" to the occupation according to O*NET. Core tasks represent the most fundamental duties an employee in that occupation is expected to perform.[38] If ChatGPT (with or without the help of another worker) can more efficiently complete the core tasks at a much lower cost, then it is likely that ChatGPT can displace the occupation as a whole.[39] In contrast, supplemental tasks are additional tasks or duties associated with the core tasks but are not considered critical or central to the occupation. If ChatGPT can more efficiently complete the supplemental tasks, then the technology can free up the time and effort for the worker to spend more time on their core tasks, increasing the worker's productivity. If the economy is satiated with the ability to complete those core tasks, then the technology may still reduce the demand for the occupation. If the economy is not satiated with the core tasks, the technology substituting supplemental tasks makes the occupation more productive and more desirable by firms. We develop two testable predictions for our labor-market analyses.

Prediction 1: Occupations with a higher task exposure to Generative AI, on average, experience lower labor demand and wages after the release of ChatGPT.

Prediction 2: Occupations with Generative AI exposure derived from core tasks experience

[38] O*NET defines a task as *core* for an occupation if the task has a relevance score above 67% and an importance rating above 3.0 for the occupation and as *supplemental* if otherwise. See details at `https://www.onetonline.org/help/online/scales#score`.

[39] This could happen either through a reduction in the number of workers in that occupation needed to complete a constant volume of core tasks, or by "unbundling" as suggested by Agrawal et al. (2023) who note that if AI replaces some tasks that were previously bundled into jobs requiring scarce skilled workers, this might permit the unbundling of related tasks into new jobs for workers without those specialized skills, enabling the firm to employ fewer skilled workers.

even lower labor demand and wages than occupations with Generative AI exposure derived from supplemental tasks after the release of ChatGPT.

B. Occupations' Generative AI substitution versus complementarity

To test the above predictions, we distinguish whether an occupation's Generative AI exposure derives from its core or supplemental tasks. Using the supplemental indicator for each task from O*NET, we measure the share of an occupation's Generative AI exposure deriving from its supplemental tasks,

$$ShareSupp^O = \frac{\sum_{T \in (O|Supplemental)} X^T}{\sum_{T \in O} X^T}. \tag{6}$$

where the numerator is the sum of task exposures to Generative AI for occupation O's supplemental tasks, and the denominator is the sum of task exposures to Generative AI from all of the occupation's tasks. For occupations with a positive Generative AI exposure, on average 23% of the exposure derives from supplemental tasks, on average, while the remainder of 77% comes from core-task exposure (see Table II). The share of Generative AI exposure that derives from supplemental tasks varies substantially across different major occupation groups - see Appendix Figure IA.4 for an illustration. For instance, "Management" occupations have relatively high overall Generative AI exposure of 35%, but 34% of that exposure comes from supplemental tasks that are do not represent core activities, while "Business and Financial Operations" have high overall exposure at 48% on average and 88% of that exposure comes from core tasks.

Occupational labor demand effects To analyze occupational labor demand, we use data from Lightcast (formerly Burning Glass) and construct firms' number of job postings for each occupation in each month from January 2019 to August 2023. We aggregate the number of job postings to the occupation-month level, resulting in a final sample with a balanced panel of 43,512 observations for occupations that ever posted jobs in our time period, where missing observations have been filled with zeros. We then test Predictions 1 and 2 using difference-in-differences regressions.

Test for Prediction 1:

$$y_{j,t} = \chi \times \text{Post}_t \times \text{GenAI Exp}_o + \text{FEs} + \varepsilon_{j,t}. \tag{7}$$

Test for Prediction 2:

$$\begin{aligned} y_{j,t} =& \lambda \times \text{Post}_t \times \text{GenAI Exp}_o \times \text{ShareSupp}_o \\ &+ \theta \times \text{Post}_t \times \text{GenAI Exp}_o \\ &+ \eta \times \text{Post}_t \times \text{ShareSupp}_o + \text{FEs} + \varepsilon_{j,t}. \end{aligned} \tag{8}$$

In these tests, $y_{j,t}$ is the number of job postings at the occupation-month level, $Post_t$ is a dummy variable that equals 1 if the month is after the release of ChatGPT on November 30, 2022, $GenAI\ Exp_o$ is the occupation's Generative AI exposure from equation (1), $ShareSupp_o$ is the share of exposure to Generative AI that comes from supplemental tasks in equation (6), and the *FEs* include fixed effects at the occupation level and year-month level. The fixed effects subsume the standalone variables $GenAI\ Exp_o$, $ShareSupp_o$, their interaction, and $Post_t$. To present the estimates as a semi-elasticity (a percentage change in job postings in response to a unit change in exposure), but without dropping observations when log-transforming the dependent variable, we follow Cohn et al. (2022) and estimate the panel relationship using a Poisson model with fixed effects.[40]

Prediction 1 predicts that $\chi < 0$ in equation (7) as occupations with a higher Generative AI exposure are, on average, more likely to be substituted by ChatGPT. Proposition 2 predicts $\lambda > 0$ in equation (8) because the substitution effects are weaker if the occupations' Generative AI exposure derives more from their supplemental tasks than from their core tasks.

Table IX reports the results. Column (1) shows that occupations with a one-standard-deviation higher Generative AI exposure experience about an 8% ($= 0.209 \times 0.382$) decline in job postings after the release of ChatGPT, relative to less exposed occupations. This result supports Proposition 1 that occupations with a higher Generative AI exposure faced lower demand from firms after the release of ChatGPT as they are, on average, substituted by ChatGPT. It also echoes many corporate reactions we observed after the release of ChatGPT. For example, IBM, the company ranked #1 in our exposure to Generative AI measure among the largest U.S. firms (see Table III), announced on May 1, 2023 that it would halt hiring of 7,800 jobs that could be replaced by AI.[41]

[40] A commonly adopted approach is to take the natural logarithm of the variable before estimating a linear model. However, the natural logarithm of zero is undefined. Prior studies therefore often transform the dependent variable into the logarithm of a constant plus the original variable. However, Cohn et al. (2022) show that estimates using this approach do not generate the correct interpretation and can even produce the wrong signs, while the Poisson model with fixed effects generates the correct interpretation of the estimated coefficient as representing percentage changes in the dependent variable. In particular, we used the PPML Stata model developed by Santos Silva and Tenreyro (2015).

[41] See https://www.businessinsider.com/ibm-halts-hiring-for-7800-jobs-that-could-be-rep

Column (2) reports the estimates of equation (8) and shows that the point estimate of λ on the interaction of Post $\times$ GenAI Exp $\times$ ShareSupp is positively significant, confirming that occupations with supplemental tasks exposed to Generative AI experience less dampening effects in the labor market than occupations with core tasks exposed to Generative AI after the release of ChatGPT. By formulating the triple interaction term using the share of supplemental tasks, the simple interaction term in the first row represents the effect from core tasks. In Column (2), the point estimate of θ is 66% ($= 0.636/0.382 - 1$) greater than the estimates of χ in equation (7) (in Column (1)), suggesting that the substitution effect from ChatGPT is much stronger if we measure an occupation's Generative AI exposure based only on core tasks than all tasks of the occupation. This finding supports Proposition 2 and our key insight that occupations with core tasks substitutable by technology are most exposed to technological displacement.

Within-firm labor reorganization One concern about the results across firms is that more exposed occupations could be suffering in terms of employment and wages because they happen to be more prevalent in firms that are not doing well during this time period. Some tech firms, for example, experienced downsizing in 2023. In that case, our finding that more exposed occupations experience declines in hiring might be unrelated to firms choosing to shift hiring away from these jobs, but rather could be explained by hiring shifting away from the firms that typically hire for exposed jobs. To address this concern, we re-aggregate the detailed job posting data to the occupation-firm-month level for the publicly traded firms for which Lightcast provided a crosswalk of job postings to Compustat firm identifiers. We then run the regressions in equations (7) and (8) while controlling for firm-month fixed effects. By controlling for firm-month fixed effects, we estimate the heterogeneous labor demand across occupations within a firm.

Columns (3) and (4) show that even when we control for firm-level trends in hiring and only consider within-firm variation in hiring, we show a shift away from more exposed occupations of a similar magnitude. This finding suggests that the hypothetical contribution of across-firm compositional shifts in hiring to the relative overall decline in job postings for exposed occupations is small: even when we hold firm hiring trends constant, firms are shifting their hiring away from the exposed occupations – with the same pattern of a larger negative effect for occupations with a larger share of exposure coming from core tasks as we found previously.

laced-by-ai-report-2023-5.

Occupational wage rate effects We corroborate our labor demand findings by further analyzing the impact of Generative AI on workers' wage rates. We do not have firms' worker-level wage data for the period after the release of ChatGPT. We instead use the monthly individual-level data from the Census Current Population Survey (CPS) from January 2019 to October 2023 (the latest available month in the CPS data).[42] We compute the more aggregated Census occupations' Generative AI exposure by applying a crosswalk and then averaging over the exposure of related SOC 6-digit occupations. Our final sample includes the natural logarithm of the hourly wage rates of about 271,000 individual-by-month observations, and their demographic characteristics, including gender, age, age squared, work experience, work experience squared, race (white versus non-white), years of education, occupation, and the CBSA metropolitan code.

We next run regression specifications similar to equations (7) and (8), where we set the dependent variable to be the hourly wage rate of the individual, we include linear controls of the above individual demographics that have been shown in the literature to be related to worker wages, we further include the fixed effects at the CBSA metropolitan level to further control for heterogeneous local labor markets (Moretti (2010) and Tuzel and Zhang (2017)), and we weight the observations using each individual's CPS sampling weight following the literature. We still run Poisson regressions for the above specifications, and hence, the coefficients should be interpreted as percentage changes in hourly wages.

Columns (5) and (6) of Table IX show results highly consistent with our labor demand findings. First, Column (5) shows that occupations with higher Generative AI exposure experience a greater relative decline in wage rates after the release of ChatGPT. A one-standard-deviation increase in Generative AI exposure is associated with a 0.6% ($= 0.209 \times 0.031$) decline in the wage rate. The smaller magnitude of the wage effect compared to that of the job posting effect is expected in that wages are known to be sticky (Barattieri et al. (2014)). Second, Column (6) shows that the impact of Generative AI exposure on wage rates is 97% ($= 0.065/0.033 - 1$) greater than the average effect if the exposure derives purely from the occupation's core tasks' susceptibility to technology displacement. This finding further supports Proposition 2 that occupations with core tasks substitutable by technology are more negatively affected by technology than occupations with supplemental tasks substitutable by technology.

— Insert Table IX about here —

[42] From the CPS data, we extract each individual's hourly wage rate, occupation code, gender, age, race, education level, and sampling weight in the survey. We require the individuals to be between 18 and 65 years old, employed in the month of the survey, and to have a non-missing hourly wage rate and non-missing occupation code. The data can be downloaded at https://cps.ipums.org/cps/.

C. *Firms' operational reaction and Generative AI exposure*

Our findings in the previous section suggest that the release of ChatGPT reduced the demand for occupations with tasks exposed to Generative AI in the U.S. economy, and more so for occupations with core tasks exposed to Generative AI. This finding has direct implications for the cross-section of *firms*: Because firms with a higher Generative AI exposure possess a higher share of such occupations, Prediction 1 implies that the release of ChatGPT provides a greater opportunity for these firms to substitute the exposed labor with Generative AI and save costs than for other firms. Furthermore, this cost-saving opportunity is particularly pronounced if these firms' high Generative AI exposure is derived from workers whose core tasks are exposed to Generative AI, as implied by Prediction 2. The cost-saving opportunities brought by the release of ChatGPT can drive up firms' future profitability and, in turn, their current market valuation shown in Section II. In this section, we examine this cost-saving mechanism by analyzing firms' job postings for highly-exposed occupations, profitability, and stock returns.

To prepare for these tests, we measure the share of a *firm*'s Generative AI exposure deriving from its workers' supplemental tasks' exposure as

$$ShareSupp^f = \frac{\sum_{O \in f} EmpShare_{f,O} \times X^O \times ShareSupp^O}{\sum_{O \in f} EmpShare_{f,O} \times X^O}, \tag{9}$$

where $ShareSupp^O$ is occupation O's share of Generative AI exposure derived from its supplemental tasks' exposure to Generative AI in equation (6), summing only over occupations with non-zero exposure.

Firms' labor demand and Generative AI exposure A prediction consistent with the conceptual framework we present is that firms with a higher Generative AI exposure have a greater cost-saving opportunity which they can take advantage of by reducing their reliance on labor to do the tasks in highly exposed occupations. We thus define an occupation as highly exposed to Generative AI if the occupation's Generative AI exposure is within the top tercile of the distribution across all occupations, labeled as *High-GenAI Exposure* occupations. From the Lightcast job posting data, we calculate firms' number of job postings of High-GenAI Exposure occupations in each month. We further require the firms to be in our stock return test sample, resulting in a final sample of 36,900 firm-month observations from January 2022 to August 2023.

Our first set of tests of the cost-saving mechanism estimates firms' heterogeneous demand

for highly exposed occupations using the following two difference-in-differences regressions:

$$\text{High-Exp Job Posting}_{it} = \beta \text{Post}_t \times \text{GenAI Exp}_i + \text{FEs} + \varepsilon_{it}, \tag{10}$$

and

$$\begin{aligned}\text{High-Exp Job Posting}_{it} = {} & \gamma \text{Post}_t \times \text{GenAI Exp}_i \times \text{ShareSupp}_i \\ & + \theta \text{Post}_t \times \text{GenAI Exp}_i + \zeta \text{Post}_t \times \text{ShareSupp}_i + \text{FEs} + \varepsilon_{it},\end{aligned} \tag{11}$$

where *High-Exp Job Posting*$_{it}$ is firm i's job posting of *High-GenAI Exposure* occupations, *GenAI Exp*$_i$ and *ShareSupp*$_i$ are firm i's Generative AI exposure and the share of its Generative AI exposure deriving from its workers' supplemental tasks' exposure, respectively, and *FEs* includes year-month fixed effects and firm fixed effects which subsume the standalone variables *Post*$_t$, *GenAI Exposure*$_i$ and *ShareSupp*$_i$.

Column (1) of Panel A in Table X confirms that firms with a higher Generative AI exposure reduce their demand for high-GenAI exposure occupations after the release of ChatGPT. Moreover, Column (2) confirms the reduced labor demand reaction is more than twice as great (i.e., comparing the coefficients -3.038 and -1.378) for firms with their Generative AI exposure derived entirely from their workers' core tasks' exposure. These findings suggest that firms with a greater share of workers that are highly substitutable by Generative AI possess a greater cost-saving opportunity after the release of ChatGPT. We next examine whether the heterogeneous cost-saving responses manifest in firms' profitability.

Firms' profitability and Generative AI exposure We compute firms' quarterly profitability using the most recent Compustat quarterly data, where profitability is the firm's revenue minus the cost of goods sold normalized by its total assets (Novy-Marx (2013)).[43] Requiring the firms to be in our stock return tests, we obtain 14,621 firm-quarter observations from 2022Q1 to 2023Q4.

Columns (3) and (4) of Panel A in Table X show results on firms' profitability that are consistent with the decline in labor demand that we documented previously, and with resulting cost-savings. In particular, we document that firms with a higher Generative AI exposure experience a relative increase in profitability after the release of ChatGPT. Moreover, this effect is more than three times greater for firms with their Generative AI exposure derived entirely from their workers' core tasks' exposure.

[43] Novy-Marx (2013) highlights that this gross profitability is "the cleanest accounting measure of true economic profitability. The farther down the income statement one goes, the more polluted profitability measures become."

Firms' market value and Generative AI exposure Finally, we revisit firms' stock market reactions to the release of ChatGPT during the event window by further distinguishing firms with their Generative AI exposure derived from their workers' core tasks' exposure versus their supplemental tasks' exposure. In particular, we estimate the event study regression on cumulative abnormal returns in equation (4) by further interacting $GenAI\ Exposure_i$ with $ShareSupp_i$:

$$\text{CAR}_i = \gamma \times \text{GenAI Exp}_i \times \text{ShareSupp} + \theta \times \text{GenAI Exp}_i + \zeta \times \text{ShareSupp}_i + \varepsilon_i. \quad (12)$$

Panel B of Table X shows the results. Compared to the benchmark estimates using all firms in Column (1), Column (2) shows that the effect of Generative AI exposure on firm value during the ChatGPT event period is a full 75% greater if firms derive their Generative AI exposure entirely from their workers' core tasks' exposure. This finding supports the idea that the increase in firm value stems largely from potential cost savings from substituting labor whose core tasks can be efficiently executed by Generative AI tools.

In summary, the consistent results across the board of our tests on firms' labor demand, profitability, and market value collectively provide strong support for the cost-saving mechanism: Firms with a higher labor exposure to Generative AI can reduce their reliance on the exposed workforce and increase their future profitability after the arrival of the Generative AI technology. Consistent with this cashflow channel, we document that these firms experience a more significant relative increase in market value during the ChatGPT event period. Further reinforcing the cost-saving mechanism, we show that all effects are more significant if we focus on firms where their labor-based Generative AI *exposure* more cleanly indicates Generative AI *substitution*, i.e., the exposures derive from their workers' core tasks' substitutability by Generative AI.

— Insert Table X about here —

IV. Conclusion

Market prices indicate that the arrival and diffusion of large language models and Generative AI represent a major technology shock with important effects on the overall value of firms, leading to winners and losers. This paper uses occupational exposures to Generative AI, along with firm-level measures of occupational composition, to assess the exposure to Generative AI innovations at the firm level for publicly traded U.S. corporations. We find that the effect of the release of ChatGPT on firm value was large, driving a difference in firm returns of 0.45% daily. Moreover, we provide evidence that firms with higher exposure

to Generative AI-driven productivity increases are more likely to communicate with their investors about these technologies in their earnings calls, and differ in their hiring behavior in recent months, reducing their overall job postings, as well as the relative share of the highly exposed occupations in their job postings. We also document a relative decline in wages and hiring for highly exposed occupations at the national level in recent months, with a more negative impact for occupations where Generative AI affects core rather than supplemental tasks. These findings raise important issues for policymakers: to the degree that shareholders benefit from increases in firm value while affected workers lose out, the new technology redistributes income in the economy, which may or may not be desirable from a welfare perspective. At the same time, our quantification of the productivity potential of Generative AI for different firms permits policymakers and entrepreneurs to better identify areas of opportunity and targets for disruption as this new technology reshapes the economic landscape. The degree to which this new technology will bring pain or plenty will depend on how firms and regulators can align in realizing the value that is promised by the financial market reaction to its release. The early labor market effects that we document highlight that there may be real challenges for policymakers in mitigating the negative impacts on employment and wages of the more affected workers.

REFERENCES

Abis, Simona, and Laura Veldkamp, 2023, The changing economics of knowledge production, *Review of Financial Studies* .

Acemoglu, Daron, and David Autor, 2011, Skills, tasks and technologies: Implications for employment and earnings, in *Handbook of labor economics*, volume 4, 1043–1171 (Elsevier).

Acemoglu, Daron, David Autor, Jonathon Hazell, and Pascual Restrepo, 2022, Artificial intelligence and jobs: evidence from online vacancies, *Journal of Labor Economics* 40, S293–S340.

Acemoglu, Daron, and Pascual Restrepo, 2018, The race between man and machine: Implications of technology for growth, factor shares, and employment, *American economic review* 108, 1488–1542.

Acemoglu, Daron, and Pascual Restrepo, 2020, Robots and jobs: Evidence from us labor markets, *Journal of political economy* 128, 2188–2244.

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

Agrawal, Ajay K, Joshua S Gans, and Avi Goldfarb, 2023, The turing transformation: Artificial intelligence, intelligence augmentation, and skill premiums, Technical report, National Bureau of Economic Research.

Akcigit, Ufuk, and Sina T Ates, 2021, Ten facts on declining business dynamism and lessons from endogenous growth theory, *American Economic Journal: Macroeconomics* 13, 257–298.

Autor, David H, Lawrence F Katz, and Melissa S Kearney, 2006, The polarization of the us labor market, *American economic review* 96, 189–194.

Autor, David H, Frank Levy, and Richard J Murnane, 2003, The skill content of recent technological change: An empirical exploration, *The Quarterly journal of economics* 118, 1279–1333.

Babina, Tania, Anastassia Fedyk, Alex He, and James Hodson, 2024, Artificial intelligence, firm growth, and product innovation, *Journal of Financial Economics* 151, 103745.

Babina, Tania, Anastassia Fedyk, Alex Xi He, and James Hodson, 2022, Firm investments in artificial intelligence technologies and changes in workforce composition, *Available at SSRN* .

Barattieri, Alessandro, Susanto Basu, and Peter Gottschalk, 2014, Some evidence on the importance of sticky wages, *American Economic Journal: Macroeconomics* 6, 70–101.

Begenau, Juliane, Maryam Farboodi, and Laura Veldkamp, 2018, Big data in finance and the growth of large firms, *Journal of Monetary Economics* 97, 71–87.

Berger, David, Kyle Herkenhoff, and Simon Mongey, 2022, Labor market power, *American Economic Review* 112, 1147–1193.

Bloom, Nicholas, Tarek Alexander Hassan, Aakash Kalyani, Josh Lerner, and Ahmed Tahoun, 2021, The diffusion of disruptive technologies, Technical report, National Bureau of Economic Research.

Caserta, Joe, Holger Harreis, Kayvaun Rowshankish, Nikhil Srinidhi, and Asin Tavakoli, 2023, The data dividend: Fueling generative ai .

Chen, AJ, Miao Ben Zhang, and Zhao Zhang, 2023, Talent market competition and firm growth, *Available at SSRN* .

Cohn, Jonathan B, Zack Liu, and Malcolm I Wardlaw, 2022, Count (and count-like) data in finance, *Journal of Financial Economics* 146, 529–551.

Donangelo, Andres, 2014, Labor mobility: Implications for asset pricing, *The Journal of Finance* 69, 1321–1346.

Eeckhout, Jan, and Laura Veldkamp, 2022, Data and market power, Technical report, National Bureau of Economic Research.

Eisfeldt, Andrea L., Antonio Falato, and Mindy Xiaolan, 2022, Human capitalists, in Martin Eichenbaum, Erik Hurst, and Valery Ramey, eds., *NBER Macroeconomics Annual*, chapter 1, 2 (University of Chicago Press, Chicago).

Eisfeldt, Andrea L, Antonio Falato, and Mindy Z Xiaolan, 2023, Human capitalists, *NBER Macroeconomics Annual* 37, 1–61.

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

Eloundou, Tyna, Sam Manning, Pamela Mishkin, and Daniel Rock, 2023, Gpts are gpts: An early look at the labor market impact potential of large language models, *arXiv preprint arXiv:2303.10130* .

Fama, Eugene F, and Kenneth R French, 2015, A five-factor asset pricing model, *Journal of financial economics* 116, 1–22.

Farboodi, Maryam, Roxana Mihet, Thomas Philippon, and Laura Veldkamp, 2019, Big data and firm dynamics, in *AEA papers and proceedings*, volume 109, 38–42, American Economic Association 2014 Broadway, Suite 305, Nashville, TN 37203.

Farboodi, Maryam, and Laura Veldkamp, 2023, Data and markets, *Annual Review of Economics* 15, 23–40.

Frey, Carl Benedikt, and Michael A Osborne, 2017, The future of employment: How susceptible are jobs to computerisation?, *Technological forecasting and social change* 114, 254–280.

Gilardi, Fabrizio, Meysam Alizadeh, and Maël Kubli, 2023, Chatgpt outperforms crowd-workers for text-annotation tasks, *arXiv preprint arXiv:2303.15056* .

Greenwood, Jeremy, 1999, The Third Industrial Revolution: Technology, Productivity, and Income Inequality, *Economic Review* 35, 2–2.

Greenwood, Jeremy, Zvi Hercowitz, and Per Krusell, 1997, Long-run implications of investment-specific technological change, *The American economic review* 342–362.

Gutiérrez, Germán, and Thomas Philippon, 2017a, Declining competition and investment in the us, Technical report, National Bureau of Economic Research.

Gutiérrez, Germán, and Thomas Philippon, 2017b, Investmentless growth: An empirical investigation, *Brookings Papers on Economic Activity* 89.

Hansen, Stephen, Peter John Lambert, Nicholas Bloom, Steven J Davis, Raffaella Sadun, and Bledi Taska, 2023, Remote work across jobs, companies, and space, Technical report, National Bureau of Economic Research.

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

Kelly, Bryan, Dimitris Papanikolaou, Amit Seru, and Matt Taddy, 2021, Measuring technological innovation over the long run, *American Economic Review: Insights* 3, 303–320.

Kogan, Leonid, and Dimitris Papanikolaou, 2014, Growth opportunities, technology shocks, and asset prices, *The journal of finance* 69, 675–718.

Kogan, Leonid, Dimitris Papanikolaou, Lawrence Schmidt, and Bryan Seegmiller, 2019, Technology-skill complementarity and labor displacement: Evidence from linking two centuries of patents with occupations, *Available at SSRN 3585676* .

Kogan, Leonid, Dimitris Papanikolaou, Lawrence DW Schmidt, and Bryan Seegmiller, 2023, Technology and labor displacement: Evidence from linking patents with worker-level data .

Krusell, Per, Lee E Ohanian, José-Víctor Ríos-Rull, and Giovanni L Violante, 2000, Capital-skill complementarity and inequality: A macroeconomic analysis, *Econometrica* 68, 1029–1053.

Li, Qin, Ben Lourie, Alexander Nekrasov, and Terry Shevlin, 2022, Employee turnover and firm performance: Large-sample archival evidence, *Management Science* 68, 5667–5683.

Lopez-Lira, Alejandro, and Yuehua Tang, 2023, Can ChatGPT Forecast Stock Price Movements? Return Predictability and Large Language Models, SSRN Scholarly Paper 4412788, Available at SSRN 4412788, Rochester, NY.

Loughran, Tim, and Bill McDonald, 2016, Textual analysis in accounting and finance: A survey, *Journal of Accounting Research* 54, 1187–1230.

Manuelli, Rodolfo E., and Ananth Seshadri, 2014, Frictionless Technology Diffusion: The Case of Tractors, *American Economic Review* 104, 1368–1391.

Moretti, Enrico, 2010, Local labor markets, Technical report, National Bureau of Economic Research.

Novy-Marx, Robert, 2013, The other side of value: The gross profitability premium, *Journal of financial economics* 108, 1–28.

Papanikolaou, Dimitris, 2011, Investment shocks and asset prices, *Journal of Political Economy* 119, 639–685.

Santos Silva, João, and Silvana Tenreyro, 2015, Ppml: Stata module to perform poisson pseudo-maximum likelihood estimation .

Seegmiller, Bryan, 2021, Valuing labor market power: The role of productivity advantages, *Available at SSRN 4412667* .

Tuzel, Selale, and Miao Ben Zhang, 2017, Local risk, local factors, and asset prices, *The Journal of Finance* 72, 325–370.

Tuzel, Selale, and Miao Ben Zhang, 2021, Economic Stimulus at the Expense of Routine-Task Jobs, *Journal of Finance* 76, 3347–3399.

Underwood, Ted, 2023, Using gpt-4 to measure the passage of time in fiction, https://tedunderwood.com/2023/03/19/using-gpt-4-to-measure-the-passage-of-time-in-fiction/.

Veldkamp, Laura, 2023, Valuing data as an asset, *Review of Finance* .

Webb, Michael, 2019, The impact of artificial intelligence on the labor market, *Available at SSRN 3482150* .

Willison, Simon, 2023, A simple python wrapper for the chatgpt api, https://til.simonwillison.net/gpt3/chatgpt-api.

Yeh, Chen, Claudia Macaluso, and Brad Hershbein, 2022, Monopsony in the us labor market, *American Economic Review* 112, 2099–2138.

Zhang, Miao Ben, 2019, Labor-technology substitution: Implications for asset pricing, *The Journal of Finance* 74, 1793–1839.

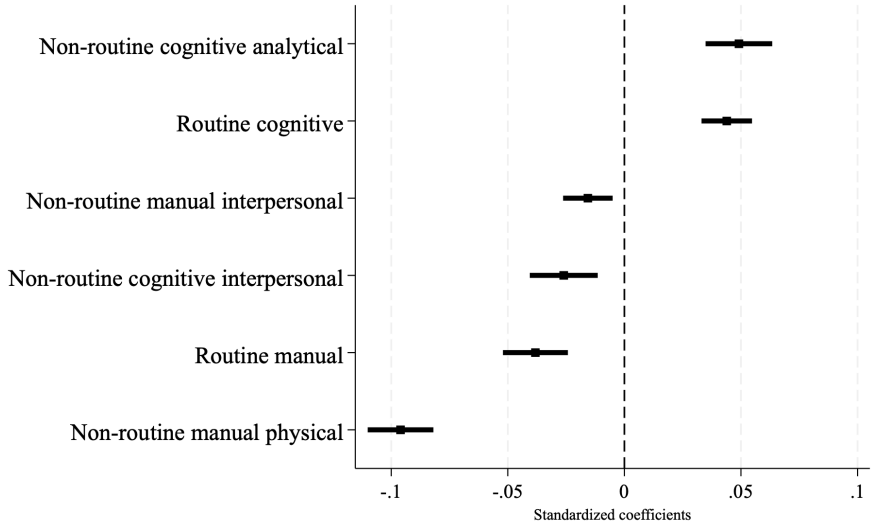


Figure 2. Occupational Generative AI Exposure and Skill Measures. This figure plots the coefficients of regressing occupations' Generative AI exposure (see equation (1)) on six measures of occupations' skills constructed by Acemoglu and Autor (2011). Equation (2) shows the regression specification. Occupation is classified at the SOC 6-digit level. The six skill measures are each standardized to have mean zero and the standard deviation of one. See Appendix C for details on the construction of these skill measures. The bar around each coefficient shows the 95% confidence interval based on heteroskedasticity-robust standard errors.

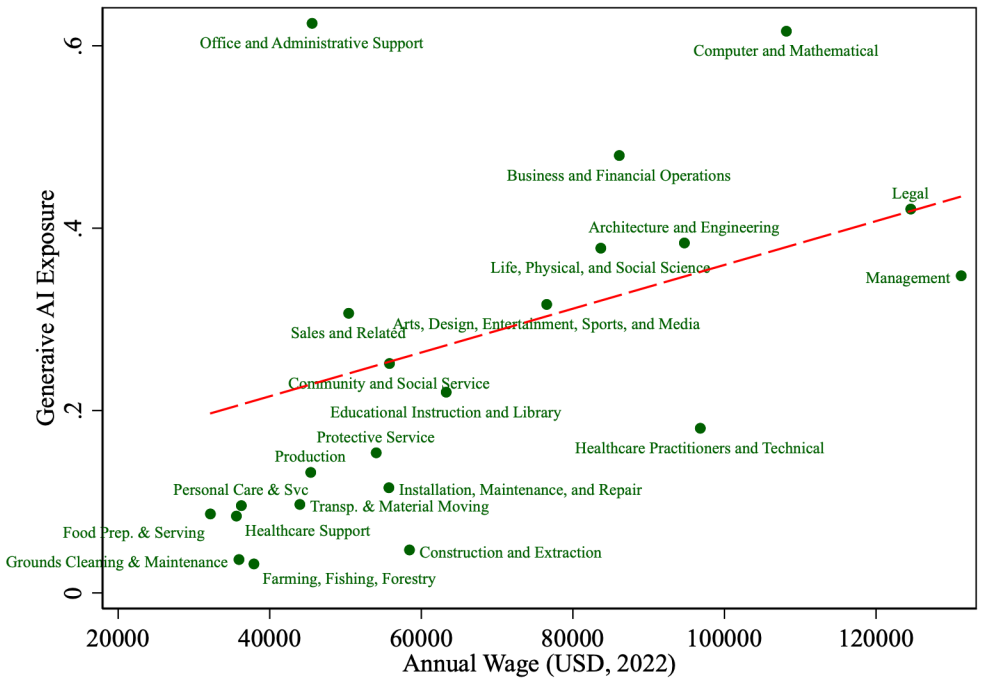


Figure 3. Generative AI Exposure and Wages by Major Occupation Group. The figure plots the relationship between Generative AI exposure and annual wages by SOC 2-digit occupation group. We aggregate the Generative AI exposure from SOC 6-digit to SOC 2-digit using the May 2022 occupational employment distribution from the Bureau of Labor Statistics Occupational Employment and Wage Statistics (OEWS) downloaded at https://www.bls.gov/oes/current/oes_stru.htm. We obtain the average annual wages of the SOC 2-digit occupations from the 2022 OEWS. The dashed red line indicates the employment-weighted linear best fit.

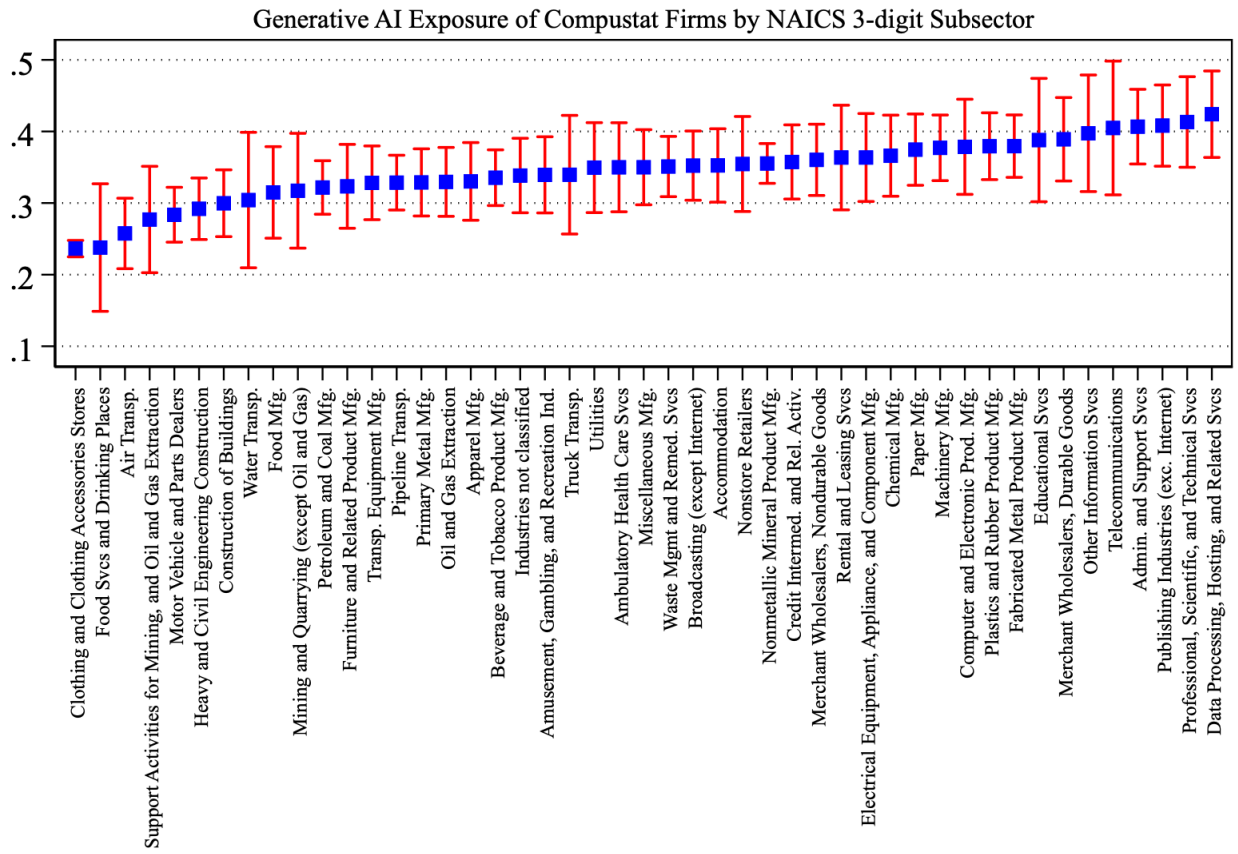


Figure 4. Generative AI Exposure Within and Across Industries. This figure plots the average and the standard deviation of Compustat firms' Generative AI exposure within each NAICS 3-digit industry in 2022. For ease of exposition, we require the industry to have at least 20 firms to be added to this figure. See Internet Appendix Figure IA.2 for the full set of subsectors.

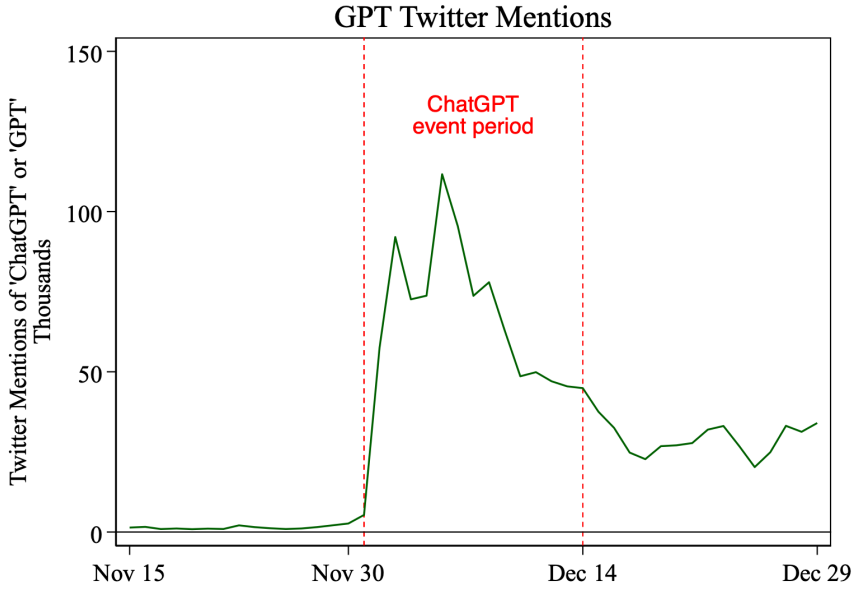
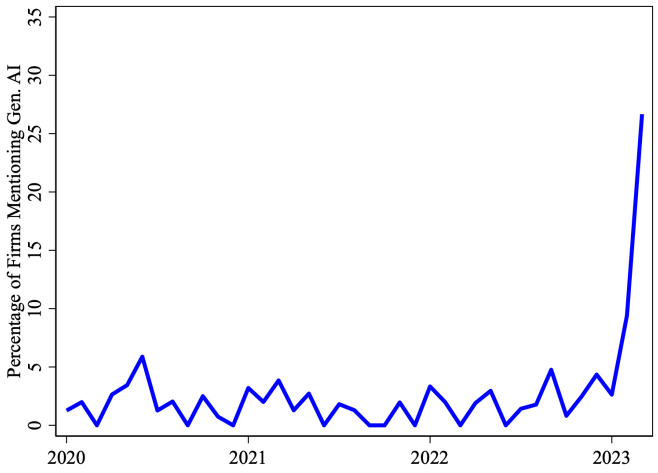


Figure 5. Social Media Attention to ChatGPT Release. The figure plots the total count of Twitter mentions of "ChatGPT" or "GPT", in thousands, on each day from November 14, 2022 to December 29, 2022. The data are from Media Cloud. The red dashed vertical lines indicate the "ChatGPT event period" for our stock market reaction analyses, which includes the two weeks following the release of ChatGPT on November 30, 2022.

Panel A: Share of Firms Mentioning Generative AI



Panel B: Firms' Generative AI Exposure and Their Mentions of Generative AI

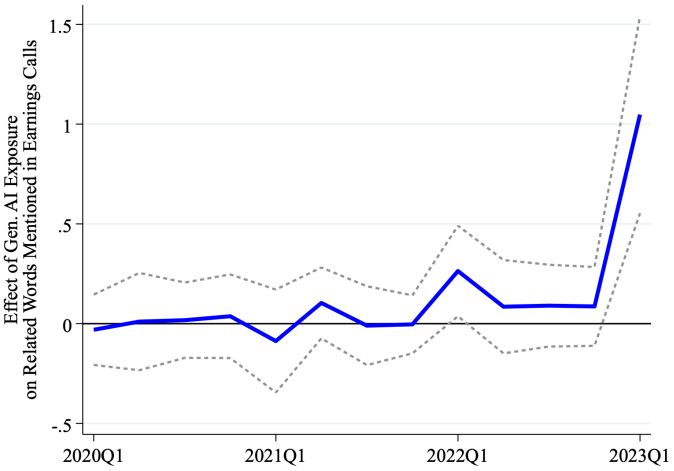


Figure 6. Firm Mentions of Generative AI in Earnings Conference Calls. Panel A plots the percentage of S&P 500 firms mentioning keywords of Generative AI in their quarterly earnings conference call transcripts in each quarter, where the data are manually collected from the *Seeking Alpha* website. We convert each call transcript into a list of lowercase unigrams and bigrams and examine whether it mentions the Generative AI keywords: "llm", "chatgpt", "gpt", "gpt3", "gpt4", "generative" and "language model". Panel B reports the cross-sectional relationship between firms' indicator of mentioning of Generative AI, $GenAI\ Mention_{i,t}$, and the labor-based Generative AI exposure, $GenAI\ Exp_i$, constructed in Section I. We estimate the following regression specification in each fiscal quarter from 2019Q1 to 2023Q1 and plot the point estimates of β_t and the 95% confidence intervals based on heteroskedasticity-robust standard errors: $\mathbb{1}[\text{GenAI Mention}]_{i,t} = \alpha_t + \beta_t \text{GenAI Exp}_i + \gamma \mathbb{1}[\text{GenAI Mention}]_{i,2019} + \varepsilon_{i,t}$. The Internet Appendix Figure IA.3 plots firms' mentions of generic "engineering" keywords as a placebo test.

Table I
Examples of GPT Scores for Tasks

This table presents examples of the scoring of labor tasks' exposure to Generative AI by comparing the tasks' statement to a rubric using the GPT 3.5 Turbo Model. We obtain from the O*NET V27.2 database task statements for each occupation and also an indicator of whether each task is "core" or "supplemental" to the occupation. An occupation has, on average, 22 tasks. We use the GPT model together with a rubric detailing the capabilities of ChatGPT-like tools to classify each task into three categories: if the task can be more efficiently completed with ChatGPT, it receives an exposure score of 1; if the task can be more efficiently completed with ChatGPT only if ChatGPT is further integrated with additional software or applications, it receives a score of 0.5 (Eloundou et al. (2023)); and if the task cannot be completed by ChatGPT, it receives a score of 0. The GPT model also provides an explanation for the classification of each task. Section I details the procedure.

Task	Occupation	Type	Score	GPT Explanation for the Score
Adjust sales scripts to better target the needs and interests of specific individuals.	Telemarketers	Core	1	The model can learn from the sales scripts and the needs and interests of specific individuals to adjust the sales scripts.
Write supporting code for Web applications or Web sites.	Web Developers	Core	1	The model can write code based on the specifications provided in the text input.
Develop and maintain media contact lists.	Fundraisers	Supp.	1	The model can learn how to develop and maintain media contact lists during training as part of its general knowledge base, and the task can be completed through text input and output.
Identify opportunities or develop channels for purchase or sale of securities or commodities.	Securities, Commodities, and Financial Services Sales Agents	Core	.5	The model can be used to analyze market trends and provide recommendations for securities or commodities to buy or sell. Additional software could be developed on top of the LLM to retrieve up-to-date information about the market and provide more specific recommendations.
Investigate facts and law of cases and search pertinent sources, such as public records and internet sources, to determine causes of action and to prepare cases.	Paralegals and Legal Assistants	Core	.5	The model can assist in searching pertinent sources and summarizing information, but it may not be able to retrieve up-to-date facts from the internet. Additional software could be developed on top of the LLM to retrieve up-to-date facts and combine them with the LLM capabilities.
Monitor budgets and payroll records, and review financial transactions to ensure that expenditures are authorized and budgeted.	Food Service Managers	Supp.	.5	The model can help with reviewing financial transactions and monitoring budgets by providing summaries and recommendations based on the input data. Additional software could be developed on top of the LLM to retrieve up-to-date financial information and provide more detailed analysis.
Connect heating or air conditioning equipment to fuel, water, or refrigerant source to form complete circuit.	Heating, AC, and Fridge Mechanics and Installers	Core	0	The task requires physical interaction with equipment and materials, which cannot be done by the LLM.
Perform therapeutic procedures, such as injections, immunizations, suturing and wound care, and infection management.	Physician Assistants	Core	0	The task requires physical interaction with patients, which cannot be done by the LLM.
Mentor new faculty.	Business Teachers, Postsecondary	Core	0	The task requires human interaction and decision-making that cannot be automated by the LLM.

Table II
Summary Statistics

This table reports summary statistics of key variations at the occupation level in Panel A and at the firm level in Panel B. *Generative AI Exposure*$_o$ is the SOC 6-digit occupation's Generative AI exposure aggregated from its tasks' Generative AI exposure (see equation (1)). *ShareSupp*$_o$ is the share of the occupation's Generative AI exposure that derives from the occupation's supplemental tasks' exposure to Generative AI as compared to its core tasks' exposure (see equation (6)). For firm-level measures, *Generative AI Exposure* is the firm's Generative AI exposure which is the average of its occupations' Generative AI exposure weighted by the firm's 2022 occupational employment shares from the Revelio Labs database. *Log Size* is the natural logarithm of total assets. *Tobin's Q* is the market value of the firm divided by book assets following Gutiérrez and Philippon (2017b). *ROA* is return on asset measured as EBITDA divided by total assets. *Labor Intensity* is the logarithm of the ratio of employment to the net value of property plant & equipment following Donangelo (2014). *Tangibility* is asset tangibility measured as the ratio of PP&E to total assets. *Mkt. Leverage* is the ratio of the firm's total value of debt to the sum of total debt and the market value of equity.

	Mean	Std. Dev.	p10	p50	p90	Obs.
Panel A: Occupation-level measures						
Generative AI Exposure$_o$	0.234	0.209	0.000	0.180	0.526	678
ShareSupp$_o$	0.227	0.275	0.000	0.141	0.667	576
Panel B: Firm-level measures						
Generative AI Exposure	0.354	0.078	0.268	0.353	0.442	2,518
Log Size	1.876	2.381	-1.091	2.063	4.636	2,517
Tobin's Q	3.667	10.081	1.354	2.176	5.832	2,380
ROA	-0.011	0.891	-0.181	0.093	0.216	2,513
Labor Intensity	0.761	1.761	-2.157	1.126	2.512	2,387
Tangibility	0.301	0.266	0.036	0.198	0.758	2,515
Market Leverage	0.239	0.215	0.011	0.179	0.553	2,379

Table III
Firms with Highest and Lowest Generative AI Exposure

This table lists the 15 firms with the highest Generative AI exposure in Panel A and the 15 firms with the lowest exposure in Panel B among the 100 largest US-headquartered publicly traded firms based on their market capitalization as of November 29, 2022. *GenAI Exp* is the firm's labor-based Generative AI exposure defined in Section I. *MktCap* is the firm's market capitalization in $ billions. *Industry* is classified at the NAICS 3-digit level.

Panel A: Top 15 large firms with the highest Generative AI exposure

Firm Name	GenAI Exp	MktCap	Industry
International Business Machines Corp	.488	128	Other Information Svcs
Intuit Inc.	.48	110	Publishing Industries (exc. Internet)
QUALCOMM Inc.	.479	123	Computer & Electronic Prod. Mfg.
Fiserv Inc.	.475	64	Data Processing, Hosting, & Related Svcs
NVIDIA Corporation	.468	360	Computer & Electronic Prod. Mfg.
S&P Global Inc	.452	108	Admin. & Support Svcs
Broadcom Inc	.449	234	Computer & Electronic Prod. Mfg.
Verizon Communications Inc	.444	165	Telecommunications
Microsoft Corp	.442	1790	Publishing Industries (exc. Internet)
3M Co	.442	66	Paper Mfg.
Advanced Micro Devices Inc	.441	104	Computer & Electronic Prod. Mfg.
ServiceNow Inc	.434	79	Publishing Industries (exc. Internet)
Adobe Inc	.427	155	Data Processing, Hosting, & Related Svcs
PayPal Holdings Inc	.418	81	Data Processing, Hosting, & Related Svcs
Thermo Fisher Scientific Inc	.411	215	Computer & Electronic Prod. Mfg.

Panel B: Bottom 15 large firms with the lowest Generative AI exposure

Firm Name	GenAI Exp	MktCap	Industry
Starbucks Corp	.119	114	Food Svcs & Drinking Places
McDonald's Corp	.194	193	Food Svcs & Drinking Places
Target Corp	.235	69	General Merchandise Stores
Walmart Inc	.235	382	General Merchandise Stores
Lowe's Cos Inc	.238	120	Build. Mat., Gard. Equip., Supplies Dealers
TJX Companies Inc (The)	.243	92	Clothing & Clothing Accessories Stores
Costco Wholesale Corp	.252	202	General Merchandise Stores
Union Pacific Corp	.253	127	Rail Transp.
CSX Corp	.256	64	Rail Transp.
United Parcel Service Inc	.256	149	Couriers & Messengers
Home Depot Inc. (The)	.261	321	Build. Mat., Gard. Equip., Supplies Dealers
Norfolk Southern Corp	.272	56	Rail Transp.
Tesla Inc	.283	390	Transp. Equipment Mfg.
Northrop Grumman Corp	.291	84	Computer & Electronic Prod. Mfg.
Mondelez International Inc	.292	91	Food Mfg.

Table IV
Reaction of Portfolios Sorted by Generative AI Exposure

This table reports daily percentage stock returns, as a measure of changes in firm value, during the ChatGPT event period for five value-weighted portfolios sorted on firms' Generative AI exposure. To form the portfolios, we sort NYSE stocks on November 29, 2022 into five quintiles by their Generative AI exposure, and we assign all stocks into the portfolios using the NYSE breakpoints. We aggregate stock returns within each portfolio using the stocks' market capitalization on the prior trading day as the weight. *AMH* is the "Artificial Minus Human" zero net investment portfolio long the portfolio with high exposure and short the portfolio with low exposure. We define the ChatGPT event period as from November 30, 2022 to December 14, 2022, i.e., the two weeks after the release of ChatGPT (see Section II for more details). Panel A reports each portfolio's raw daily return in excess of the risk-free rate. Panels B and C report the CAPM market-adjusted alphas and the Fama-French 5-factor-justed alphas, where the factor loadings of each portfolio are estimated using data from the previous six months. Panel D reports the Fama-French 5-factor-adjusted alphas for a sample that excludes firms in the tech sectors, NAICS 51 (information) and NAICS 54 (business services), following Acemoglu et al. (2022). *t*-statistics are shown in parentheses. See Section II.B for more details.

Q1 (H)	Q2	Q3	Q4	Q5 (A)	AMH
A: Excess returns (%)					
-0.169	0.020	0.141	0.035	0.281	0.450
(-0.62)	(0.06)	(0.41)	(0.11)	(0.74)	(3.53)
B: Market-factor-adjusted alpha (%)					
-0.234	-0.047	0.074	-0.028	0.208	0.442
(-3.91)	(-0.45)	(0.97)	(-0.93)	(4.54)	(4.70)
C: FF 5-factor-adjusted alpha (%)					
-0.188	0.125	-0.017	-0.026	0.166	0.354
(-2.86)	(2.31)	(-0.35)	(-0.82)	(4.23)	(3.85)
D: FF 5-factor-adjusted alpha ex. tech sector (%)					
-0.193	0.131	0.008	0.003	0.192	0.384
(-2.83)	(2.26)	(0.22)	(0.09)	(2.44)	(4.77)

Table V
Cumulative Abnormal Returns and Generative AI Exposure

This table reports the relation between firms' cumulative abnormal returns during the ChatGPT event period, $[-1, 10]$ trading days around the release of ChatGPT on November 30, 2022, (CAR $[-1, 10]$) and firms' Generative AI exposure, defined in Section I. To compute the cumulative abnormal return of each stock, we first subtract the daily market return from Kenneth French's website from the daily stock return from Yahoo Finance, and then we cumulate the daily abnormal returns over $[-1, 10]$. See Table II for the definitions of firm characteristic variables. Equation (4) describes the regression specification and the weighting of firms. t-statistics are computed using heteroskedasticity-robust standard errors and reported in parentheses. ***, **, and * denote significance at 1%, 5%, and 10% levels, respectively.

Dep. var.:	CAR[-1,10]							
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
GenAI Exp	0.216***	0.226***	0.210**	0.214***	0.212**	0.208**	0.199**	0.206**
	(2.621)	(2.672)	(2.170)	(2.747)	(2.472)	(2.238)	(2.374)	(2.192)
Log Size		-0.006***						-0.006**
		(-2.689)						(-2.517)
Tobin's Q			0.001					-0.002
			(0.291)					(-0.579)
ROA				0.081**				0.081
				(1.974)				(1.356)
Labor Intensity					0.002			-0.000
					(0.760)			(-0.114)
Tangibility						-0.005		-0.010
						(-0.226)		(-0.411)
Mkt. Leverage							-0.066**	-0.055*
							(-2.169)	(-1.901)
R^2	0.05	0.08	0.06	0.08	0.06	0.06	0.09	0.12
Observations	2,085	2,078	2,078	2,073	2,054	2,069	2,073	2,047

Table VI
Reaction of Within-Industry Sorted Portfolios

This table reports daily percentage stock returns, as a measure of changes in firm value, during the ChatGPT event period for three value-weighted portfolios sorted on firms' Generative AI exposure within the industry. To form the portfolios, we sort NYSE stocks on November 29, 2022 into three terciles by their Generative AI exposure across all stocks in Panels A1 and B1, within the NAICS 3-digit industry in Panels A2 and B2, and within the Hoberg and Phillips (2016) FIC 50-industry in Panels A3 and B3, and we assign all stocks into the portfolios using the NYSE breakpoints. We aggregate stock returns within each portfolio using the stocks' market capitalization on the prior trading day as the weight. *AMH* is the "Artificial Minus Human" zero net investment portfolio long the portfolio with high exposure and short the portfolio with low exposure. We define the ChatGPT event period as from November 30, 2022 to December 14, 2022, i.e., the two weeks after the release of ChatGPT (see Section II for more details). Panels A1-A3 report the CAPM market-adjusted alphas, and Panels B1-B3 report the Fama-French 5-factor-adjusted alphas. t-statistics are shown in parentheses. See Section II.C for more details.

Q1 (H)	Q2	Q3 (A)	AMH
A1: Market-factor-adjusted alpha (%)			
-0.174	0.069	0.124	0.298
(-2.87)	(1.36)	(2.86)	(3.70)
A2: NAICS 3-digit ind.-neutral market-factor-adjusted alpha (%)			
-0.165	0.029	0.145	0.311
(-3.98)	(0.50)	(9.43)	(6.06)
A3: FIC 50 ind.-neutral market-factor-adjusted alpha (%)			
-0.183	0.045	0.123	0.306
(-3.45)	(1.82)	(2.04)	(2.97)
B1: FF 5-factor-adjusted alpha (%)			
-0.059	0.012	0.105	0.164
(-1.12)	(0.32)	(2.41)	(2.03)
B2: NAICS 3-digit ind.-neutral FF 5-factor-adjusted alpha (%)			
-0.146	0.063	0.156	0.302
(-3.78)	(0.86)	(5.32)	(6.31)
B3: FIC 50 ind.-neutral FF 5-factor-adjusted alpha (%)			
-0.106	0.081	0.121	0.226
(-2.48)	(3.61)	(2.22)	(2.57)

Table VII
Product Exposure vs. Labor Exposure to Generative AI

This table reports the relation between firms' cumulative abnormal returns during the ChatGPT event period, $[-1, 10]$ trading days around the release of ChatGPT on November 30, 2022, (CAR $[-1, 10]$) and firms' labor-based Generative AI exposure (our main measure defined in Section I) while controlling for various proxies for firms' product exposure to Generative AI. To compute the cumulative abnormal return of each stock, we first subtract the daily market return from Kenneth French's website from the daily stock return from Yahoo Finance, and then we cumulate the daily abnormal returns over $[-1, 10]$. The definitions of the four proxies of firms' product exposure to Generative AI are described in detail in Section II.C.3. *GPT10 Product Exp* is based on GPT 3.5 Turbo assessment of firms' annual report's description of their products. *Count10K Product Exp* is based on counting AI-related keywords in firms' annual reports. *GS Product Exp* is based on Goldman Sachs' classification of firms as "near-term beneficiaries of AI". *BFHH Product Exp* is based on firms' share of AI-skilled workers constructed by Babina et al. (2024). All regressions include NAICS 3-digit fixed effects. Equation (5) describes the regression specification and the weighting of firms. t-statistics are computed using heteroskedasticity-robust standard errors and reported in parentheses. ***, **, and * denote significance at 1%, 5%, and 10% levels, respectively.

Dep. var.:	CAR[-1,10]					
	(1)	(2)	(3)	(4)	(5)	(6)
GenAI Exp	0.288***	0.291***	0.187**	0.264***	0.200**	0.175**
	(2.770)	(2.835)	(2.246)	(3.074)	(2.223)	(2.029)
GPT10K Product Exp		0.011				-0.016
		(0.614)				(-0.845)
Count10K Product Exp			0.002***			0.002
			(5.521)			(1.608)
GS Product Exp				0.041**		0.013
				(2.293)		(0.437)
BFHH Product Exp					1.251***	0.081
					(5.012)	(0.164)
R^2	0.36	0.37	0.40	0.38	0.38	0.40
Observations	2,080	1,905	1,905	2,080	1,493	1,374
NAICS3 FE	X	X	X	X	X	X

Table VIII
Generative AI Exposure and Firm Data Assets

This table reports coefficients from regressions of cumulative abnormal returns of firms during the ChatGPT event period, $[-1, 10]$ trading days around the release of ChatGPT on November 30, 2022, on firms' Generative AI exposure interacted with proxies for their data assets. To compute the cumulative abnormal return of each stock, we first subtract the daily market return from Kenneth French's website from the daily stock return from Yahoo Finance, and then we cumulate the daily abnormal returns over the $[-1, 10]$ period. Section II.D details the construction of the two proxies for firms' data assets. *10K Data Assets* is a score of 0, 1, 2, or 3, indicating the firm is not likely, slightly likely, moderate likely, or highly likely to have data that can be used as an input into large language model analytics based on the assessment of firms' annual reports using the GPT 3.5 Turbo Model. *AV Data Assets* is constructed following Abis and Veldkamp (2023) and measures the intensity of data management skills in the firm's workforce. All regressions include NAICS 3-digit fixed effects and weight firms by their market capitalization as of November 29, 2022. t-statistics are computed using heteroskedasticity-robust standard errors and reported in parentheses. ***, **, and * denote significance at 1%, 5%, and 10% levels, respectively.

Dep. var.:	CAR[-1,10]	
	(1)	(2)
GenAI Exp	-0.035	-0.085
	(-0.403)	(-0.992)
GenAI Exp × *10K Data Assets*	0.216**	
	(2.144)	
GenAI Exp × *AV Data Assets*		12.122**
		(2.339)
10K Data Assets	-0.073*	
	(-1.816)	
AV Data Assets		-3.815
		(-1.644)
R^2	0.11	0.12
Observations	1,910	2,043
NAICS3 FE	X	X

Table IX
Occupational Generative AI Exposure: Substitute vs. Complement

This table reports the impact of the ChatGPT release on job postings and wages of occupations. In Columns (1) and (2), the sample includes the total number of job postings at the occupation-month level in the U.S. from January 2019 to October 2023 aggregated from the Lightcast (formerly Burning Glass) job posting database. In Columns (3) and (4), the sample includes the total number of job postings at the more granular occupation-firm-month level for the U.S. publicly-traded firms. In Columns (5) and (6), the sample includes the worker-level hourly wage rate from the Census Current Population Survey (CPS) monthly panels from January 2019 to October 2023. Equations (7) and (8) describe the regression specifications. *Post* is a dummy variable indicating the months after the release of ChatGPT on November 30, 2022. $GenAI\ Exp_o$ is the occupation-level Generative AI exposure, defined in equation (1). $ShareSupp_o$ is the occupation's share of Generative AI exposure deriving from the occupation's supplemental (as opposed to core) tasks' exposure to Generative AI in equation (6). Worker demographic controls in Columns (5) and (6) include the worker's gender, age, age squared, work experience, experience squared, race (white vs. not), and years of education. Individual observations in Columns (5) and (6) are weighted by the CPS sampling weight. The job posting panels include many zeros, which makes it impractical to log-transform the dependent variable. Instead, we estimate all regressions using the Poisson model with fixed effects following Cohn et al. (2022) which means that coefficients can be interpreted as semi-elasticities, i.e., in terms of percentage changes in the dependent variable. t-statistics in parentheses are computed using standard errors double clustered at the month level and the occupation level in Columns (1)-(2) and (5)-(6) and triple clustered at the month, occupation, and firm levels in Columns (3)-(4). ***, **, and * denote significance at 1%, 5%, and 10% levels, respectively. See Section III.B for more details.

Dep var.:	Job Postings				Hourly Wage	
Panel Unit:	Occ. × Month		Occ. × Firm × Month		Worker× Month	
	(1)	(2)	(3)	(4)	(5)	(6)
Post × GenAI Exp	-0.382***	-0.636***	-0.362**	-0.701***	-0.033**	-0.065***
	(-4.035)	(-4.955)	(-2.166)	(-5.391)	(-2.287)	(-4.568)
Post × ShareSupp		-0.274***		-0.389**		0.003
		(-2.830)		(-2.218)		(0.262)
Post × GenAI Exp × ShareSupp		0.943***		1.495***		0.108**
		(3.410)		(3.194)		(2.293)
Observations	43,512	36,176	91,944,741	76,340,404	271,000	247,362
Occupation FE	X	X	X	X	X	X
Month FE	X	X			X	X
Firm× Month FE			X	X		
Metropolitan FE					X	X
Worker Demographic Control					X	X

Table X
Firms' Generative AI Exposure: Substitute vs. Complement

This table reports the reactions to the release of ChatGPT of firm fundamentals (Panel A) and firm values (Panel B). *ShareSupp* is the firm's share of Generative AI exposure deriving from its workers' supplemental (as opposed to core) tasks' exposure to Generative AI, defined in equation (9). *Post* is a dummy variable indicating the time period after the release of ChatGPT on November 30, 2022. *GenAI Exp* is the firm's Generative AI exposure, defined in equation (3). The dependent variable in Columns (1) and (2) of Panel A is the firm's number of job postings for occupations highly exposed to Generative AI, defined in Section III.C. We estimate these two regressions using the Poisson model with fixed effects following Cohn et al. (2022) which allows the coefficients to be interpreted as percentage changes in the dependent variable. The dependent variable in Columns (3) and (4) of Panel A is the firm's gross profitability defined as $(Revenue - COGS)/Assets$ following Novy-Marx (2013). All regressions from Columns (1) to (4) pf Panel A control for firm fixed effects and time fixed effects. Columns (1) and (2) further control for firms' total number of job postings in the month. Panel B reports coefficients from cross-sectional regressions of firms' cumulative abnormal returns during the ChatGPT event period (CAR$[-1, 10]$) on *GenAI Exp* as well as its interaction with *ShareSupp*. See Table V for details on the construction of CAR. t-statistics in parentheses are based on standard errors clustered at both the firm level and time level in Panel A and based on heteroskedasticity-robust standard errors in Panel B. ***, **, and * denote significance at 1%, 5%, and 10% levels, respectively. See Section III.C for more details.

Panel A: Effects on Firm Fundamentals

Dep var.:	High-Exp Occ Job Posting		Profitability	
Panel Unit:	Firm×Month		Firm×Qtr	
	(1)	(2)	(3)	(4)
Post × GenAI Exp	-1.378***	-3.038***	0.016**	0.056***
	(-3.747)	(-2.991)	(2.265)	(3.189)
Post × ShareSupp		-2.302**		0.059**
		(-2.029)		(2.007)
Post × GenAI Exp × ShareSupp		8.100**		-0.231***
		(2.037)		(-2.946)
Observations	36,900	36,880	14,621	14,614
Time FE	X	X	X	X
Firm FE	X	X	X	X

Panel B: Effects on Firm Values

Dep var.:	CAR[-1,10]	
	(1)	(2)
GenAI Exp	0.216***	0.380***
	(2.621)	(2.579)
ShareSupp		0.272
		(1.447)
GenAI Exp × ShareSupp		-1.937***
		(-3.226)
Observations	2,085	2,084

Internet Appendix for

"The Labor Impact of Generative AI on Firm Value"

Andrea L. Eisfeldt Gregor Schubert Bledi Taska Miao Ben Zhang

Appendix A. Appendix Figures

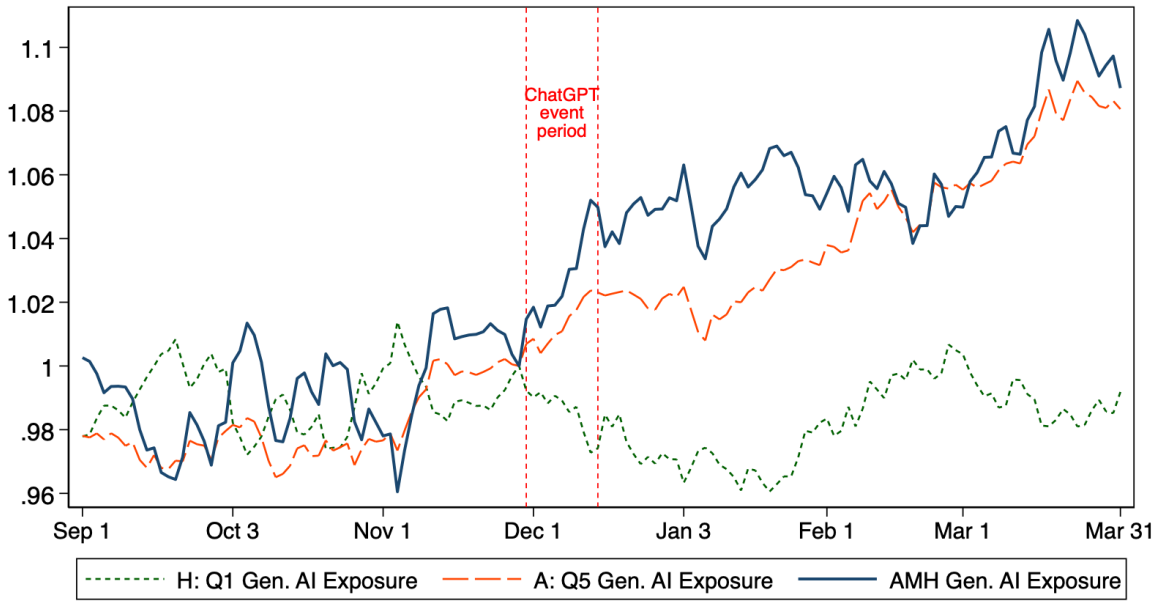


Figure IA.1. Cumulative Abnormal Returns by Generative AI Exposure: Extended Pre-Period The figure plots the cumulative abnormal returns (CARs) of value-weighted quintile portfolios sorted by firms' labor-based Generative AI exposure. The graph shows the CARs of the lowest-exposure quintile portfolio, "Human" (H), the highest-exposure quintile portfolio, "Artificial" (A), and the zero investment portfolio that goes long A and shorts H, "Artificial-minus-Human" (AMH). Market-adjusted daily abnormal returns are cumulated from November 29, 2022, the day before the release of ChatGPT, and are based on factor exposures computed over the 4-month period preceding the period shown in the graph. Daily stock returns are from Yahoo Finance. The dashed vertical lines indicate the "ChatGPT event period" from November 30, 2022, to December 14, 2022. See details of the definition of firms' Generative AI exposure in Section I and the construction of the portfolios, the calculation of portfolios' CARs, and the determination of the ChatGPT event period in Section II.

Figure IA.2. Generative AI exposure across and within subsectors: complete list. This figure plots the average and the standard deviation of Compustat firms' Generative AI exposure within each NAICS 3-digit subsector.

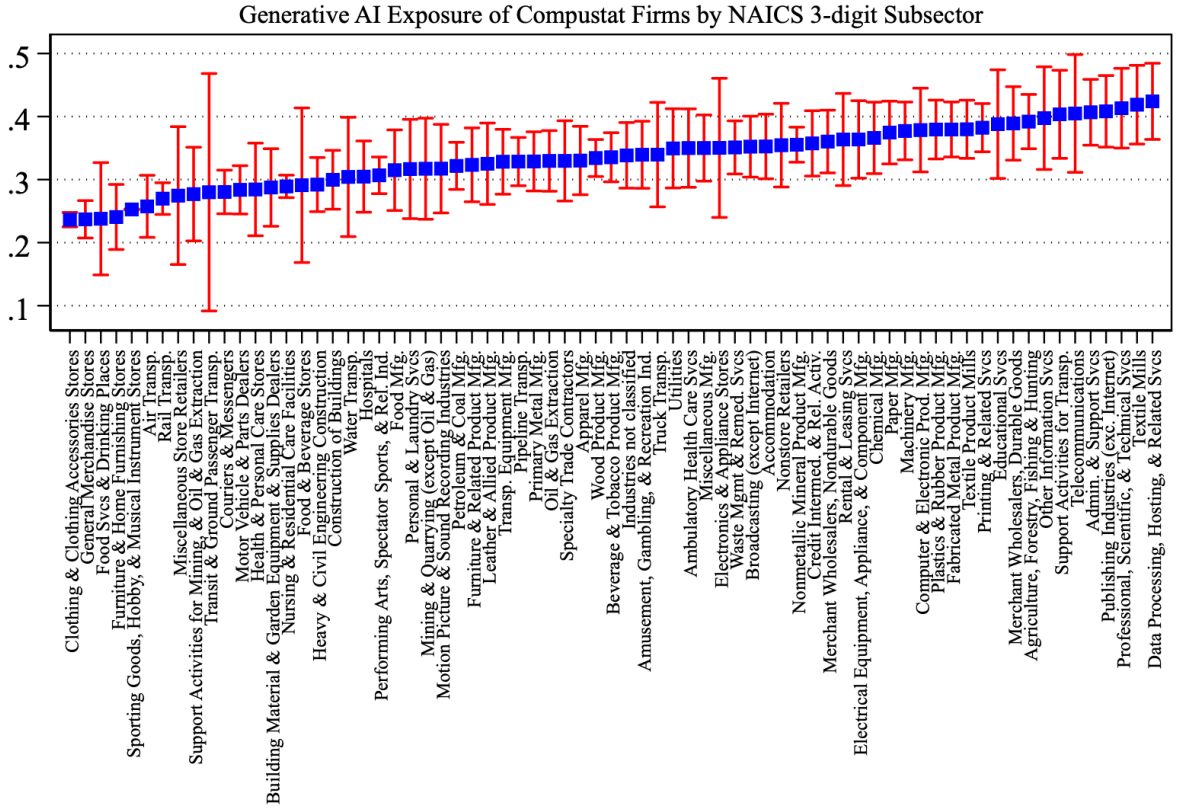
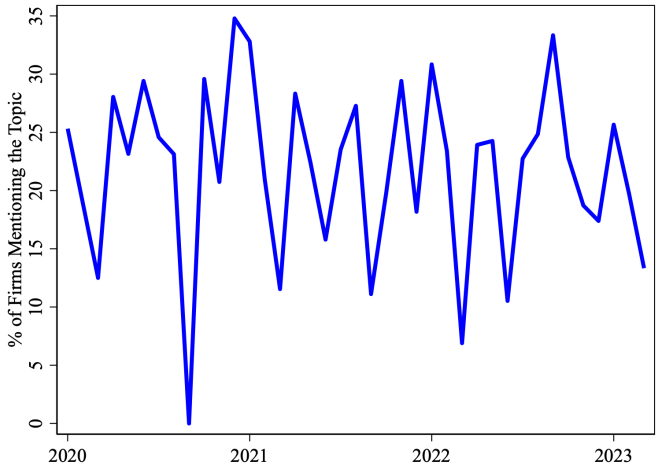


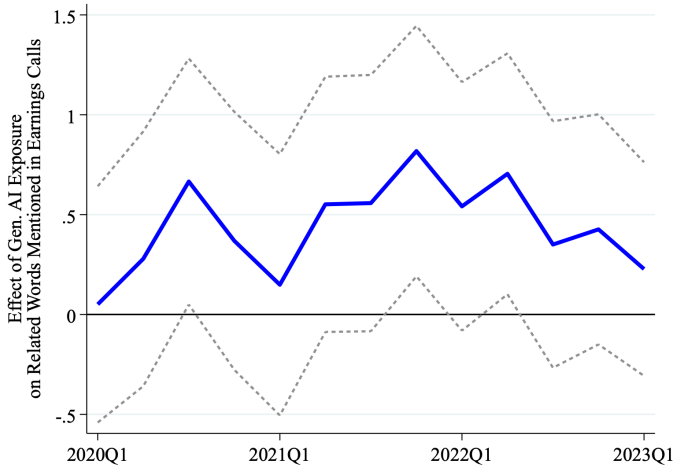
Figure IA.3. Firm-level Gen. AI exposure and "engineering" topic mentions in company earnings conference calls. These graphs use data on the share of S&P 500 firms' earnings calls that mention a particular topic. The quarterly earnings conference call transcripts for S&P 500 firms' are manually collected from the *Seeking Alpha* website. Each call transcript is converted into a list of lower-case unigrams and bigrams. Panel A shows the share of earnings calls mentioning keywords about engineering: "engineer" and "engineering". Panel B then shows the result of estimating regression specifications of the form

$$\mathbb{1}[\text{Engineering Topic}]_{i,t} = \alpha_t + \beta_t^X \text{GenAI Exp}_i + \gamma \mathbb{1}[\text{Engineering Topic}]_{i,2019} + \varepsilon_{i,t}$$

for cross-sections in each fiscal quarter from 2019 Q1 to 2023 Q1. The graph shows the estimates $\hat{\beta}_t^X$ that represent the effect of higher Generative AI exposure on the likelihood that a firm mentions "engineering" as a topic in that quarter's earnings call, controlling for whether the firm already mentioned the topic in any 2019 earnings call. Dotted lines show 95% confidence intervals based on heteroskedasticity-robust standard errors.



(A) Share of earnings calls mentioning Engineering



(B) Effect of Gen. AI exposure on firm mentioning Engineering

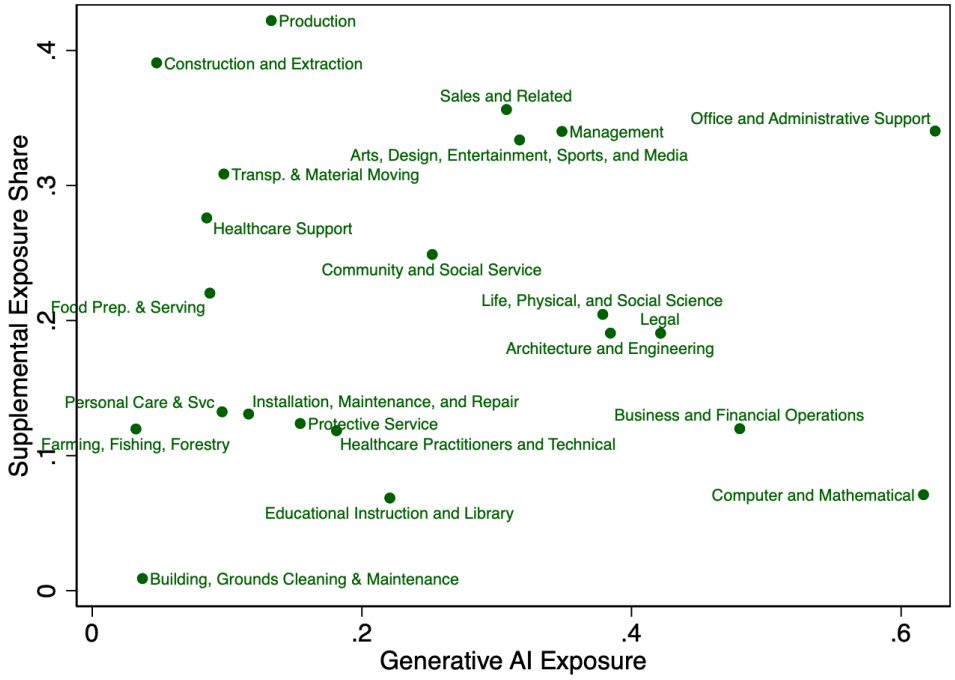


Figure IA.4. Generative AI Exposure and Supplemental Exposure Share by Major Occupation Group. The figure plots the relationship between Generative AI exposure and the supplemental task share of exposure by SOC 2-digit occupation group. We aggregate the Generative AI exposure and supplemental task exposure shares from SOC 6-digit to SOC 2-digit using the May 2022 occupational employment distribution from the Bureau of Labor Statistics Occupational Employment and Wage Statistics (OEWS) downloaded at https://www.bls.gov/oes/current/oes_stru.htm.

Appendix B. Appendix Tables

SOC Code	Occupation Title	Exposure Score
41-9041	Telemarketers	.96
43-9081	Proofreaders and copy markers	.95
43-3031	Bookkeeping, accounting, and auditing clerks	.87
15-2021	Mathematicians	.86
15-1251	Computer programmers	.85
43-9022	Word processors and typists	.85
43-3011	Bill and account collectors	.83
27-3091	Interpreters and translators	.82
43-9111	Statistical assistants	.82
15-1254	Web developers	.81
43-6011	Executive secretaries and executive administrative assistants	.77
43-3051	Payroll and timekeeping clerks	.77
43-6014	Secretaries and administrative assistants, except legal, medical, and executive	.77
43-5061	Production, planning, and expediting clerks	.76
15-1212	Information security analysts	.75
43-6013	Medical secretaries and administrative assistants	.75
27-3043	Writers and authors	.75
43-4021	Correspondence clerks	.74
43-9061	Office clerks, general	.74
41-3091	Sales representatives of services, except advertising, insurance, financial services, and travel	.73
⋮	⋮	⋮
39-5093	Shampooers	0
51-6041	Shoe and leather workers and repairers	0
51-6042	Shoe machine operators and tenders	0
51-3023	Slaughterers and meat packers	0
47-2022	Stonemasons	0
47-2221	Structural iron and steel workers	0
51-2041	Structural metal fabricators and fitters	0
29-9093	Surgical assistants	0
51-6052	Tailors, dressmakers, and custom sewers	0
47-2082	Tapers	0
49-9052	Telecommunications line installers and repairers	0
47-2053	Terrazzo workers and finishers	0
51-6064	Textile winding, twisting, and drawing out machine setters, operators, and tenders	0
47-2044	Tile and stone setters	0
51-9197	Tire builders	0
49-3093	Tire repairers and changers	0
51-4194	Tool grinders, filers, and sharpeners	0
39-3031	Ushers, lobby attendants, and ticket takers	0
49-9064	Watch and clock repairers	0
53-7073	Wellhead pumpers	0

Table IA.1 Highest and lowest Generative AI exposure score occupations. See Section I.B for details. Note that 17% (132 of 785) of the occupations have zero exposure, so only a subset of the lowest exposure occupations is shown.

NAICS Subsector	Industry Title	Exposure Score
518	Data Processing, Hosting, and Related Svcs	.42
313	Textile Mills	.42
541	Professional, Scientific, and Technical Svcs	.41
511	Publishing Industries (exc. Internet)	.41
561	Admin. and Support Svcs	.41
517	Telecommunications	.4
488	Support Activities for Transp.	.4
519	Other Information Svcs	.4
111	Agriculture, Forestry, Fishing and Hunting	.39
423	Merchant Wholesalers, Durable Goods	.39
611	Educational Svcs	.39
323	Printing and Related Svcs	.38
314	Textile Product Mills	.38
332	Fabricated Metal Product Mfg.	.38
326	Plastics and Rubber Product Mfg.	.38
334	Computer and Electronic Prod. Mfg.	.38
333	Machinery Mfg.	.38
322	Paper Mfg.	.37
325	Chemical Mfg.	.37
335	Electrical Equipment, Appliance, and Component Mfg.	.36
532	Rental and Leasing Svcs	.36
424	Merchant Wholesalers, Nondurable Goods	.36
522	Credit Intermed. and Rel. Activ.	.36
327	Nonmetallic Mineral Product Mfg.	.36
454	Nonstore Retailers	.35
721	Accommodation	.35
515	Broadcasting (except Internet)	.35
562	Waste Mgmt and Remed. Svcs	.35
443	Electronics and Appliance Stores	.35
339	Miscellaneous Mfg.	.35
621	Ambulatory Health Care Svcs	.35
221	Utilities	.35
484	Truck Transp.	.34
713	Amusement, Gambling, and Recreation Ind.	.34
999	Industries not classified	.34
312	Beverage and Tobacco Product Mfg.	.34
321	Wood Product Mfg.	.33
315	Apparel Mfg.	.33
238	Specialty Trade Contractors	.33
211	Oil and Gas Extraction	.33
331	Primary Metal Mfg.	.33
486	Pipeline Transp.	.33
336	Transp. Equipment Mfg.	.33
316	Leather and Allied Product Mfg.	.33
337	Furniture and Related Product Mfg.	.32
324	Petroleum and Coal Mfg.	.32
512	Motion Picture and Sound Recording Industries	.32
212	Mining and Quarrying (except Oil and Gas)	.32
812	Personal and Laundry Svcs	.32
311	Food Mfg.	.31
711	Performing Arts, Spectator Sports, and Rel. Ind.	.31
622	Hospitals	.3
483	Water Transp.	.3
236	Construction of Buildings	.3
237	Heavy and Civil Engineering Construction	.29
445	Food and Beverage Stores	.29
623	Nursing and Residential Care Facilities	.29
444	Building Material and Garden Equipment and Supplies Dealers	.29
446	Health and Personal Care Stores	.28
441	Motor Vehicle and Parts Dealers	.28
492	Couriers and Messengers	.28
485	Transit and Ground Passenger Transp.	.28
213	Support Activities for Mining, and Oil and Gas Extraction	.28
453	Miscellaneous Store Retailers	.27
482	Rail Transp.	.27
481	Air Transp.	.26
451	Sporting Goods, Hobby, and Musical Instrument Stores	.25
442	Furniture and Home Furnishing Stores	.24
722	Food Svcs and Drinking Places	.24
452	General Merchandise Stores	.24
448	Clothing and Clothing Accessories Stores	.24

Table IA.2 Generative AI exposure scores by NAICS 3-digit industry

Table IA.3 Generative AI exposure for the Largest 100 U.S. Firms This table lists the Generative AI exposure scores for the largest 100 publicly-traded firms with headquarters in the U.S., where size is measured as the market capitalization as of November 1, 2022. *Generative AI exposure* is the firm's labor exposure defined in Section I. *MktCap* is the firm's market capitalization as of November 1, 2022, in $B. *Subsector* is defined at the NAICS 3-digit level.

Company Name	Gen. AI exposure	MktCap	Subsector
International Business Machines Corp	.488	128	Other Information Svcs
Intuit Inc.	.48	110	Publishing Industries (exc. Internet)
QUALCOMM Inc.	.479	123	Computer & Electronic Prod. Mfg.
Fiserv Inc.	.475	64	Data Processing, Hosting, & Related Svcs
NVIDIA Corporation	.468	360	Computer & Electronic Prod. Mfg.
S&P Global Inc	.452	108	Admin. & Support Svcs
Broadcom Inc	.449	234	Computer & Electronic Prod. Mfg.
Verizon Communications Inc	.444	165	Telecommunications
Microsoft Corp	.442	1790	Publishing Industries (exc. Internet)
3M Co	.442	66	Paper Mfg.
Advanced Micro Devices Inc	.441	104	Computer & Electronic Prod. Mfg.
ServiceNow Inc	.434	79	Publishing Industries (exc. Internet)
Adobe Inc	.427	155	Data Processing, Hosting, & Related Svcs
PayPal Holdings Inc	.418	81	Data Processing, Hosting, & Related Svcs
Thermo Fisher Scientific Inc	.411	215	Computer & Electronic Prod. Mfg.
Intuitive Surgical Inc	.404	93	Computer & Electronic Prod. Mfg.
Automatic Data Processing Inc	.398	99	Data Processing, Hosting, & Related Svcs
Comcast Corp	.396	148	Broadcasting (except Internet)
Vertex Pharmaceuticals Inc	.395	74	Chemical Mfg.
Analog Devices Inc	.392	84	Computer & Electronic Prod. Mfg.
AbbVie Inc	.391	286	Chemical Mfg.
Regeneron Pharmaceuticals Inc	.39	79	Chemical Mfg.
Gilead Sciences Inc	.388	107	Chemical Mfg.
Micron Technology Inc.	.388	55	Computer & Electronic Prod. Mfg.
Intel Corp	.386	109	Computer & Electronic Prod. Mfg.
Bristol-Myers Squibb Co	.385	151	Chemical Mfg.
Illinois Tool Works Inc.	.382	67	Machinery Mfg.
Netflix Inc	.381	131	Rental & Leasing Svcs
Meta Platforms Inc	.381	315	Other Information Svcs
Lam Research Corp	.38	58	Machinery Mfg.
SALESFORCE INC	.379	130	Publishing Industries (exc. Internet)
General Dynamics Corp	.378	68	Transp. Equipment Mfg.
Abbott Laboratories	.376	191	Computer & Electronic Prod. Mfg.
AT&T Inc	.375	131	Telecommunications
Applied Materials Inc	.374	82	Machinery Mfg.
Booking Holdings Inc	.373	76	Other Information Svcs
General Electric Co	.373	91	Industries not classified
Merck & Co Inc	.372	282	Chemical Mfg.
T-Mobile US Inc	.371	173	Telecommunications
Johnson & Johnson	.371	462	Chemical Mfg.
Honeywell International Inc	.368	143	Transp. Equipment Mfg.
Alphabet Inc	.366	1134	Data Processing, Hosting, & Related Svcs
Amgen Inc	.365	140	Chemical Mfg.
Eli Lilly and Co	.364	329	Chemical Mfg.
Apple Inc	.364	2072	Computer & Electronic Prod. Mfg.
Philip Morris International Inc	.364	157	Beverage & Tobacco Product Mfg.
DEERE & COMPANY	.364	128	Machinery Mfg.
Texas Instruments Inc	.363	150	Computer & Electronic Prod. Mfg.
Emerson Electric Co.	.363	57	Computer & Electronic Prod. Mfg.
Caterpillar Inc	.358	124	Machinery Mfg.
CVS Health Corp	.356	121	Ambulatory Health Care Svcs
Cisco Systems Inc	.355	196	Computer & Electronic Prod. Mfg.
Zoetis Inc	.355	68	Chemical Mfg.
Pfizer Inc	.352	288	Chemical Mfg.
Southern Co (The)	.351	78	Utilities
Danaher Corp	.35	193	Computer & Electronic Prod. Mfg.
Procter & Gamble Co (The)	.342	363	Chemical Mfg.
Raytheon Technologies Corp	.339	148	Transp. Equipment Mfg.
Colgate-Palmolive Co	.337	65	Chemical Mfg.
Becton Dickinson and Co	.331	72	Miscellaneous Mfg.
NextEra Energy Inc	.329	166	Utilities
Walt Disney Co (The)	.328	158	Broadcasting (except Internet)
Altria Group Inc	.327	82	Beverage & Tobacco Product Mfg.
Air Products and Chemicals Inc.	.326	68	Chemical Mfg.
Waste Management Inc.	.325	64	Waste Mgmt & Remed. Svcs
Duke Energy Corp	.322	79	Utilities
EOG Resources Inc.	.322	76	Oil & Gas Extraction
Exxon Mobil Corp	.32	450	Petroleum & Coal Mfg.
Estee Lauder Cos Inc (The)	.32	89	Chemical Mfg.
Amazon.com Inc	.317	860	Nonstore Retailers
Stryker Corp	.317	93	Miscellaneous Mfg.
Schlumberger Ltd	.316	76	Support Activities for Mining, & Oil & Gas Extraction
Conocophillips	.316	144	Oil & Gas Extraction
HCA Healthcare Inc	.312	67	Hospitals
Occidental Petroleum Corp	.307	57	Oil & Gas Extraction
Coca-Cola Co (The)	.306	275	Beverage & Tobacco Product Mfg.
Boston Scientific Corp	.305	66	Computer & Electronic Prod. Mfg.
Freeport-McMoRan Inc	.305	54	Mining & Quarrying (except Oil & Gas)
PepsiCo Inc	.303	249	Beverage & Tobacco Product Mfg.
Chevron Corp	.301	344	Petroleum & Coal Mfg.
Berkshire Hathaway Inc	.3	676	Industries not classified
Lockheed Martin Corp	.299	124	Transp. Equipment Mfg.
Boeing Co	.298	114	Transp. Equipment Mfg.
Sherwin-Williams Co (The)	.296	61	Chemical Mfg.
Activision Blizzard Inc	.295	60	Publishing Industries (exc. Internet)
Mondelez International Inc	.292	91	Food Mfg.
Northrop Grumman Corp	.291	84	Computer & Electronic Prod. Mfg.
Tesla Inc	.283	390	Transp. Equipment Mfg.
Norfolk Southern Corp	.272	56	Rail Transp.
Home Depot Inc. (The)	.261	321	Build. Mat., Gard. Equip., Supplies Dealers
United Parcel Service Inc	.256	149	Couriers & Messengers
CSX Corp	.256	64	Rail Transp.
Union Pacific Corp	.253	127	Rail Transp.
Costco Wholesale Corp	.252	202	General Merchandise Stores
TJX Companies Inc (The)	.243	92	Clothing & Clothing Accessories Stores
Lowe's Cos Inc	.238	120	Build. Mat., Gard. Equip., Supplies Dealers
Walmart Inc	.235	382	General Merchandise Stores
Target Corp	.235	69	General Merchandise Stores
McDonald's Corp	.194	193	Food Svcs & Drinking Places
Starbucks Corp	.119	114	Food Svcs & Drinking Places

Table IA.4 Firm Generative AI exposure and firm characteristics. This table regresses our firms' Generative AI exposure measure on firm characteristics using the cross-section of U.S. publicly traded firms in 2022. See Table II for variable definitions. Panel B controls for fixed effects at the NAICS 3-digit level. All variables are winsorized at the top and bottom 1%. Standard errors are clustered at the industry level and reported in parentheses. *, **, and *** indicate significance at the 10%, 5%, and 1% level, respectively.

Panel A: Across All Firms						
	(1)	(2)	(3)	(4)	(5)	(6)
Log Size	-2.653**					
	(1.199)					
Tobin's Q		3.076***				
		(0.798)				
ROA			-21.516*			
			(11.100)			
Labor Intensity				7.892***		
				(2.325)		
Tangibility					-89.931***	
					(20.961)	
Mkt. Leverage						-61.872***
						(12.899)
Observations	2517	2380	2513	2387	2515	2379
Adjusted R^2	0.006	0.013	0.006	0.038	0.107	0.032
Panel B: Within-Industry						
	(1)	(2)	(3)	(4)	(5)	(6)
Log Size	-0.730					
	(1.021)					
Tobin's Q		0.524				
		(0.661)				
ROA			-7.303			
			(10.112)			
Labor Intensity				5.593***		
				(1.559)		
Tangibility					-61.794**	
					(25.082)	
Mkt. Leverage						-22.958***
						(7.092)
Observations	2517	2380	2513	2387	2515	2379
Adjusted R^2	0.280	0.292	0.279	0.313	0.298	0.296

Table IA.5
Reaction of Portfolios Sorted by Generative AI Exposure: Shorter Event Period

This table reports daily percentage stock returns, as a measure of changes in firm value, during the ChatGPT event period for five value-weighted portfolios sorted on firms' Generative AI exposure. To form the portfolios, we sort NYSE stocks on November 29, 2022 into five quintiles by their Generative AI exposure, and we assign all stocks into the portfolios using the NYSE breakpoints. We aggregate stock returns within each portfolio using the stocks' market capitalization on the prior trading day as the weight. *AMH* is the "Artificial Minus Human" zero net investment portfolio long the portfolio with high exposure and short the portfolio with low exposure. We define the shorter ChatGPT event period here as from November 30, 2022 to December 7, 2022, i.e., the 1 week after the release of ChatGPT (see Section II for more details). Panel A reports each portfolios' raw daily return in excess of the risk free rate. Panels B and C report the CAPM market-adjusted alphas and the Fama-French 5-factor-justed alphas, where the factor loadings of each portfolio are estimated using data from the previous six months. Panel D reports the Fama-French 5-factor-adjusted alphas for a sample that excludes firms in the tech sectors, NAICS 51 (information) and NAICS 54 (business services), following Acemoglu et al. (2022). *t*-statistics are shown in parentheses. See Section II.B for more details.

Q1 (H)	Q2	Q3	Q4	Q5 (A)	AMH
A: Excess returns (%)					
-0.246	-0.253	-0.077	-0.087	0.002	0.248
(-0.49)	(-0.52)	(-0.15)	(-0.17)	(0.00)	(1.89)
B: Market-factor-adjusted alpha (%)					
-0.117	-0.120	0.056	0.038	0.148	0.265
(-2.40)	(-1.93)	(0.69)	(1.12)	(2.83)	(2.96)
C: FF 5-factor-adjusted alpha (%)					
-0.080	0.073	-0.042	0.013	0.138	0.218
(-1.87)	(0.72)	(-0.83)	(0.31)	(3.02)	(2.76)
D: FF 5-factor-adjusted alpha ex. tech sector (%)					
-0.085	0.082	-0.035	0.057	0.191	0.276
(-1.86)	(0.79)	(-0.79)	(1.54)	(2.83)	(3.59)

Table IA.6Reaction of Portfolios sorted by Generative AI exposure: Non-event period.

This table reports daily percentage stock returns, as a measure of changes in firm value, in the days surrounding, but outside of, the ChatGPT event period for five value-weighted portfolios sorted on firms' Generative AI exposure. To form the portfolios, we sort NYSE stocks on November 29, 2022 into five quintiles by their Generative AI exposure, and we assign all stocks into the portfolios using the NYSE breakpoints. We aggregate stock returns within each portfolio using the stocks' market capitalization on the prior trading day as the weight. *AMH* is the "Artificial Minus Human" zero net investment portfolio long the portfolio with high exposure and short the portfolio with low exposure. We define the ChatGPT event period as from November 30, 2022 to December 14, 2022, i.e., the two weeks after the release of ChatGPT (see Section II for more details). The table then shows placebo event study returns for the ten trading days before (Nov. 14- Nov. 29, 2022) and after (Dec 15 - Dec 29, 2022) the ChatGPT release event period. Panel A reports each portfolios' raw daily return in excess of the risk free rate. Panels B and C report the CAPM market-adjusted alphas and the Fama-French 5-factor-justed alphas, where the factor loadings of each portfolio are estimated using data from the previous six months. Panel D reports the Fama-French 5-factor-adjusted alphas for a sample that excludes firms in the tech sectors, NAICS 51 (information) and NAICS 54 (business services), following Acemoglu et al. (2022). t-statistics are shown in parentheses. See Section II.B for more details.

Q1 (H)	Q2	Q3	Q4	Q5 (A)	AMH
A: Excess returns (%)					
-0.140	-0.177	-0.290	-0.110	-0.232	-0.092
(-0.67)	(-0.92)	(-1.09)	(-0.56)	(-0.93)	(-0.88)
B: Market-factor-adjusted alpha (%)					
0.062	0.032	-0.081	0.087	-0.004	-0.066
(0.65)	(0.50)	(-1.08)	(2.09)	(-0.14)	(-0.72)
C: FF 5-factor-adjusted alpha (%)					
0.069	0.049	-0.129	0.071	0.019	-0.050
(0.73)	(1.12)	(-3.12)	(2.06)	(0.71)	(-0.51)
D: FF 5-factor-adjusted alpha ex. tech sector (%)					
0.065	0.047	-0.151	0.041	0.023	-0.042
(0.68)	(0.94)	(-3.78)	(1.24)	(0.60)	(-0.50)

Table IA.7 Reaction of Within-industry Portfolios sorted by Generative AI exposure: Non-event period.

This table reports daily percentage stock returns, as a measure of changes in firm value, in the days surrounding, but outside of, the ChatGPT event period for three value-weighted portfolios sorted on firms' Generative AI exposure within the industry. To form the portfolios, we sort NYSE stocks on November 29, 2022 into three terciles by their Generative AI exposure across all stocks in Panels A1 and B1, within the NAICS 3-digit industry in Panels A2 and B2, and within the Hoberg and Phillips (2016) FIC 50-industry in Panels A3 and B3, and we assign all stocks into the portfolios using the NYSE breakpoints. We aggregate stock returns within each portfolio using the stocks' market capitalization on the prior trading day as the weight. *AMH* is the "Artificial Minus Human" zero net investment portfolio long the portfolio with high exposure and short the portfolio with low exposure. We define the ChatGPT event period as from November 30, 2022 to December 14, 2022, i.e., the two weeks after the release of ChatGPT (see Section II for more details). The table then shows placebo event study returns for the ten trading days before (Nov. 14- Nov. 29, 2022) and after (Dec 15 - Dec 29, 2022) the ChatGPT release event period. Panels A1-A3 report the CAPM market-adjusted alphas, and Panels B1-B3 report the Fama-French 5-factor-adjusted alphas. t-statistics are shown in parentheses. See Section II.C for more details.

Q1 (H)	Q2	Q3 (A)	AMH
A1: Market-factor-adjusted alpha (%)			
0.035	-0.019	0.029	-0.006
(0.53)	(-0.39)	(1.25)	(-0.08)
A2: NAICS 3-digit ind.-neutral market-factor-adjusted alpha (%)			
-0.013	0.069	0.020	0.033
(-0.35)	(1.51)	(0.80)	(0.78)
A3: FIC 50 ind.-neutral market-factor-adjusted alpha (%)			
-0.005	0.048	0.052	0.057
(-0.09)	(2.11)	(1.91)	(0.85)
B1: FF 5-factor-adjusted alpha (%)			
0.062	-0.060	0.037	-0.024
(1.33)	(-2.30)	(1.58)	(-0.41)
B2: NAICS 3-digit ind.-neutral FF 5-factor-adjusted alpha (%)			
0.009	0.089	0.022	0.014
(0.32)	(2.80)	(0.83)	(0.30)
B3: FIC 50 ind.-neutral FF 5-factor-adjusted alpha (%)			
0.015	0.021	0.070	0.054
(0.24)	(0.77)	(1.89)	(0.83)

Table IA.8 Cumulative Abnormal Returns and Generative AI Exposure: other abnormal return models.
This table reports the relation between firms' cumulative abnormal returns during the ChatGPT event period, $[-1, 10]$ trading days around the release of ChatGPT on November 30, 2022, (CAR $[-1, 10]$) and firms' labor-based Generative AI exposure defined in Section I. To compute the baseline cumulative abnormal return of each stock, we first subtract the daily market return from Kenneth French's website from the daily stock return from Yahoo Finance, and then we cumulate the daily abnormal returns over $[-1, 10]$. Columns (1)-(2) show abnormal returns as defined in the baseline model. Columns (3)-(4) compute abnormal returns relative to the CAPM model, and columns (5)-(6) relative to the Fama French 5-factor model, where the factor. These abnormal returns are computed by first estimating factor loadings for each firm separately either for the market factor or the Fama French 5-Factor model using data for the 6 months from May 18, 2022 - Nov 14, 2022 (ending two weeks before the GPT release), and then computing abnormal returns for later periods as the difference between the raw excess returns and the returns predicted by the factor loadings. Where indicated, regressions include NAICS 3-digit fixed effects. Equation (5) describes the regression specification and the weighting of firms. t-statistics are computed using heteroskedasticity-robust standard errors and reported in parentheses. ***, **, and * denote significance at 1%, 5%, and 10% levels, respectively.

Dep. var.:	CAR[-1,10]					
Abnormal returns:	Baseline		CAPM		FF5	
	(1)	(2)	(3)	(4)	(5)	(6)
GenAI Exp	0.216***	0.288***	0.229***	0.284***	0.158**	0.278**
	(2.621)	(2.770)	(2.817)	(2.693)	(1.961)	(2.195)
R-squared	0.05	0.36	0.06	0.35	0.03	0.26
Observations	2085	2080	2085	2080	2085	2080
NAICS 3-dig. FE		X		X		X

Table IA.9 Correlations between firm-level Gen. AI exposure, AI product market exposure, and data assets.
This table reports the correlation between the firm-level Generative AI exposure measure defined in Section I.D , the different product market exposure ("PE") measures defined in Section II.C.3, and the data asset ("DA") measures defined in Section II.D.

	GenAI Exp	GPT10K PE	Count10K PE	GS PE	BFHH PE	10K DA	AV DA
GenAI Exp	1.00						
GPT10K PE	0.16	1.00					
Count10K PE	0.12	0.28	1.00				
GS PE	0.05	0.16	0.21	1.00			
BFHH PE	0.15	0.18	0.42	0.20	1.00		
10K DA	0.21	0.38	0.17	0.14	0.13	1.00	
AV DA	0.58	0.36	0.28	0.19	0.31	0.43	1.00

Table IA.10 Product market AI exposure proxies and reported AI use in production by industry. These regressions show the relationship between different product market AI exposure proxies and the reported AI use by firms. The AI use in production by industry is measured as the average rate of "yes" responses to the question *"In the last two weeks, did this business use Artificial Intelligence (AI) in producing goods or services?"* by 3-digit NAICS subsector in the U.S. Census Bureau Business Trends and Outlook Survey from October 23, 2023. The product market AI exposure proxies in the different panels are: (A) *10-K Based Product GenAI Exp* is a dummy variable of whether the firm's products would benefit from a Generative AI boom, as assessed by reviewing the business description section in the firm's recent annual report using a large language model (see more details in Appendix C). (B) *10K-based AI Mentions* is the count of mentions of "AI" or "artificial intelligence" in the business description in the firm's recent annual report. (D) *Share AI Skill Workers* is a measure of the share of a firm's workers with AI skills on their resume constructed by Babina et al. (2024), measured in the last year for which that firm has resume data in the 2018-2021 period. t-statistics based on heteroskedasticity-robust standard errors are shown in parentheses. ***, **, and * denote significance at 1%, 5%, and 10% levels, respectively.

Dep. var.:	AI Use in Last 2 Wks %, Oct. 2023 (BTOS Survey)				
	(1)	(2)	(3)	(4)	(5)
10k-based AI Product Mkt Exp.	18.919***				17.325**
	(3.609)				(2.574)
10K-based AI Mentions		1.262**			-0.499
		(2.229)			(-0.974)
Goldman Sachs Gen. AI Exp.			11.795*		-0.985
			(1.916)		(-0.303)
Share AI Skill Workers				1128.449**	629.899*
				(2.631)	(1.760)
R-squared	0.50	0.33	0.18	0.35	0.57
Observations	57	57	57	57	57

Appendix C. Appendix: Methodology Notes

GPT prompt for task-based exposure scoring

The following prompt structure was based on the rubric language by Eloundou et al. (2023), as well as insights by Willison (2023) and Underwood (2023) about how to best structure API calls for GPT classification. We conducted this categorization on March 29, 2023. To rule out misclassification due to the model's confusion of ChatGPT's known capabilities with ChatGPT's in-development but not well-known capabilities such as image recognition, we follow Eloundou et al. (2023) and add a fourth category (E3) in addition to the main categories used in our analysis, indicating that the task is not exposed to ChatGPT, either directly or by software integration, but can be completed more efficiently if image capabilities of ChatGPT were accessible. Only 0.6% of the tasks are categorized as E3. We consider tasks that are exposed to image recognition to be not exposed throughout our study. Here are the instruction prompts submitted before asking GPT 3.5 Turbo to classify each task statement (using the version as of March 29th, 2023). Note that the order in which the two user-assistant interactions are provided to the API is randomized for each task, and the GPT "temperature" parameter is set to 0:

systemprompt = "Consider the most powerful OpenAI large language model (LLM). This model can complete many tasks that can be formulated as having text input and text output where the context for the input can be captured in 2000 words. The model also cannot draw up-to-date facts (those from <1 year ago) unless they are captured in the input. Assume you are a worker with an average level of expertise in your role trying to complete the given task. You have access to the LLM as well as any other existing software or computer hardware tools mentioned in the task. You also have access to any commonly available technical tools accessible via a laptop (e.g. a microphone, speakers, etc.). You do not have access to any other physical tools or materials. You are a helpful research assistant who wants to label the given tasks according to the rubric below. Equivalent quality means someone reviewing the work would not be able to tell whether a human completed it on their own or with assistance from the LLM. If you aren't sure how to judge the amount of time a task takes, consider whether the tools described exposed the majority of subtasks associated with the task.
# Exposure rubric:
E1 - Direct exposure: Label tasks E1 if direct access to the LLM through an interface like ChatGPT or the OpenAI playground alone can reduce the time it takes to complete the task with equivalent quality by at least half. This includes tasks that can be reduced

to: - Writing and transforming text and code according to complex instructions, - Providing edits to existing text or code following specifications, - Writing code that can help perform a task that used to be done by hand, - Translating text between languages, - Summarizing medium-length documents, - Providing feedback on documents, - Answering questions about a document, - Generating questions a user might want to ask about a document, - Writing questions for an interview or assessment, - Writing and responding to emails, including ones that involve refuting information or engaging in a negotiation (but only if the negotiation is via written correspondence), - Maintain records of written data, - Prepare training materials based on general knowledge, or - Inform anyone of any information via any written or spoken medium.

E2 - Exposure by LLM-powered applications: Label tasks E2 if having access to the LLM alone may not reduce the time it takes to complete the task by at least half, but it is easy to imagine additional software that could be developed on top of the LLM that would reduce the time it takes to complete the task by half. This software may include capabilities such as: - Summarizing documents longer than 2000 words and answering questions about those documents, - Retrieving up-to-date facts from the Internet and using those facts in combination with the LLM capabilities, - Searching over an organization's existing knowledge, data, or documents and retreiving information, - Retrieving highly specialized domain knowledge, - Make recommendations given data or written input, - Analyze written information to inform decisions, - Prepare training materials based on highly specialized knowledge, - Provide counsel on issues, and - Maintain complex databases. ## E3 - Exposure given image capabilities: Suppose you had access to both the LLM and a system that could view, caption, and create images as well as any systems powered by the LLM (those in E2 above). This system cannot take video as an input and it cannot produce video as an output. This system cannot accurately retrieve very detailed information from image inputs, such as measurements of dimensions within an image. Label tasks as E3 if there is a significant reduction in the time it takes to complete the task given access to a LLM and these image capabilities: - Reading text from PDFs, - Scanning images, or - Creating or editing digital images according to instructions. The images can be realistic but they should not be detailed. The model can identify objects in the image but not relationships between those options

E0 - No exposure: Label tasks E0 if none of the above clearly decrease the time it takes for an experienced worker to complete the task with high quality by at least half. Some examples: - If a task requires a high degree of human interaction (for example, in-person demonstrations) then it should be classified as E0. - If a task requires precise measurements then it should be classified as E0. - If a task requires reviewing visuals in detail then it

should be classified as E0. - If a task requires any use of a hand or walking then it should be classified as E0. - Tools built on top of the LLM cannot make any decisions that might impact human livelihood (e.g.hiring, grading, etc.). If any part of the task involves collecting inputs to make a final decision (as opposed to analyzing data to inform a decision or make a recommendation) then it should be classified as E0. The LLM can make recommendations. - Even if tools built on top of the LLM can do a task, if using those tools would not save an experienced worker significant time completing the task, then it should be classified as E0. - The LLM and systems built on top of it cannot do anything that legally requires a human to perform the task. - If there is existing technology not powered by an LLM that is commonly used and can complete the task then you should mark the task E0 if using an LLM or LLM-powered tool will not further reduce the time to complete the task. When in doubt, you should default to E0."

user_prompts = ["", "", "", ""] user_prompts[0] = "Read the following occupation title and one of the tasks that belong to that occupation. Then do three things. 1: Reason step by step to decide which of the labels (E0/E1/E2/E3) from the exposure rubric you were given applies to the task's exposure to LLM. Report. Give an explanation for which label applies and report the label that you think fits best. Do not say zero or N/A.
2: Report only the label that you determined for the task, which should match the label in step 1. Do not reply N/A.
3: Given the amount of speculation required in step 1, describe your certainty about the estimate–either high, moderate, or low.
The task to label follows:
Occupation: Inspectors, Testers, Sorters, Samplers, and Weighers. Task: Adjust, clean, or repair products or processing equipment to correct defects found during inspections."

assistant_prompts = ["", "", "", ""] assistant_prompts[0] = "1: Label (E0/E1/E2/E3): E0. Explanation: The model does not have access to any kind of physicality, and more than half of the task (adjusting, cleaning and repairing equipment) described requires hands or other embodiment.
2: E0.
3: High confidence."

user_prompts[1] = "Read the following occupation title and one of the tasks that belong to that occupation. Then do three things. 1: Reason step by step to decide which of the labels (E0/E1/E2/E3) from the exposure rubric you were given applies to the task's exposure to LLM. Report. Give an explanation for which label applies and report the label that you think fits best. Do not say zero or N/A.
2: Report only the label that you determined for the task, which should match the label in

step 1. Do not reply N/A.
3: Given the amount of speculation required in step 1, describe your certainty about the estimate–either high, moderate, or low.
The task to label follows:
Occupation: Computer and Information Research Scientists. Task: Apply theoretical expertise and innovation to create or apply new technology, such as adapting principles for applying computers to new uses."

assistant_prompts[1] = "1: Label (E0/E1/E2/E3): E1. Explanation: The model can learn theoretical expertise during training as part of its general knowledge base, and the principles to adapt can be captured in the text input to the model.
2: E1.
3: Medium confidence."

Consistency of Generative AI scoring

To validate the consistency and replicability of our procedure that employs the GPT API for classification, we compare the scores assigned across 3 different GPT runs (which may vary in results due to the randomized order of example cases provided, or non-deterministic features of the underlying LLM) for a randomly selected subsample of 100 task statements. We compare the different sets of scores as follows: First, we construct 3 different classifications for each task based on the assigned score: (1) "Current exposure": score 1 has been assigned. (2) "Expected exposure:" Either score 1 or 2 has been assigned. Then, we compute the agreement between different scoring runs with regard to which tasks belong in these categories. The comparison between different runs is shown in Appendix Table IA.1. We find that the agreement between different GPT runs is very high - they arrive at the same score for at least 88% of all cases independent of the exposure classification considered. This validates that GPT reliably provides classifications that are highly consistent across different runs.

Table IA.1 Exposure score variation across GPT scoring runs

	Agreement %	
Score comparison	Current Exposure	Expected exposure
GPT #1 vs. GPT #2	95	90
GPT #1 vs. GPT #3	93	88
GPT #2 vs. GPT #3	96	88

Acemoglu and Autor (2011) measures of skill requirements of occupations

For the analysis shown in Figure 2, we need measures of the skill requirements of different occupations. We draw on standard measures from the literature that allow for comparability of our results to the characteristics of previous waves of automation.

Acemoglu and Autor (2011) construct six skill measures for each SOC occupation based on O*Net measures in the following steps. First, they assign detailed skill requirements from the O*Net's database to each of the six aggregated skill measures, using the following O*Net measurements for each occupation:

Non-routine cognitive: analytical skill:

4.A.2.a.4 Analyzing data/information

4.A.2.b.2 Thinking creatively

4.A.4.a.1 Interpreting information for others

Non-routine cognitive: interpersonal skill:

4.A.4.a.4 Establishing and maintaining personal relationships

4.A.4.b.4 Guiding, directing and motivating subordinates

4.A.4.b.5 Coaching/developing others

Routine cognitive skill:

4.C.3.b.7 Importance of repeating the same tasks

4.C.3.b.4 Importance of being exact or accurate

4.C.3.b.8 Structured v. Unstructured work (taking the reverse value)

Routine manual skill:

4.C.3.d.3 Pace determined by speed of equipment

4.A.3.a.3 Controlling machines and processes

4.C.2.d.1.i Spend time making repetitive motions

Non-routine manual: physical skill

4.A.3.a.4 Operating vehicles, mechanized devices, or equipment

4.C.2.d.1.g Spend time using hands to handle, control or feel objects, tools or controls

1.A.2.a.2 Manual dexterity

1.A.1.f.1 Spatial orientation

Non-routine manual: interpersonal skill

2.B.1.a Social perceptiveness

Second, they obtain an importance scale of each detailed skill requirement for each SOC8 occupation, and standardize the importance scale of each detailed skill to have a mean of 0 and a standard deviation of 1 across occupations.

Third, they compute an occupation's six skill measures a the average of the standardized importance scales of their corresponding detailed skill measures.

Generative AI exposure portfolio construction.

Portfolios for the main realized return analysis are formed from quintiles of stocks that have Yahoo Finance data for Nov. 15, 2022 - March 31, 2023. Quintile thresholds that define value-weighted portfolios within industries or for all stocks are solely based on the sample of stocks listed on NYSE as of the sorting date. All portfolios are formed based on equal weighted sorts on November 29, 2022, and weights for computing portfolio returns are adjusted based on daily returns to mimic passive buy-and-hold exposure.

Industry-neutral portfolios are computed by first forming within-industry equal-weighted tercile portfolios, and then averaging portfolio returns for the same terciles across industries. Returns for within-industry terciles and for all global (not industry-neutral) portfolio quintile sorts are value-weighted, while across-industry averages are industry market-cap. weighted. *AMH* is the "Artificial Minus Human" zero net investment portfolio long highest exposure quantile (A) stocks and short lowest exposure quantile (H) stocks. The data set for estimating portfolio returns consists of daily stock returns from Yahoo Finance and the Fama-French factors, including the market factor and risk free returns are obtained from Ken French's website.

We define the ChatGPT event period as from November 30, 2022 to December 14, 2022, i.e., the two weeks after the release of ChatGPT (see Section II for more details). CAPM market-adjusted alphas and the Fama-French 5-factor-justed alphas are based on the factor loadings of each portfolio estimated using data from the six months preceding Nov. 15, 2022. Abnormal returns are computed as the difference between raw portfolio returns and the product of factor loadings and the factor returns on each day of the event period. Alphas for the event period are then computed as the intercept in a regression of abnormal returns over the event period on a constant, with Newey West standard errors with five lags.

Product Market AI Exposure Measures

We create a number of new measures to try and capture the degree to which a company has product market exposure to AI technology innovations. The data sources and construction of these measures is detailed below.

Company annual report data. The 10-K annual reports filed by companies at the SEC's EDGAR system are obtained in pre-cleaned text files from Bill McDonald's "Software Repository for Accounting and Finance" website, [44] based on the work in Loughran and McDonald (2016). We then use regular expressions to break up the text into the different

[44] See https://sraf.nd.edu/

"items" contained within the report.[45], focusing only on reports filed in 2022, and on the "Business" section of each report.

AI-related business description keywords. For our first product market AI exposure measures we follow Hoberg and Phillips (2016) and use the "Business" description section of a firm's 10-K annual report to infer information about its product markets. We construct a simple measure of AI relevance for the firm's business by tokenizing the text and counting the number of mentions of "AI" or "artificial intelligence". We use the total count of mentions as a proxy for a firm's products either using AI or depending on the use of it by other actors in the value chain. We do this by first lower-casing and tokenizing the text in the "Business" section, constructing lists of all possible unigrams and bigrams contained in the text, and then counting the occurrences of "ai" and "artificial intelligence" in the text.

GPT assessment of business product exposure to AI. The method of counting keywords potentially discards relevant information contained in the full text and the context and interdependence of the AI-related concepts discussed in it. Therefore, we also use a method that uses a large language model to evaluate the text of the business description in the firm's annual report, asking it to determine whether there would be direct positive product market impact of a Generative AI boom on the described company. This allows us to keep the text in its original form and take into account the context within which AI or related topics are discussed. We develop a rubric that provides guidance to the model in the form of asking it to consider whether the firm's products might be involved in enabling or scaling AI technologies, or might benefit from a direct incorporation of the new AI capabilities (the full text of the prompt is shown below). The model is then given two examples of scores applied to company business descriptions and given a new business description and asked to apply a binary label of whether the firm is "directly product market exposed to AI" or not, and also to provide an explanation for the score, which allows for an audit of the model's reasoning. To economize on computing resources and API costs, we only do this analysis for annual reports in our sample that belong to firms for which we have previously computed a task-based AI exposure score, which have stock price data in our sample, and for which the "Business" section contains the word "and" at least three times (which removes annual reports that omitted the "Business" section or where our text extraction parsed a snippet that is too small). We submit the prompt for scoring together with a "Business" section to OpenAI's GPT 3.5 Turbo API for evaluation. We are limited by the context window to evaluating only the first 3000 words of the business description for each firm, which is rarely binding and allows for ample business description for almost all companies - as the

[45]We built on code provided by Yu Zhu at https://yuzhu.run/how-to-parse-10x/ and coding support from ChatGPT in this analysis.

beginning of the section tends to provide the general overview of the company. The result of this procedure are binary product market AI exposure scores for $\sim$ 2.2K companies. Moreover, in early trial runs we audited the explanations provided by the model to ensure that the prompt leads to scoring that closely corresponds to a human scorer's interpretation of the business impact.

Goldman Sachs "near-term AI beneficiaries". This classification is based on the list of firms in the report by Goldman Sachs US Equity Views, August 21, 2023, "The (AI) trade after the trade: Identifying potential long-term EPS beneficiaries of AI adoption", which identifies 11 large firms "directly exposed to the development of AI technology." While this report is produced ex post, this should, if anything, bias it towards better capturing the (at that point) observed product market exposure that leads to a stock market reaction. We use the report to code a binary variable on whether or not a firm is on the following list: NVIDIA (NVDA), Meta (META), Amazon (AMZN), Salesforce (CRM), Marvell Technology (MRVL), Adobe(ADBE), Alphabet (GOOGL), ServiceNow (NOW), Microsoft Corporation (MSFT), Intuit (INTU) and Credo Technology (CRDO).

Resume AI skill share from Babina et al. (2024): we use a measure of the share of workers at a firm that have AI skills on their resumes constructed by Babina et al. (2024), who show that this measure is a good predictor of AI-related product innovations and R&D spending during the pre-ChatGPT wave of AI advancements. We use both the last available data point on the stock of AI skills (as a share of firm employment) as proxies for a firm's level of investment in using AI-related tools pre-ChatGPT. We use the replication data available from the authors and follow their methodology to construct these regression variables based on the Cognism resume AI skill share data provided. The last year of available skill share data is 2021 for the majority of firms in the sample, and we only keep skill share level data if the last available data is from no earlier than 2018.

Product Market AI Exposure: GPT Prompt for Scoring

General context prompt submitted with the completion task to GPT 3.5 Turbo to score company annual report text from the "Business" section for whether it suggests the company has direct product market AI exposure:

```
systemprompt = "You are a financial analyst who is an expert in evaluating company filings for their business
implications. \
    You are evaluating whether a given company's business description in their annual report suggests that the
company's products will be in higher demand as a result of a recent boom in generative artificial intelligence
applications.\
     You are only interested in companies whose product sales might be directly impacted by the Generative AI
boom, not companies that might benefit from labor productivity gains due to incorporating Generative AI into
their processes.\
Among the companies that are likely to see their product market positively affected by the Generative AI boom,
ALL of the following might be reasons why they have positive direct exposure:\
    1) Empowered users. Technology companies leveraging Generative AI technology to amplify their product
functionality  will be directly positively exposed.\n \
    2) Enablers: Select makers of semiconductors and related equipment or other infrastructure needed to build
or deploy Generative AI technology will be directly positively exposed.\n \
    3) Hyperscalers:  Large companies using their extensive cloud computing infrastructures to commercialize
Generative AI on a large scale will be directly positively exposed. \n \
    You are reasoning through how a company's product as described in the business description section of the
annual report might be affected by a boom in demand for Generative artificial intelligence applications, taking
into account the possible impact channels just described.\
    Then,  you use that reasoning to decide if a company is: \n \
        P1: directly positively affected in its product market, which requires that it either supplies products
(e.g. equipment, chips, infrastructure) that will be needed to support the Generative AI boom, or has immediate
benefits in terms of adding AI capabilities into its product, e.g. because of access to proprietary data, or
providing solutions to user problems for which Generative AI is highly complementary. The company should fall
into one of the categories of Empowered User, Enabler, or Hyperscaler to get a P1 exposure score. \n \
        P0: no direct positive impact in its product market, as it does not sell inputs into the Generative AI
activities of other firms and does not have immediate ability to incorporate new AI capabilities into its
products. This includes more remote indirect impacts from generative AI leading to broader increases in
innovation or demand for the industry but not necessarily for this company in particular. \n \
         Note that you should assign P1 only when there is immediate product market impact (e.g. the company
product involves to a large degree chatbots or text generation that Generative AI can easily improve, or
supplies key inputs), but should assign P0 (not exposed) when the AI impact on the firm's product capabilities
is more remote, or only somewhat probable, such as requiring future AI-based tools to be built by other
companies that might then complement the firm's product. Also note that companies that provide critical inputs
into Generative AI usage, such as infrastructure or hardware can also be directly affected and should get a
score of P1, as those components are essential for using AI and will see higher demand in a boom."
```

Prompt preceding the "Business" text submission":

```
prefix = "Read the following business description from a company's annual report. Then do two things. \
1: Reason step by step to apply a score of P0 (not directly positively exposed) or P1 (directly positively
exposed) to rate its product market exposure. Use the rubric and generative AI product market exposure
categories provided to you before (considering whether the company might be an Empowered User, Enabler, or
Hyperscaler, all of which positively exposed) to decide whether the company's product market is going to be
positively affected by the boom in Generative artifical intelligence. \
Report the label that you think fits best and provide up to two sentences that summarize your reasoning for the
product market exposure score that you assigned. Do not say N/A.\n \
    2: Report only the product market exposure score that you determined for the company, which should match
the score in step 1. Do not reply N/A. \n\
   The business description follows:\n  "
```

Overall prompt structure (here, *fulltext* is the "Business" description to be scored):

```
prompt = [{"role": "system", "content": systemprompt},
      {"role": "user", "content": user_prompts[0]},
        {"role": "assistant", "content": assistant_prompts[0]},
       {"role": "user", "content": user_prompts[1]},
       {"role": "assistant", "content": assistant_prompts[1]}]
```

Measuring Firm "Data Value" for Generative AI

In order to measure the "data value" of a company that might contribute to its ability to productively deploy Generative AI tools and their analytics capabilities, we develop a number or new measures to quantify the amount of data that a company has effective access to. Similar to the product market Generative AI exposure measurement approach described above, we again review the business description section in the firm's recent annual report using both traditional NLP approaches and a large language model. Again, we follow Hoberg and Phillips (2016) and use the "Business" description section of a firm's 10-K annual report

to infer information about its product markets. We also develop an alternative approach of measuring data value based on the predicted share of "data management" roles in a firm's employment structure, which is inspired by the analysis in Abis and Veldkamp (2023).

Data-related business description keywords. We construct a simple measure of data value for the firm's business by tokenizing the text of the business description in the firm's annual report filed in 2022 and counting the number of mentions of "data". We use the total count of mentions as a proxy for the importance of data in firm's existing business. We do this by first lower-casing and tokenizing the text in the "Business" section, constructing lists of all possible unigrams contained in the text, and then counting the occurrences of "data" in the text. In the regression analysis we use an indicator of a non-zero count of "data" mentions as our proxy for a firm having valuable data.

GPT assessment of business "data value" for LLMs. Again, we also use a method that uses a large language model to evaluate the text of the business description in the firm's annual report. We develop a rubric that provides guidance to the model in the form of asking it to first consider subcategories of data relevance. In particular, the LLM is asked to first assess whether a firm's business description suggests access to relevant data based on its coverage of 6 categories: the general nature of the company's business, the scale and reach of the firm, data collection mechanisms, data utilization, data infrastructure & management, and data regulation and privacy (full prompt shown below). The model is asked to assign a score of 0, 1, 2, or 3 (no, little, moderate, or high relevance) in each category. Only then is the model asked to also provide an overall score for the degree to which a firm is likely to have data that can be used as an input into LLM analytics (low, moderate, or high data value).

We submit the prompt for scoring together with a "Business" section to OpenAI's GPT 3.5 Turbo 16k API for evaluation. We submit only the first 3000 words of the business description for each firm, which is rarely binding and allows for ample business description for almost all companies - as the beginning of the section tends to provide the general overview of the company. In our analysis based on this scoring, we use both the LLM's 0-3 assessment of the "overall" relevance of the company's data for LLM analytics, as well as a binary indicator for whether any of the subcategory scores was assessed as a 3 (high relevance).

Data Management Skill. As an alternative measure of a company's effective ability to leverage data in combination with LLM analytical capabilities, we build on the insight in Abis and Veldkamp (2023) that the prevalence of "data management" skills in a firm's employment indicates the accumulation of valuable data. First, we predict the likelihood that a U.S. job posting for a 6-digit SOC occupation mentions at least 3 "data management

skills". The se skills are classified based on a list of words indicating relevant skills in job postings from Abis and Veldkamp (2023), which we fuzzy match to skill tags in Lightcast job posting data using Stata's *matchit* command. We retain all matches with a similarity score above 0.7, and then manually inspect whether the matched Lightcast skill tag actually corresponds to a data management skill or was a spurious match.

Then, we count the number of such data management skills present in each job posting for 2017-2021—the 5 years preceding the year when ChatGPT was released. We aggregate the posting-level data into occupations to compute the probability that a U.S. job posting for a 6-digit SOC occupation mentions at least 3 data management skills in this time period—which we define as the likelihood that a job in this occupation has "high data management skill."

Last, we again use the LinkedIn occupational employment distribution at each firm in 2022, together with these occupation-level expected shares of high data management skill jobs, to predict the probability that a job at a firm is a "high data management skill" position. That is, this measure represents the predicted share of a firm's positions on LinkedIn that are "data management"-intensive. In our analysis, we use both this predicted share and a binary variable for firms in the top tercile of this predicted data management intensity among its workers—as proxies for the degree to which company is likely to have valuable data based on its employment structure.

Data Value Assessment: GPT Prompt for Scoring

General context prompt submitted with the completion task to GPT 3.5 Turbo 16k to score company annual report text from the "Business" section for whether it suggests the company has data that would be valuable as an input for LLM analytics:

```
systemprompt = "You are an equity analyst who is an expert in assessing companies' intangible assets.  Analyze
the 'Business Description' section of the provided 10-K annual report. You are to evaluate specifically the
company's potential data value as it would pertain to complementing Large Language Model-driven analytics. \n \
    To ensure a thorough analysis, consider the following dimensions in which the company's business
description might reveal the value of  a company's data : \n \
    1) Nature of Business (NOB): Examine and specify if the company is in a data-centric industry, such as e-
commerce, or a more traditional industry. Detail the nature of its data interactions. \n \
    2) Scale and Reach (SAR): Assess the geographical scope and customer breadth of the company. Does it
operate globally with a diverse user base, or is its impact more regional or niche? Provide specific examples.
\n \
    3) Data Collection Mechanisms (DCM): Elaborate on how the company collects its data. Highlight if it
actively collects through proprietary platforms or devices. Identify any passive collection methods or
partnerships focused on data. \n \
    4) Data Utilization (DU)): Dive deep into the company's application of its data. Describe if there are
methods of direct monetization or if the data supports business processes or provides insights. Identify and
discuss any future plans mentioned about leveraging data. \n \
    5) Infrastructure & Management (IM): Extract and comment on any mentions of investments in data-centric
infrastructure or tools, collaborations related to data, or any dedicated internal teams focused on data. \n \
    6) Regulation & Privacy (RP)): Detail the company's approach to data regulation and privacy. Outline
mentions of data security, potential legal hurdles related to data, or initiatives ensuring ethical data usage.
\n  \
    For each of the six points above, provide a clear reasoning for your conclusions. \n \
    Score the company on a scale of 0-3 for each point, where 0 is 'No mention or not applicable', 1 is
'Limited evidence', 2 is 'Moderate evidence', and 3 is 'Strong evidence'. Conclude only after you've addressed
all points, and base your summary on the evidence provided, specifically focusing on the company's data value
potential for LLM analytics. \n \
    Put your overall summary on a scale of 0-3,  where 0 is 'No valuable data',  1 is 'Limited data value', 2
is 'Moderate data value', 3 is 'High data value'.    "
```

Prompt preceding the "Business" text submission":

```
prefix = "Read the following business description from a company's annual report. Then do two things. \
1: Reason step by step to apply a score of 0,1,2, or 3 in each of the six categories provided, which highlight
different ways that data value relevant for LLM analytics might be revealed in the business description. Also
apply a summary data value score of 0,1, 2, or 3. \
Report your scores for each category, and your summary score, and provide one sentence that summarizes your
reasoning for each category. \n \
    2: Report only the scores for each of the six categories and the overall sore, preceding them by the
abbreviated category names (NOB, SAR, DCM, DU, IM, RP, Overall),  which should match the scores reported in
step 1. Do not reply N/A. \n\
   The business description follows:\n \  "
```