NBER WORKING PAPER SERIES

ARTIFICIAL INTELLIGENCE AND THE LABOR MARKET

Menaka Hampole
Dimitris Papanikolaou
Lawrence D.W. Schmidt
Bryan Seegmiller

Working Paper 33509
http://www.nber.org/papers/w33509

NATIONAL BUREAU OF ECONOMIC RESEARCH
1050 Massachusetts Avenue
Cambridge, MA 02138
February 2025

We are grateful to Huben Liu, Weizhe Sun, and Tim Zhang for excellent research assistance. The views expressed herein are those of the authors and do not necessarily reflect the views of the National Bureau of Economic Research.

Artificial Intelligence and the Labor Market
Menaka Hampole, Dimitris Papanikolaou, Lawrence D.W. Schmidt, and Bryan Seegmiller
NBER Working Paper No. 33509
February 2025
JEL No. E20, J01, J24, O3, O33

ABSTRACT

We leverage recent advances in NLP to construct measures of workers' task exposure to AI and machine learning technologies over the 2010 to 2023 period that vary across firms and time. Using a theoretical framework that allows for a labor-saving technology to affect worker productivity both directly and indirectly, we show that the impact on wage earnings and employment can be summarized by two statistics. First, labor demand decreases in the average exposure of workers' tasks to AI technologies; second, holding the average exposure constant, labor demand increases in the dispersion of task exposures to AI, as workers shift effort to tasks that are not displaced by AI. Exploiting exogenous variation in our measures based on pre-existing hiring practices across firms, we find empirical support for these predictions, together with a lower demand for skills affected by AI. Overall, we find muted effects of AI on employment due to offsetting effects: highly-exposed occupations experience relatively lower demand compared to less exposed occupations, but the resulting increase in firm productivity increases overall employment across all occupations.

Menaka Hampole
Yale School of Management
menaka.hampole@yale.edu

Dimitris Papanikolaou
Kellogg School of Management
Northwestern University
2211 Campus Drive, Office 4319
Evanston, IL 60208
and NBER
d-papanikolaou@kellogg.northwestern.edu

Lawrence D.W. Schmidt
Sloan School of Management
Massachusetts Institute of Technology
100 Main Street
Cambridge, MA
ldws@mit.edu

Bryan Seegmiller
Northwestern University
Kellogg School of Management
bryan.seegmiller@kellogg.northwestern.edu

Recent advances in artificial intelligence have re-ignited the perennial concern that technology will automate away most tasks performed by workers and lead to large declines in labor demand, depressed wages and diminished job opportunities for workers. In contrast to prior waves of technological change, which have largely exposed middle- and low-skilled occupations (Autor, Katz, and Kearney, 2006; Autor and Dorn, 2013; Kogan, Papanikolaou, Schmidt, and Seegmiller, 2023), AI exposure appears to be concentrated in white-collar jobs (Webb, 2020; Eloundou, Manning, Mishkin, and Rock, 2023). However, despite the fact that firm investments in artificial intelligence have been underway for well over a decade, measuring the impact of AI improvements on labor demand has been elusive.[1] Part of the challenge is that, advances in AI that are related to an occupation's tasks may actually increase demand for the job, for instance if it increases firm productivity.[2] Our goal is to shed light on the distinct channels through which AI affects overall labor demand by using theory to guide measurement.

We begin by introducing a model that nests several direct and indirect channels. In our model an occupation performs a collection of tasks. A new technology that affects a specific occupation is characterized by the degree of improvements in a task-specific (intangible) capital that is a substitute for labor in each task. Some technologies may affect all the tasks that a given occupation performs (such as a customer service chat box), or their application could be limited to a small set of tasks (such as a system for automatically filing expense reports). Importantly, workers can optimally allocate their time across these tasks. An improvement in the automation technology for a specific task has a direct effect on the price for that task, but also indirect effects on the other tasks performed by the same worker. These indirect effects depend on the ease of reallocating effort across tasks and several elasticities: the elasticity of substitution between capital and labor, and the degree of complementarity across tasks within an occupation, and across occupations within a firm.

The key new insight delivered by our model is that the labor market impact of a specific AI technology can be summarized by a two key statistics. First, the *mean* exposure of an occupation's tasks to AI is in general negatively related to demand for that occupation. Thus, a moderate improvement in a technology that is related to all the tasks of a particular occupation will lower demand—a customer service chatbox with modest capabilities will still reduce demand for customer service agents. Second, the *dispersion* in occupational task exposure to AI increases labor demand for that particular occupation, holding the mean exposure constant. An automated system for generating expense reports allows workers to reallocate their effort towards other tasks. These two

[1] For example, while Acemoglu, Autor, Hazell, and Restrepo (2022) find that firms with AI-exposed workforces have reduced job postings for non-AI positions, any aggregate impacts of AI-labor substitution on employment and wage growth in more exposed occupations and industries have been too small to detect.

[2] See for instance Acemoglu and Restrepo (2018, 2021). In addition, some technologies may complement rather than substitute for labor (Autor, Chin, Salomons, and Seegmiller, 2024; Kogan et al., 2023), which further complicates the link between a job's exposure to AI and labor demand.

measures are sufficient to characterize the shift in labor demand within a firm. The overall effect, however, also depends on the resulting increase in firm productivity due to AI adoption.

In sum, the model implies that whether AI actually displaces an AI-exposed occupation is highly ambiguous, as it depends on the relative strength of direct substitution effects, indirect effects that operate through reallocation of effort in potentially complementary tasks, and aggregate effects that operate through changes in firm productivity. Specifically, even with a potentially high degree of capital-labor substitution at the task level, these opposing forces make the net effect on relative occupational labor demand within the firm unclear, as it depends on several other elasticities: the elasticity of substitution across tasks within an occupation and the elasticity of substitution across occupations within a firm. Additionally, since occupations within a firm are not perfect substitutes, improvements in technology that affect a subset of occupations can increase firm productivity and ultimately labor demand even for the directly affected occupations.

The main part of the paper focuses on teasing out the relative importance of these forces. We do so by leveraging recent advances in large language models (LLMs) and natural language processing (NLP) techniques applied to a rich corpus of resume and job posting data from Revelio Labs. The first step consists of constructing direct analogues of the model-implied measures and therefore separate the average task-level exposure to AI from the dispersion in task-level exposure to AI. Our resume data allows us to create proxies for the adoption of specific AI applications by specific firms. We use LLMs to identify the exact workers who are implementing AI technologies at which firms, as well as *how* they are applying AI at the firm. Using modern NLP methods, we then estimate the semantic similarity between these AI applications and individual tasks performed by specific occupations from ONET. Our operating assumption is that AI applications that are similar to specific tasks performed by a specific occupation are a *a substitute* for these tasks.

We find that higher-paid occupations are, on average, more exposed to AI—the workers at the 90th percentile of the pay distribution have the highest mean exposure. This pattern is similar if we instead examine the dispersion in task exposures to AI within an occupation. Combining these similarity scores with a measure of intensity of AI adoption at the firm level—based on the number of employees implementing AI applications—we then construct the empirical equivalent of our model implied measures: the extent to which workers in a specific occupation, in a particular firm, at a point in time are exposed to AI across all of their tasks (the mean) or highly exposed but only in some of their tasks (the dispersion).

A key challenge in our empirical analysis is that AI adoption is not necessarily random across firms. Indeed, examining how our measure of AI utilization correlates with firm characteristics during the 2014 to 2023 period among the sample of publicly traded firms, we see patterns that are largely consistent with prior evidence. In particular, consistent with the findings of Acemoglu, Anderson, Beede, Buffington, Childress, Dinlersoz, Foster, Goldschlag, Haltiwanger, Kroff, Restrepo, and Zolas (2023a), firms with high AI utilization tend to be larger, more productive, and pay higher wages. In addition, firms that utilize AI tend to experience faster growth in sales, productivity, profits, and employment, consistent with Acemoglu et al. (2023a); Babina, Fedyk, He, and Hodson (2024). Thus, to understand the impact of AI on productivity, we need to separate selective adoption from the causal impact of AI on productivity.

To address the endogeneity of firm adoption, we instrument for firm's employment of AI-focused employees with the growth in AI employees in other firms that the treated firms tends to hire workers from. Thus, if United Health tends to hire workers from Target, and Target tends to employ a lot of workers that are developing AI applications, some of these workers will be hired by United Health. The identifying assumption is that these pre-existing networks are exogenous to the benefits and costs of AI adoption (Goldsmith-Pinkham, Sorkin, and Swift, 2020), so we exclude AI-implementing employees when constructing firm to firm job flows. Using our identification strategy leads to significantly larger effects of AI adoption on firms' growth rates compared to our OLS estimates.

We next turn our attention to the demand for workers. An advantage of our resume data is that it provides a rich snapshot of employment of workers in specific occupations by particular firms. Consistent with the model, we find that the average exposure of an occupation's tasks to the AI applications adopted by their employer is significantly negatively related to subsequent employment growth. By contrast, greater dispersion in the occupation's exposure to these AI applications is associated with greater employment growth. The granularity of our AI exposure measures allow us to saturate our specifications with a rich set of controls, including the interaction of firm and occupation with calendar year fixed effects, with the estimates largely comparable across specifications. Put differently, our specifications compare employment growth in differentially exposed occupations within the same firm, but also in the same occupation across different firms that adopt AI with different intensities or focus.

Naturally, AI-based automation can be directed to specific occupations, for instance if these workers are scarce. Accordingly, we also construct a shift-share instrument for our firm-level measures that is based on the mean and dispersion in task exposure to AI across all applications in the same period (the shift) times the predicted intensity of AI adoption at the firm level—using our firm-level instrument based on hiring networks. The IV results are comparable to our OLS results qualitatively, but are quantitatively larger. For instance, in our preferred IV specification, a standard deviation increase in our measure of mean task-level exposure to AI leads to approximately a 14 percent decline in the within-firm employment share of affected occupations, while a standard deviation increase in the dispersion measure leads to an 8 percent increase in the within-firm employment share. Overall, these results are consistent with both the presence of strong AI–task substitution, but also productivity spillovers across tasks within an occupation, which considerably dampen this direct substitution effect. The end result is that the overall effect of AI-exposure on the relative demand for affected occupations within the firm is muted.

Next, we examine the specific implications of our model regarding the impact of AI technologies on the nature of tasks that specific occupations perform. In particular, improvements in technology that automate specific tasks should imply that workers will devote less effort in those tasks. Using the Revelio job posting data, we therefore examine whether the specific skills that are related to the particular tasks that are directly affected by AI are less likely to appear in subsequent job postings. We find that, within a treated occupation–firm pair, a one standard deviation increase in *task*-level AI exposure reduces by 4.5% the relative demand for skills related to that task, as a share of the total skill requirements for that job.

The last part of our analysis uses our empirical estimates to quantify the net impact of AI on firm labor demand, both overall, but also across different occupations, job types, and worker earnings levels. In particular, we compute the total net marginal effect of AI use using our coefficient estimates and the joint empirical distributions of the mean and the dispersion of task-level AI exposure combined with our estimate firm-level AI utilization. Overall, we find that the aggregate impact of AI on firm labor demand is muted, due to the presence of counter-veiling forces. Even though the direct substitution effect (the mean task exposure) is quantitatively stronger, there are significant labor-augmenting effects arising due to task reallocation (the dispersion across task exposures) and the impact of increased firm productivity on firm labor demand.

In addition, these effects vary somewhat across the worker pay distribution, but the net effects are more homogeneous across jobs than the underlying forces would suggest. In particular, the labor-substitution part of AI is stronger across higher-paid occupations, but then so is the reallocative effect. Thus, focusing on within-firm job reallocation, we find that employment of highly-exposed occupations (those at the 90-th percentile of the pay distribution) declines by about 3.5% relative to the employment of the last exposed workers (those at the bottom of the pay distribution). After taking into account the firm productivity effect, this effect is mildly reversed, since the jobs that are more exposed to AI are more likely to exist in firms that adopt AI and realize productivity gains, Thus, the jobs that are most exposed to AI actually experience a *slight increase* in their share of aggregate employment compared to less exposed occupations.

Further, the most adversely impacted occupations fall under the business, financial, and engineering categories. In terms of share of overall employment, our estimates imply that these jobs experienced a decline of 2% to 2.5% over a five year period. As before, the existence of task complementarities and a strong firm productivity effect act as offsetting factors to the significant substitution effect at the task-level. At the same time, however, our estimates also imply declines in the overall employment share of less exposed occupations, such as 'Food preparation and serving'. Even though these occupations face very little task exposure to AI, the fact that their employers do not utilize AI implies that they grow slower than the firms that do, leading to lower labor demand overall.

Overall, this exercise illustrates why it may be difficult to detect the impact of AI technologies on labor reallocation overall, despite the strong evidence of AI–labor substitution at the task level. That said, however, we do find that our measures of AI exposure still have meaningful explanatory power for the realized reallocation of workers across jobs. Comparing the realized employment share growth at the occupation level with the total AI–predicted changes, we find that about 14 percent of the variation in observed aggregate occupational employment share changes can be attributed to the net effects of AI exposure. About half of this is driven by the reallocation away from AI–exposed occupations within the firm, with the remainder coming from the average firm–level AI use by occupations' employers.

Our work contributes broadly to existing work in economics studying the impact of technological change on the labor market. Closest to our paper is work that explores the impact of machine learning AI on firm productivity and the labor market. Acemoglu et al. (2022) find some evidence that AI substitutes for labor at the establishment level, but they find essentially no impact of employment and wage growth for exposed occupations. Babina et al. (2024) find that AI adoption has an impact on firm growth, but find no impact that it leads to productivity gains or job automation. Acemoglu et al. (2023a) argue that the correlation between AI adoption and firm growth is driven by selection of which firms adopt advanced technologies. Last, Gathmann, Grimm, and Winkler (2024) argue that, in contrast to robots, AI has reduced the demand for abstract tasks and increased the demand for certain routine tasks.

Our measure of AI ends in 2023, and therefore it largely excludes the recent rise of generative AI (GenAI). Despite the relatively short period since GenAI first became broadly available (November 2022), some early work has studied its effect on firms and workers. Eloundou et al. (2023) construct an occupation-level exposure measure to generative AI (GenAI). They document that most occupation have significant exposure to GenAI, and unlike past instances of automation, high-wage occupations are significantly more exposed than low- or middle-wage jobs. Eisfeldt, Schubert, Taska, and Zhang (2023) build on Eloundou et al. (2023) and construct a firm-level measure of workforce exposure to GenAI. Using data on job postings and average wages at the occupation level, they argue that GenAI has helped firms improve profitability by reducing labor costs. In contrast to Eloundou et al. (2023); Eisfeldt et al. (2023) , Auer, Köpfer, and Sveda (2024) argue that GenAI can be a complement for high-wage workers, and argue that low-wage jobs are more likely to be displaced in the future. Humlum and Vestergaard (2024) find that the adoption of GenAI tools by workers is widespread but uneven, with increased adoption among younger, higher-achieving and male workers. Acemoglu (2024) argues that generative AI will likely have small effects on productivity, using a version of Hulten's theorem.

Compared to the existing body of work, our work has some key advantages on the measurement side. Specifically, our data and methodology allows us to construct direct measures of the exposure of a particular occupation in a specific firm to Artificial Intelligence; our measures are grounded in theory and they differentiate between the direct substitution effect of AI technologies on labor demand from the reallocative effect. Similar to Babina et al. (2024); Babina, Fedyk, He, and Hodson (2023), we use information in online resumes to determine whether a firm is actually adopting AI; in contrast to their work, however, our exposure measures are based on how each particular AI application adopted by a specific firm is related to the particular tasks a given occupation performs. The detailed nature of our measure allows us to separate the impact of AI from occupation-specific trends in employment, unlike existing measures that vary only across occupations (Webb, 2020; Felten, Raj, and Seamans, 2018; Brynjolfsson, Mitchell, and Rock, 2018; Eloundou et al., 2023; Eisfeldt et al., 2023). Further, our focus on actual changes in employment, rather than job postings (Acemoglu et al., 2022; Eisfeldt et al., 2023), helps address the concern that firms may hire multiple employees from one job posting. The increased granularity of our data helps us understand why the aggregate employment effects of AI have been muted, despite clear evidence for substitution at the micro level.

More broadly, our work connects to the literature studying the broad effects of labor-substituting technologies. An important stream has emphasized that automation can substitute for routine tasks, and therefore an occupations share of routine's task is indicative of its exposure to labor-saving technologies over the last several decades (Autor et al., 2006; Acemoglu and Autor, 2011; Goos, Manning, and Salomons, 2014). Another stream has emphasized constructing direct measures of labor-saving technologies and exploring their impact on the labor market, which are either based on robots (Acemoglu and Restrepo, 2021; Graetz and Michaels, 2018; Humlum, 2019; Dauth, Findeisen, Suedekum, and Woessner, 2021; Koch, Manuylov, and Smolka, 2021; Bonfiglioli, Crinò, Fadinger, and Gancia, 2020; de Souza and Li, 2023; Benmelech and Zator, 2022) or on textual analysis of patent documents (Webb, 2020; Kogan et al., 2023; Autor et al., 2024; Mann and Püttmann, 2023; Dechezleprêtre, Hémous, Olsen, and Zanella, 2021).

1 Theoretical Framework

To guide measurement, we first begin with a simple model framework based on Acemoglu and Autor (2011); Acemoglu and Restrepo (2018); Caunedo, Jaume, and Keller (2023); Kogan et al. (2023).

Relative to prior work, our model describes the productivity effects of dispersion in exposure to technological change within an occupation, and allows for endogenous re-organization of effort to emphasize different tasks in response to labor-substituting technological shocks to exposed tasks. The model features workers of different occupations who each perform different tasks. The model has the following key features: labor and capital are substitutes in production, and capital is specific to each task. Technological improvements are reflected in declines in the (quality-adjusted) price of capital. Within a job, workers allocate their time among different tasks. Workers optimally choose across jobs taking occupation- and firm-level wages as given along with idiosyncratic preference shocks.

The model has a simple prediction: improvements in technology that substitutes for labor in particular tasks can either increase or decrease labor demand for a specific occupation depending how these technology improvements affect the capital that is specific in her tasks. In particular, if a technology uniformly improves the capital that is specific to most of the worker's tasks, then labor demand for that occupation decreases, since capital substitutes for associated labor tasks. By contrast, if the technology improves capital in a disparate fashion—some tasks are greatly affected while others are not—then labor demand for an occupation is likely to increase. The reason is that tasks are complements and workers can endogenously allocate their time among tasks. Thus, if a certain improvement is significantly better than labor in performing a specific task, the productivity of the other tasks will increase and workers will re-allocate more time to them.

1.1 Setup

There is a continuum of firms that produce aggregate output $\bar{Y}$ as a CES composite of the output Y_f of different firms,

$$\bar{Y} = \left(\int_{\mathcal{F}} Y_f^{\frac{\theta-1}{\theta}} df \right)^{\frac{\theta}{\theta-1}} . \tag{1}$$

Here, θ captures the elasticity of substitution across firms. Each firm produces a differentiated good by combining the output of many occupations,

$$Y_f = \left(\int_{O} Y(o,f)^{\frac{\chi-1}{\chi}} \right)^{\frac{\chi}{\chi-1}} . \tag{2}$$

Firms make profits because of imperfect competition, reflecting both monopolistic competition in product markets and monopsonistic power in labor markets. We denote the firm's markup over marginal cost by $\Theta = \frac{\theta}{\theta-1}1$. Due to the presence of monopsony power, the firm's marginal cost will exceed its average cost and the firm will mark down the wage it pays below the marginal cost of labor.

Workers in occupation o employed in firm f produce output $Y(o,f)$ by combining the output of J individual tasks. The total output of occupation o in firm f is given by

$$Y(o,f) = \left(\sum_j y(j)^{\frac{\psi-1}{\psi}} \right)^{\frac{\psi}{\psi-1}} \tag{3}$$

Here, ψ denotes the elasticity of substitution across tasks within a given job, which determines the elasticity of labor demand for each task. To simplify the notation, we will suppress the firm subscript and occupation subscripts unless needed.

Each task j in job (o,f) is produced by a labor input $l(j)$ and a capital input $k(j)$,

$$y(j) = \left(\gamma_j \, l(j)^{\frac{\nu-1}{\nu}} + (1-\gamma_j)\, k(j)^{\frac{\nu-1}{\nu}} \right)^{\frac{\nu}{\nu-1}}. \tag{4}$$

In the context of our application, we should think of $k(j)$ as intangible capital (e.g. software algorithms) that can substitute for labor in a specific task. Here, ν gives the elasticity of substitution between capital $k(j)$ and labor $l(j)$, while ψ denotes the elasticity of substitution across tasks within an occupation. In what follows, we will be assuming that $\nu > \psi$, which will imply that improvements in the technology that is specific to task j are likely to be labor-saving.

We model the impact of technological innovation as a reduction in $q(j)$, the quality-adjusted price of intangible capital $k(j)$ that is specific to task j,

$$\Delta \log q(j) = -\varepsilon(j). \tag{5}$$

A specific technology is potentially applicable to several tasks within a job. A given technological improvement that is applicable to job (o,f) can therefore be represented as a firm- and occupation-specific vector $\varepsilon \equiv [\varepsilon_1 \ldots \varepsilon_J]$ of weakly positive random variables. If $\varepsilon(j) > 0$, that implies that the firm is adopting an improved (or cheaper) labor-saving technology that is specific to task j.

Workers in job (o,f) optimally choose the amount of time they allocate in each task. The effective supply of labor by worker i in task j is given by

$$\ell(j) = h(j)^{1-\beta}. \tag{6}$$

Here, the parameter $\beta \in (0,1)$ captures the degree of decreasing returns to effort at the task level. If $\beta \to 0$ then there are no decreasing returns at the task level, whereas if $\beta \to 1$ then the quantity of effort is essentially fixed in each task. More generally, a smaller value of β implies that there is more scope of reallocating effort across tasks.

The total number of hours a worker can supply across all J tasks is equal to one. Given the above, the optimal time a worker in job (o, f) will allocate to task j is equal to

$$h(j) = \frac{w(j)^{\frac{1}{\beta}}}{\sum_{j \in J} w(j)^{\frac{1}{\beta}}}, \tag{7}$$

where $w(j)$ is the firm- and occupation-specific wage in task j. Thus, a worker's total earnings in job (o, f) are equal to

$$W(o, f) \equiv \sum_{j \in J_o} h(j)^{1-\beta} w(j), \tag{8}$$

which are a function of her allocation of time and the (job-specific) task prices $w(j)$.

There is a continuum of measure one of ex-ante identical workers who chose a firm–occupation pair based on the total earnings of that occupation and an idiosyncratic taste shock. As in Eaton and Kortum (2002), each worker draws a set of job-specific taste shocks that are independent and identically distributed according to a Fréchet distribution with scale parameter 1 and a shape parameter $1 + \zeta$. Using the properties of the Fréchet distribution, the measure of workers that choose job (o, f) is given by

$$N(o, f) = \bar{\zeta}\, W(o, f)^{\zeta}, \tag{9}$$

where $\bar{\zeta}$ is a constant defined in the Appendix.

1.2 Model Implications

To derive the model's implication for measurement, we can examine the impact of a given technology ε on equilibrium earnings and employment. Denote by

$$\eta_o \equiv \frac{\partial \log w(j)}{\partial \varepsilon(j)} \tag{10}$$

and

$$\eta_c \equiv \frac{\partial \log w(j)}{\partial \varepsilon(j')} \tag{11}$$

the own- and cross-elasticity of task-specific prices $w(j)$ to improvements $\varepsilon(j)$ and $\varepsilon(j')$ to along the symmetric equilibrium with $\gamma_j = \gamma$, $w(j) = w$, and $q(j) = q$ for all $j \in J$. We solve for these elasticities by a log-linear approximation around the symmetric equilibrium—see Appendix A.1 for details. We obtain

$$\eta_o = -\frac{s_k}{J} \frac{(\nu - \chi)(1-\beta) + \beta \Big((J-1)(\nu - \psi)\zeta + \nu(\psi - \chi) + J\,(\nu - \psi)(s_k\,\nu + s_l\,\chi) \Big)}{\Big(s_k\,\nu + s_l\,\chi + \zeta \Big) \Big(1 - \beta\,(1 - s_k\,\nu - \psi\,s_l) \Big)}, \tag{12}$$

and

$$\eta_c = -\frac{s_k}{J}\frac{(\nu-\chi)(1-\beta) - \beta\Big(\nu\,(\chi-\psi) + \zeta(\nu-\psi)\Big)}{\Big(s_k\,\nu + s_l\,\chi + \zeta\Big)\Big(1-\beta\,(1-s_k\,\nu - \psi\,s_l)\,\Big)}. \tag{13}$$

The first elasticity η_o captures the impact of the technology improvement that is specific to task j on the price of task j. As long as the elasticity of substitution between capital and labor ν is sufficiently high, the elasticity η_o is negative: improvements in the technology that substitutes for labor in task j leads to lower labor demand for that task and therefore a lower task price $w(j)$. In the limit where the technology becomes a perfect substitute for labor ($\nu \to \infty$ then the elasticity η_o converges to minus one as wages fall one to one with technology improvements. More broadly, a sufficient condition for the elasticity to be negative is that the number of tasks is sufficiently large and $\nu > \psi$.

The second elasticity η_c captures cross-task spillovers. In general, improvements in technology that are specific to task j can have spillovers on tasks that are not directly affected. The strength, and sign, of these spillovers is in general ambiguous as they depend on the elasticity of substitution across tasks ψ and the rate of decreasing returns to effort at the task level b. For most parameter values, the cross-elasticity η_c is negative—unless tasks are sufficient complements, that is ψ is sufficiently low. In addition, as long as the elasticity of substitution between capital and labor ν is sufficiently high, it is increasing as the ease of reallocation increases (β falls).

Given the above, we can express the resulting change Δ_ε of log wage earnings (8) in job (o, f) in response to technology ε using a second order approximation around the symmetric equilibrium.

$$\Delta_\varepsilon \log W(o,f) \approx \Big(\eta_o + \eta_c\,(J-1)\Big)m(\varepsilon) + \frac{1}{2\,\beta}\,(\eta_o+\eta_c)^2\;V(\varepsilon) \tag{14}$$

where $m(\varepsilon)$ denotes the mean improvement of the technology across all tasks,

$$m(\varepsilon) \equiv \frac{1}{J}\sum_{j\in J}\varepsilon(j), \tag{15}$$

and $V(\varepsilon)$ denotes the dispersion across tasks,

$$V(\varepsilon) \equiv \frac{1}{J}\sum_{j\in J}\Big(\varepsilon(j) - m(\varepsilon)\Big)^2. \tag{16}$$

Examining equation (14), we see that it has two terms that are related to the mean level of technology improvement and its dispersion, respectively. The first term combines two effects: the direct effect of the average improvement of the technology on worker earnings, which is a function of the two elasticities η_o and η_c. The latter elasticity is multiplied by the number of tasks $J-1$

because η_c is decreasing in J, while the product converges to a positive constant as $J \to \infty$.

The second term in equation (14) is increasing in the dispersion of $\varepsilon(j)$. Technologies that greatly improve the (intangible) capital used only in some tasks but not others have a labor augmenting effect. Note that this term is over and beyond the standard Jensen's inequality effect that arises because we are approximating the growth in log wages: crucially it is a function of β, the coefficient capturing the decreasing returns to scale at the task level. Holding the elasticities η_c and η_o constant, a decrease in β (easier reallocation across tasks) implies a larger impact of the dispersion term on worker earnings. In general, as long as the elasticity of substitution ν is sufficiently high, both elasticities decrease with β implying the overall term rises as β falls.

The growth in employment is directly related to the growth in wages through the labor supply—equation (9). That is, the growth in employment in job (o, f) is equal to

$$\Delta_\varepsilon \log N(o,f) \approx \zeta \Big(\eta_o + \eta_c\,(J-1)\Big) m(\varepsilon) + \frac{\zeta}{2\,\beta}\,(\eta_o + \eta_c)^2\; V(\varepsilon) \tag{17}$$

The same forces that give rise to changes in earnings also directly lead to changes in employment growth.

We can also examine the impact of a change in firm productivity on worker's wage earnings and employment. In our model, we can express the productivity of firm f as

$$Z_f \equiv \left(\int_O X(o,f)^{\chi-1}\right)^{\frac{1}{\chi-1}}, \tag{18}$$

where $X(o, f)$ is the productivity of occupation o employed in firm f. If a given technology affects the productivity of multiple occupations in firm f at the same time, then this is an additional force that will affect wage earnings that is not captured by the two elasticities η_o and η_c above. After log-linearizing around the symmetric equilibrium, we can calculate the elasticity of task prices $w(j)$ to an increase in firm productivity,

$$\eta_z \equiv \frac{\partial \log w(j)}{\partial \log Z_f} = \frac{\theta - \chi}{s_k\,\nu + s_l\,\chi + \zeta} \tag{19}$$

As long as the elasticity of substitution across firm products is greater than the elasticity of substitution across occupations within a given firm—a reasonable assumption—then an increase in firm productivity leads to an increase in the wage of each task, which directly translates into higher wage earnings for workers on the job,

$$\Delta_z \log W(o,f) \approx \eta_z\,\Delta_z \log Z_f \tag{20}$$

and a corresponding increase in employment

$$\Delta_z \log N(o, f) \approx \zeta\, \eta_z\, \Delta_z \log Z_f \tag{21}$$

Last, the model has a direct implication for how task-level hours respond. In particular, hours growth is approximately proportional to the task-specific cost shock minus the occupational average cost shock:

$$\Delta \log h(j) \approx \frac{\eta_o - \eta_c}{b}\, (\varepsilon(j) - m(\varepsilon)) \tag{22}$$

Putting everything together, the effect of technology improvements on firm labor demand can be summarized by equations (17) and (21). The first equation derives the impact of a given technology ε on labor demand holding firm productivity constant, whereas the second equation illustrates the effect of changes in firm productivity holding everything else constant. Furthermore, equation (22) shows that at the task-level, the amount of effort allocated to a particular task within an occupation should be affected by how exposed the task is relative to the average exposure across all tasks.

2 Measurement

Our model in the previous section implies that the impact of a specific technology on firm labor demand for a specific occupation can be summarized by the two statistics in equations (15) and (16): the average improvement in the labor-saving technology across the tasks performed by workers in that job, and the dispersion in these improvements. In this section, we construct empirical analogues of these objects in the data.

2.1 Data

Our primary dataset comes from Revelio labs, a leading workforce database provider that has collected the near-universe of Linkedin Profiles as well as job postings, which we link to publicly traded firms in Compustat. The resume data includes comprehensive details on individuals' educational and employment histories, including universities attended, fields of study, employers, job titles, employment dates, and self-reported descriptions of jobs. The job postings data includes company name and other identifiers, job taxonomies (e.g., SOC, NAICS), posting dates, removal dates, seniority, salary (as estimated by Revelio if missing), job location (state, city, MSA, and ZIP code), full posting text, and, if the firm is public, relevant stock market identifiers (e.g., CUSIP).

We focus our main analysis on job positions that were active between 2014 and 2023, the most recent full decade for which we have data, a period marked by both improved resume coverage and the rising prevalence of artificial intelligence in the workforce. However, when constructing our instrument, we also incorporate job positions dating back to 2011. For the job postings data, we include records from 2010 to 2023, leveraging earlier years to construct our measures of task reallocation, as detailed in Section 3.4. To reduce the size of the raw job postings dataset, which exceeds a billion records before filtering, we randomly sample up to 10 postings per year for each occupation-firm pair. We further restrict our analysis to publicly traded companies, allowing us to merge labor market data with firm-level financial and performance metrics.

For tagging AI applications, we limit our analysis to resume job positions with a valid job description and a U.S. location, ensuring thath positions descriptions are in English. To count employment in non-AI positions within a firm, we additionally require a valid occupation identifier. In both cases, positions must be linked to a gvkey in Compustat. A position is classified as 'active' in a given year if the recorded start and end dates indicate that the worker held the position for at least six months within that year. After applying these restrictions, our dataset includes approximately 58 million LinkedIn profiles and 14 million job postings that are linked to Compustat firms and meet the sampling criteria.

Additionally we supplement our main data with O*NET and Compustat. We use O*NET's (Occupational Information Network), a comprehensive database developed by the U.S. Department of Labor that provides detailed descriptions of occupations, including the specific tasks, skills, and knowledge required for each job. We use Compustat in order to identify publicly traded firms and other key financial information used in the analysis.

2.2 Extracting AI applications from resumes and linking to worker tasks

To identify which firms are using AI and for what purposes, we first examine job descriptions where workers explicitly mention AI usage. We provide a brief overview of this process here and detail it further in Appendix A.2. In the initial step, we search for AI-related keywords within job position descriptions in resumes from the Revelio database to identify roles where AI may be utilized. We then employ a large language model (Llama 3.1 70B) to refine and extract specific phrases within job postings that describe how AI is being used. This process involves filtering job positions through multiple queries to the language model to systematically tag and clean AI-related information.

The first query identifies and cleans relevant phrases. Since our focus is on *how* AI is being used rather than the specific tools being mentioned, the second query removes references to particular AI tools from the extracted text. It also filters out vague AI applications that do not specify a concrete use case (e.g., a phrase like 'AI tools are being used to deploy computer vision models' lacks details on the intended purpose of the models and is therefore discarded). A final, third query eliminates any remaining non-specific AI applications that were not successfully filtered in the second step. This process yields over 1 million distinct AI use cases derived from approximately 500,000 job positions.

After identifying AI use cases, we next measure the extent to which non-AI workers' tasks are exposed to these applications. To do this, we use text embeddings, which encode the semantic meaning of a document as a geometric vector representation. When generated from the same embeddings model, documents with similar textual content will exhibit high cosine similarity. Earlier word-specific embeddings models, such as GloVe and word2vec (Pennington, Socher, and Manning, 2014; Mikolov, Sutskever, Chen, Corrado, and Dean, 2013), have been widely used in economic research to measure document similarity (Seegmiller, Papanikolaou, and Schmidt, 2023; Kogan et al., 2023; Autor et al., 2024), but these models assign a fixed vector to each word regardless of context. Recent advancements in embeddings technology allow for models that capture the full contextual meaning of text. We utilize the GTE-Large embeddings model, developed by Alibaba DAMO Academy, which encodes text into a 1096-dimensional vector and has demonstrated strong performance across a range of document similarity tasks compared to other models of similar scale.

Using the GTE embeddings, we represent each of the approximately 1 million cleaned AI application as a 1096-dimensional vector, and we do the same thing for each of the roughly 20,000 job tasks in the O*NET database by representing them by a 1096-dimensional vector. We impose that the vast majority of AI application/occupation task pairs should be unrelated. We only consider a task exposed to a particular AI application if the cosine similarity of the two text embeddings are above the unconditional 95th percentile in the overall task–AI application distribution of similarity scores.

Figure 1 illustrates this process with a specific example from the resume of a worker in an AI-implementing role at JP Morgan. The worker describes their position as follows:

> Technology delivery lead for risk and fraud forecasting models in auto, card, and home lending businesses. AI/ML model delivery in public cloud, private cloud and on prem. managing credit risk deployment services platform with continuous delivery, development and deployment of quantitative risk models that serve regulatory and credit risk assessments.

From this description, we first tag the terms "AI/ML" as AI-related. Our LLM queries identify "Technology delivery lead for risk and fraud forecasting models in auto, card, and home lending businesses" and "development and deployment of quantitative risk models that serve regulatory and credit risk assessments" as phrases which encode distinct ways the worker uses AI. Our LLM queries then clean and process the AI applications. The first AI-related phrase becomes the following concrete AI application: "Forecast risk and fraud in various lending businesses, including auto, card, and home lending"; the second phrase becomes "Assess credit risk and provide regulatory compliance across different lines of business."

Our GTE embeddings model identifies tasks performed by credit analysts as being particularly exposed to both of these AI applications: for the former example, among all 20,000 tasks in the O*NET database, the task with the highest cosine similarity is "Prepare reports that include the degree of risk involved in extending credit or lending money." Similarly, the task "Analyze credit data and financial statements to determine the degree of risk involved in extending credit or lending money" is the most related of all tasks to the latter AI application.

2.3 Constructing the AI Exposure Measures

Here, we detail the construction of our key measures of AI exposure at the occupation–firm level.

Occupation-Level Exposure to AI: Mean vs. Dispersion

We now outline our approach for tagging AI applications and exposed tasks to align with model-consistent definitions of occupational AI exposure. We define an occupation using the detailed 6-digit SOC codes. While the O*NET database includes approximately 800 such codes, in our dataset, Revelio assigns resumes to 335 distinct 6-digit occupation codes, resulting in a slightly more aggregated classification in practice.

Let j represent tasks, o occupations, f firms, and i an AI application identified from resumes. We classify an AI application as being in use at firm f during a given year t if the corresponding job position on the resume was active for at least six months within that year. We denote $N_{f,t}$ as the number of AI applications in use at firm f during year t. We then define the probability that a given occupation is exposed to an AI application in use at firm f at time t as follows:

$$\text{Exposure Probability}_{j,f,t} = \frac{1}{N_{f,t}} \sum_{i=1}^{N_{f,t}} I_{j,i}^{95} \tag{23}$$

The variable $I_{j,i}^{95}$ takes a value of 1 if the cosine similarity of the GTE text embedding for the task j and AI application i are above the 95th percentile in the distribution across all potential (i,j) pairs. We then compute the weighted average exposure probability across an occupation's tasks:

$$\mu_{o,f,t} = \sum_{j} \omega_{o,j} \text{ Exposure Probability}_{j,f,t} \tag{24}$$

The weights $\omega_{o,j}$ are taken from the O*NET task importance scores. In particular, O*NET assigns a score between 1 and 5 on each task performed by a specific occupation to indicate how central is the task as part of the job. We take these weights $\omega_{o,j}$ from these importance score scores and rescale them so that they sum to one within each occupation.

How does average AI exposure probability vary across wage ranks? In Figure 3, we rank occupation–firm pairs in the Revelio data based on their average salary as imputed by Revelio. We then plot the average of $\mu_{o,f,t}$ across employment-weighted salary percentile ranks. Consistent with Webb (2020), we find that AI exposure rises with wage rank up to approximately the 90th percentile, after which it declines. This pattern contrasts with previous technological shifts, which primarily affected middle-skill labor (Autor et al., 2006; Kogan et al., 2023).

Given $\mu_{o,f,t}$, we can likewise define the dispersion of AI exposure across tasks as:

$$\sigma^2_{o,f,t} = \sum_j \omega_{o,j} \left(\text{Exposure Probability}_{j,f,t} - \mu_{o,f,t}\right)^2 \tag{25}$$

The measures $\mu_{o,f,t}$ and $\sigma^2_{o,f,t}$ capture the mean and dispersion, respectively, of the likelihood that a given task within an occupation is exposed to a given AI application at firm f and time t. To take into account that some firms may utilize AI applications more intensively than others, we adjust these measures with an estimate of the AI intensity in firm f. Accordingly, our exposure measures are given by

$$\text{AI Exposure Average}_{o,f,t} = \mu_{o,f,t} \times \log(1 + N_{f,t}) \tag{26}$$

$$\text{AI Exposure Dispersion}_{o,f,t} = \sigma^2_{o,f,t} \times \log(1 + N_{f,t}) \tag{27}$$

Equations (26) and (27) are the empirical equivalents of equations (15) and (16) in the model. Importantly, the vary across firms not only because different firms develop different AI applications, but also because different firms adopt AI more intensively than others—a proxy for the size of the ε shocks in the model.[3]

Our model implies that an increase in (26) will lead to a decrease employment *within* the firm, while holding (26) constant, an increase in (27) will lead to higher employment within the firm. The first prediction stems from the direct task-level substitution of labor with AI, while the latter comes from cross-task productivity spillovers generated by the cost improvements coming from AI applications targeted towards the occupation's tasks.

Firm-Level AI Utilization

In the model, the firm-specific productivity index Z_f is increasing in the average occupational productivity $X(o, f)$, or the inverse cost of one unit of occupational output. Since the number of AI

[3] We multiply by the log of one plus, rather than incorporating $N_{f,t}$ directly into the definition of (23), because the distribution of the number $N_{f,t}$ of AI applications is highly skewed within firms—some firms have AI uses in the several thousands while others have very few. Our specification implicitly takes a stand on the mapping between the number of AI applications and the decline ε in the quality-adjusted price of labor-saving technologies.

uses at the firm level affects the exposure of each occupation, it consequently raises this productivity index. Therefore, we use the uninteracted firm-level measure $\log(1+N_{f,t})$ as a predictor of firm-wide AI-induced cost improvements, and use it to proxy for $\Delta \log Z_f$. However, because this index is likely to be mechanically larger for bigger firms or those with more comprehensive resume coverage in the Revelio data, we always include a control for the log total number of Revelio employees at the firm in year t in any specification that includes $\log(1 + N_{f,t})$ as an independent variable.

3 The Impact of AI on Labor Demand

We now present our empirical findings. Identifying the causal impact of AI adoption on firm-level and labor market outcomes is challenging, as firms endogenously decide whether to adopt AI. To address this, we develop an identification strategy that leverages pre-existing differences in firms' hiring networks.

3.1 Instrumental Variables

AI adoption at the firm level is inherently endogenous, as the same factors that drive a firm's decision to implement AI also shape its economic performance. Larger, more productive, and more profitable firms are more likely to adopt AI, as we later show, raising concerns about selection bias since these firms tend to follow distinct growth trajectories.Specifically, such firms tend to have *lower* growth, naturally (Evans, 1987). As a result, firms with both observable and unobservable characteristics that increase their likelihood of adopting AI, may follow different growth trajectories for reasons unrelated to AI adoption, potentially exhibiting lower growth rates. Second, unobserved heterogeneity may bias estimates if firms with stronger management practices or more advanced technological infrastructure are both more likely to adopt AI and better positioned to benefit from it. This would likely result in an upward bias in the estimated effect of AI adoption.

Together, these issues highlight the need for an instrumental variables approach to isolate exogenous variation in AI adoption to establish a causal relationship with firm outcomes. Our strategy for identifying the effects of firm-level AI use on firm outcomes exploits variation in AI adoption across firms driven by hiring spillovers from AI-intensive firms from pre-existing hiring networks across firms among non-AI employees. Specifically, we construct an instrumental variables (IV) strategy that leverages variation in AI exposure at the occupation level across firms and a shift-share approach to predict AI employment at the firm level.

The concerns from measuring AI exposure within firm but across occupations are somewhat different. First of all, some selection bias potentially arises from firms selectively implementing AI in occupations or tasks where the expected cost savings are highest (for where firms expect increased labor scarcity or diminished labor productivity), leading to a non-random distribution of AI adoption across firms. Such selection would tend to bias our estimates of the effect of average AI exposure in a particular manner: it would naturally induce a negative bias in the coefficients on both the mean and the dispersion of AI exposure, meaning that after instrumenting, the IV coefficients on average AI exposure should become less more negative and the dispersion of AI exposure should be more positive relative to the OLS estimates. Another possibility is that there is some measurement error in occupational AI exposure within firms due to limited data on actual AI implementation, which can attenuate estimates of AI's impact on occupational employment towards zero. Given the highly granular nature of our measurement and the relatively few AI uses within firms we often have to compute our measures, there is potential for the presence of such attenuation. Our IV strategy for our occupation-level exposure measures leverages variation in the mean and dispersion of occupational exposure to AI across *all* firms to predict the mean and dispersion of AI exposure within the firm. By using information on AI exposure across many AI applications outside the firm, this simultaneously helps mitigate potential bias due to both measurement error and the targeted adoption of AI toward certain occupations but within the firm.

The key variation in our IV strategy for the effects of AI use comes from labor mobility spillovers: firms that have previously hired non–AI employees from AI-intensive firms are more likely to be able to adopt AI themselves because of labor market interconnectedness with other AI–using firms. This allows us to capture AI adoption patterns that are driven by external AI labor supply exposure rather than purely selection-driven internal firm characteristics. We predict the number of AI workers at the firm as follows:

$$N_{f,t}^{AI,pred} = N_{f,t}^{total} \times p_{f,t}^{AI} \tag{28}$$

Here $N_{f,t}^{w,total}$ is the total number of resumes in active job positions at firm f at time t, while $p_{f,t}^{AI}$ is the predicted probability of a given worker implementing AI at the firm. Our estimate of $p_{f,t}^{AI}$ is

$$p_{f,t}^{AI} = \sum_{f' \in \mathcal{F}_{f,t}} w_{f' \to f,t} \times \frac{N_{f',t-1}^{AI}}{N_{f',t-1}^{total}} \tag{29}$$

Here $N_{f',t-1}^{AI}$ is the actual number of AI workers in firm f and year $t-1$, and $N_{f,t}^{w,total}$ is the number of total workers in firm f' year t; $\mathcal{F}$ is the set of all firms that firm f hires from over the year $t-3$ to t. We further denote by $w_{f' \to f,t}$ the share of all firm f's non-AI hires from $t-3$ to $t-1$ which came from f'. We compute $w_{f' \to f,t}$ using within-individual transitions in job positions represented in Revelio resumes.

Figure 2 shows the top 10 firms in the hiring network of UnitedHealth in 2018, which is based off their outside hiring from 2015 to 2017. From 2015 to 2017, UnitedHealth's top 10 sources for external hires in non-AI roles included firms such as McKesson Corporation, Aetna, Wells Fargo, Accenture, and Target. Our IV estimates the extent to which UnitedHealth's AI adoption is influenced by the share of AI workers at these firms. This approach assumes that UnitedHealth's established hiring networks for non-AI workers serve as a proxy for the availability of AI talent from these same sources. However, these hiring patterns are assumed to be independent of UnitedHealth's direct incentives or capacity to invest in AI.

We ultimately define our instrument for the number of AI uses, $\log(1 + N_{f,t})$, as:

$$\log(1 + N_{f,t})^{IV} \equiv \log(1 + N_{f,t}^{AI,pred}) \tag{30}$$

which operates under the plausible assumption that the number of ways a firm can use AI is increasing in the (predicted) number of workers implementing AI within the firm. This predicted AI employment variable captures exogenous shifts in AI adoption based on pre-existing labor market networks that are plausibly unrelated to technology adoption considerations. Because firms do not directly control other firms' AI employment, this instrument reduces endogeneity concerns while. In the next section we show the instrument has a high F-stat for predicting actual AI utilization, so it also satisfies the relevance condition.

Next, we compute the occupation-level instruments. We let N_t denote the number of AI applications extracted across *all* firms in year t, and we define

$$\text{Exposure Probability}_{j,t} = \frac{1}{N_t} \sum_{i=1}^{N_{f,t}} \text{Above 95th percentile indicator}_{j,i} \tag{31}$$

and

$$\mu_{o,t} = \sum_j \omega_{o,j} \text{Exposure Probability}_{j,t} \tag{32}$$

$$\sigma_{o,t}^2 = \sum_j \omega_{o,j} \left(\text{Exposure Probability}_{j,t} - \mu_{o,t}\right)^2 \tag{33}$$

which are respectively the task-level exposure probabilities, and the means and dispersions of the exposure probabilities across all firms in year t. Finally, we arrive at our instruments for the mean and dispersion of occupational task exposure to AI:

$$\text{AI Exposure Average}\ _{o,f,t}^{IV} \equiv \mu_{o,t} \times \log(1 + \text{Predicted AI Employees})_{f,t} \tag{34}$$

$$\text{AI Exposure Dispersion}_{o,f,t}^{IV} \equiv \sigma_{o,t}^2 \times \log(1 + \text{Predicted AI Employees})_{f,t} \tag{35}$$

The logic behind these instruments is that the average and dispersion of AI exposure across all firms within an occupation instead reflects broader technological diffusion of AI towards certain tasks, rather than any selected targeting of particular occupations for purely firm-specific reasons. Additionally, by computing the means and dispersions across a broad set of firms, the estimation noise in the mean and dispersion go down considerably, which helps with attenuation concerns. As we demonstrate in 3.3, this attenuation appears especially impactful for the coefficient estiumates on the AI exposure dispersion.

3.2 Firm-Level Effects

Before studying the occupation-level implications of our measurement, we first briefly examine what types of firms tend to use AI and its impact on subsequent firm outcomes in the 2014-2023 period that is the focus of our analysis. In Table 1 we examine the relationship between $\log(1+N_{f,t})$ and log sales per worker; log sales; log profits (defined as sales minus cost of goods sold); log TFP (which we estimate for Compustat firms following the procedure outlined in İmrohoroğlu and Şelale Tüzel (2014)); and finally, the log average salary for each position at the firm based off Revelio–imputed salaries to see which types of firm tend to use AI technologies. We control for 3-digit NAICS industry by year fixed effects and the log of total Revelio employee resumes in each specification because of the potentially mechanical relationship between the number of resumes we can observe and the number of distinct AI uses we find at the firm level. We cluster standard errors by firm.

In Table 1 we find a strongly positive relationship between the extent of firms' AI utilization and productivity measures like log sales per worker or TFP, as well sales, profits, and average pay. We scale the main independent variable to unit standard deviation. We find that a standard deviation increase in AI utilization is associated with roughly 12% increases in both productivity measures, while sales and profits increase by about 28% and 38%, respectively; the marginal effect on log average salary to a standard deviation increase in AI use is about 10%. This table illustrates that AI–using firms are quite different from other firms, being larger, more productive, and higher paying. This echoes survey evidence on the cross-sectional relationship between advanced technology adoption and firm characteristics (Acemoglu, Anderson, Beede, Buffington, Childress, Dinlersoz, Foster, Goldschlag, Haltiwanger, Kroff, Restrepo, and Zolas, 2023b; McElheran, Li, Brynjolfsson, Kroff, Dinlersoz, Foster, and Zolas, 2023). We emphasize that these relationships are only correlational, and could reflect both causal relationships as well as selection into which types of firms are able to take advantage of AI.

With this in mind, we now look at the impact of current AI use on subsequent firm outcomes. In particular, we estimate regressions of the form

$$\log Y_{f,t+5} - \log Y_{f,t} = \beta \log(1+N_{f,t}) + \alpha_{Ind,t} + \delta X_{f,t} + \epsilon_{f,t} \tag{36}$$

In our controls, we include the log of Compustat employment and the log total number of resumes in the Revelio database to net out the association between AI use driven by the and the. Because AI–adopting firms may also be more innovative and subsequently grow as a result, we control for the market value of innovation scaled by book value of assets Kogan, Papanikolaou, Seru, and Stoffman (2017). As noted by Kogan et al. (2017) firms with different characteristics may mechanically growth faster or slower, and given the correlation with firm characteristics uncovered in Table 1, we follow Kogan et al. (2017) by controlling for the lagged level of $\log Y$, as well its the lagged growth rate. Finally, we add industry $\times$ year fixed effects to account for industry–specific trends, and we again cluster standard errors at the firm level.

We present our estimates of equation (36) in Table 2. We find a strong positive link between the number of the AI uses in firm f in year t and the subsequent growth in firm sales, employment, profits, and total factor productivity over the next five years. In the first four columns we focus on OLS specifications, while in the last four columns we report the estimates from two-stage least squares, in which we estimate (36) by instrumenting for the number of AI applications in firm f in year t with our instrument based on pre-existing hired networks described in equation (30) above.

Examining Table 2, several findings stand out. First, all coefficient estimates are strongly positive and statistically significant across both the OLS and IV specifications. To the extent that our instrument is valid, these results imply a significantly positive impact of AI adoption on firm growth. These strong positive relations are consistent with our interpretation of the number of AI applications as a shifter in the firm productivity index Z_f in the model. By providing benefits to the firm in the form of higher cost efficiency, AI utilization causes firms to expand and be more productive.

Second, our IV estimates are quantitatively larger than our OLS specifications. This fact would arise if firms which tend to adopt AI also are the types of firms which tend to grow more slowly, as it would tend to bias the OLS estimates towards zero. This possibility is consistent with survey evidence (Acemoglu et al., 2023b) that shows that advanced technology adopters tend to be larger and older to begin with, meaning they are likely to naturally be in the "low–growth" phase of their life-cycle (Evans, 1987). Our instrumental variables strategy therefore tends to mitigate this downward bias. Also of note are the large F-statistics on the instrument in each of the IV specifications, alleviating any potential concerns related to weak instruments.

Third, we find sizeable effects of AI adoption on future firm growth. Focusing on the IV specifications, a standard deviation increase in the extent of AI use at the firm is associated with 9.5% higher sales growth, 6% higher employment growth, 8.5% profit growth, and 7.6% higher revenue TFP growth. In the case of employment growth, we emphasize that this feature of the data should not be interpreted as evidence that AI is directly complementary to worker tasks.

Through the lens of the model, the impact of firm productivity on labor demand can be large when the elasticity of substitution between firms (θ in the model) is much higher than the elasticity of substitution between occupations within firms (χ in the model), generating positive net labor demand effects even when the task-level AI–labor substitution ν is large. To infer the degree of labor substitution, we need to examine more granular evidence at the occupation–firm level, which is the focus of the next section.

3.3 Effects on Occupational Labor Demand

Having established that AI has a causal impact on firm productivity and growth, we next turn to the main part of our analysis that focuses on labor demand for specific jobs. To do so, we leverage the granularity of our exposure measures in (26) and (27) that establish which occupations are more or less exposed to AI applications within a given firm.

Our main outcome variable is employment of a particular occupation in a specific firm at a point in time. We measure employment at the 6-digit SOC occupation level within a firm using the Revelio resume data, which allows us to keep track of how many positions were active at the given firm and occupation each year. We use the weights provided by Revelio to reweigh observations to account for the skew in the data towards certain occupations. We consider a position active in a given year if it was in place for at least half of the year. Our model focuses on the labor-substituting effects of AI at the task level. Therefore, when measuring employment and the occupation–firm level, we exclude occupations from the 2-digit SOC code 15 (Computer and Mathematical occupations) because these are by far the most likely broad occupation group to be tagged as AI implementers, representing about 50% of all AI positions in our data. Moreover, even if not directly running the AI models, workers in these occupations are also those who would implement and maintain the software and hardware necessary to use AI at the firm.

Our first specification exploits both within- and between-firm variation. Specifically, we estimate

$$\begin{aligned}\log(Emp_{f,o,t+5}) - \log(Emp_{f,o,t}) = {} & \beta \text{AI Exposure Average}_{f,o,t} + \delta \text{AI Exposure Dispersion}_{f,o,t} \\ & + \gamma \log(1 + N_{f,t}) + \Gamma X_{f,o,t} + \alpha_t \epsilon_{f,t}\end{aligned} \tag{37}$$

In this specification, we include only calendar year fixed effects and directly control for firm-level AI utilization $\log(1+N)_{f,t}$, which allows us to examine the direct effects of shifting firm productivity Z_f in the model. In this specification, the vector of controls $X_{f,o,t}$ includes the log of firm employment (based on Revelio and Compustat); the growth in employment over the last year for that particular occupation in that specific firm; and the market value of innovation to book assets from Kogan et al. (2017), similar to equation (36) above. We cluster standard errors by occupation–firm, as this is the level at which our main measures vary, and we weight observations by the share of yearly total employment in the occupation–firm cell. We scale the independent variables to have unit standard deviation, and we multiply the dependent variable by 100 so that coefficients read as the percent increase in employment growth for a standard deviation change in exposure.

Our second specification focuses on within-firm patterns, by differencing out shocks to the firm,

$$\begin{aligned}\log(Emp_{f,o,t+5}) - \log(Emp_{f,o,t}) = \beta \text{AI Exposure Average}_{f,o,t} + \delta \text{AI Exposure Dispersion}_{f,o,t} \\ + \Gamma X_{f,o,t} + \alpha_{f,t} + (\alpha_{o,t}) + \epsilon_{f,t}\end{aligned} \tag{38}$$

Specifically, in equation (38) we include the interaction of firm–year fixed effects $\alpha_{f,t}$ to the specification in (37). The cost of doing so is that we can no longer identify the coefficient γ capturing the impact of firm productivity improvements due to AI on labor demand. The advantage, however, is that by leveraging cross-occupation and within-firm variation, this specification nets out all endogeneity driven by general firm-level selection into AI adoption, and isolates the relative demand effects stemming from differences in occupational exposure within the firm. As a variant of this specification, we also include granular occupation (6-digit SOC) $\times$ year fixed effects $\alpha_{o,t}$. This variant further absorbs occupation–specific trends by exploiting variation in the degree to which a particular occupation is more exposed to AI within a given firm compared to the typical exposure of that occupation to AI among all firms. This specification is feasible due to the granular nature of our measure compared to existing measures that identify only cross-occupation differences in exposure (Webb, 2020; Eisfeldt et al., 2023; Eloundou et al., 2023; Brynjolfsson et al., 2018).

We report our estimates in Table 3. In columns 1 through 3 we report the estimated coefficients from (37) and (38) using OLS. In columns 4 through 6 we estimate the same specifications through two-stage least squares using the instrumental variables described in section 3.1.

Confirming the predictions of the model, the estimated coefficients on $\widehat{\beta}$ on AI Exposure Average$_{f,o,t}$ is negatively and highly significant, while the estimated coefficient $\widehat{\delta}$ on AI Exposure Dispersion$_{f,o,t}$ is also positive and highly significant across specifications. When we do not include firm–year fixed effects in columns 1 and 4, we can separately estimate the impact of firm-level AI induced productivity gains, which also has highly significantly positive relationship with employment growth. Coefficient estimates are reasonably stable when comparing within OLS or within IV specifications, though they are unsurprisingly the smallest when we include occupation–year fixed effects in the most strict specifications in columns 3 and 6. Instrumented specifications again have F-statistics that are far beyond typical weak instruments thresholds, which diminishes any such concerns.

In terms of magnitudes, there are several points worth noting. The range of coefficient estimates on the average AI exposure double from between negative 5 and 8 percent for the OLS specifications, to minus 10 to 16 percent for the IV specifications. This pattern is even more stark for the dispersion of AI exposure, which increases by 5 to 10 times when comparing IV to OLS specifications, going from a significantly positive but small 1 to 1.6 percent in OLS specs versus a large and highly significant positive 7 to 10 percent across IV specs. If the primary source of endogeneity were selective adoption within firms towards occupations and tasks where labor scarcity is expected to increase or labor productivity to decline, we would then expect to see a positive bias in both OLS coefficients: the coefficient on AI Exposure Average$_{f,o,t}$ would become less negative and go towards zero after being instrumented, while AI Exposure Dispersion$_{f,o,t}$ would become more positive. Instead, both coefficient estimates move away from zero and in opposite directions after being instrumented, which is more consistent with measurement noise causing attenuation in our estimates (especially for $\sigma^2_{f,o,t}$). While we cannot rule out that there is some targeted AI adoption within the firm, it doesn't seem to be the primary source of bias in the OLS.

The negative effect stemming from average AI exposure is directly in line with the model's predictions under the assumption of labor–AI substitution at the task level. The positive coefficient on AI Exposure Dispersion$_{f,o,t}$ suggests that there are indeed productivity spillovers across tasks within an occupation which mitigate some of this negative direct substitution effect. The model allows for either force to dominate the other depending on the parameter mix, most especially the various elasticities of substitution. In the data we find that this direct substitution effect appears to be quantitatively larger. On the other hand, the firm-level productivity effect of AI use raises employment in general. Thus whether AI use on net erodes or increases relative employment for particular groups is an empirical question that depends on the empirical distribution of each measure, as well as the relative coefficient estimates.

3.4 Task Reallocation

A key channel through which improvements in AI technologies affect workers is through the changing nature of a job. In the model, this is reflected through the endogenous allocation of effort across tasks performed by an occupation. Recalling equation (22) in the model, the change in effort devoted to task j is proportional to the task-specific technology improvement $\varepsilon(j)$ relative to the occupational average mean exposure $m(\varepsilon)$. The hours response to AI exposure will be negative when the own elasticity $\eta_o < 0$, which occurs when the labor–AI capital elasticity substitution ν is sufficiently large, as discussed in section 1.

Thus, to study how AI changes the nature of a job, the right object to focus on is the extent to which *specific tasks* are exposed to AI applications within a given firm, after netting out the occupation-level average exposure has been netted out with occupation–firm fixed effects. We construct a proxy for the the intensity of task utilization using the composition of skill demands in firms' online job postings. We obtain texts of online job postings for occupations from Revelio, and we tag each job posting with the list of skills the job posting requires using the Open Skills API provided by LightCast. A growing literature uses the LightCast (formerly Burning Glass) lists of job posting skills.[4] After tagging the skills in each job posting using the LightCast API, the average job posting in our dataset lists 17 distinct skills.

The Revelio job posting texts are linked to the firm that posted the job and the relevant occupation. Additionally, LightCast provides a textual description of each skill, which allows us to textually link the skills with associated occupation job tasks in O*NET. LightCast identifies approximately 30,000 distinct skills, and we impose sparsity in the textual linkages by connecting a skill as applicable to a given task if the cosine similarity of the task and skill description's GTE embeddings are in the top percentile of the distribution across all pairs, implying that the typical task has around 300 associated skills that could be applicable when performing the task. We then examine how the share of total skills demanded across job postings listed by firm f for occupation o are linked to task j. Namely, we compute

$$Share_{j,o,f,\tau} = \frac{\text{\# of skills linked to task } j \text{ in occupation } o \text{ job postings at firm } f \text{ over time period } \tau}{\text{Total \# of skills in occupation } o \text{ job postings at firm } f \text{ over time period } \tau}$$

For a given year t, we take τ to be the 5-year window from $t-4$ through t inclusive. We then look at the Davis, Haltiwanger, and Schuh (1996) changes in these shares between consecutive 5 year periods. The Davis et al. (1996) (DHS) change is a second-order approximation to the log change, so coefficients can be interpreted in units of percentage changes, but it also accommodates cases where one of the shares is equal to zero. Formally, the DHS change is

$$\Delta_{DHS} Share_{j,o,f,\tau+1} = \frac{Share_{j,o,f,\tau+1} - Share_{j,o,f,\tau}}{0.5 \times (Share_{j,o,f,\tau+1} + Share_{j,o,\tau})} \tag{39}$$

Here τ denotes the 5-year period that includes the years $t-4$ through t, inclusive. We use $\Delta_{DHS} Share_{j,o,f,\tau+1}$ to approximate the growth in the importance of task j to occupation o at firm f. Note that this measure now varies at the *task*–by–firm level, rather than the occupation–by–firm level as before. In a parallel manner, we construct a task–level exposure measure; since our unit of analysis is now at the task–level within an occupation, we only need to include the task–level exposure (proxy $\epsilon(j)$ in the model). To be consistent with our definition of occupation–level average exposure in (26), we allow task–level exposure to be a function of how likely a task is to be exposed

[4] A few examples include Deming and Kahn (2018), Deming and Noray (2020), Acemoglu et al. (2022), and Braxton and Taska (2023).

to a given AI application at the firm, multiplied by how intensively the firm utilizes AI:

$$\text{Task-Level AI Exposure}_{j,f,t} = \text{Exposure Probability}_{j,f,t} \times \log(1+N)_{f,t} \tag{40}$$

Where Exposure $\text{Probability}_{j,f,t}$ is defined in (23). Similarly, our task–level instrument now becomes the likelihood of task j being exposed to an AI application from any firm, and interacted with the predicted number of AI employees:

$$\text{Task-Level AI Exposure}^{IV}_{j,f,t} = \text{Exposure Probability}_{j,t} \times \log(1 + \text{Predicted AI Employees})_{f,t} \tag{41}$$

With these in hand, we estimate the following specification.

$$\Delta_{DHS} Share_{j,o,f,\tau+1} = \beta \text{Task-Level AI Exposure}_{j,f,t} + \alpha_{o,f,t} + \delta X_{j,o,f,t} + \epsilon_{j,o,f,t} \tag{42}$$

The controls X include the task importance weight $\omega_{o,j}$ derived from O*NET task importance scores as explained in section 2.3, which we take to be a rough notion of how much effort someone in this occupation would typically to allocate to this task on average. In specifications without the full complement of occupation × firm × year fixed effects, we also control for the occupational average AI exposure, since (7) makes clear that hours growth is proportional to task-level exposure relative to the occupational average. We cluster standard errors at the occupation–firm level, and weight each observation by the number of workers in the occupation–firm cell multiplied by the task importance weight. We scale Task-Level AI $\text{Exposure}_{j,f,t}$ to unit standard deviation, and we multiply the dependent variable to 100 so that coefficient estimates can be interpreted as the percentage response to a standard deviation increase in task–level AI exposure.

Table 5 reports our estimates of equation (42). The first three columns are estimated via ordinary least squares and the last three are estimated via two-stage least squares with Task-Level AI $\text{Exposure}^{IV}_{j,f,t}$ above as the IV. In the first and fourth columns we only include form × year fixed effects; in the second and fifth columns we include firm × year and occupation × year fixed effects separately. In the third and sixth columns we instead control for the firm × occupation × year fixed effects to fully leverage the within occupation–firm variation that our task–level analysis allows.

Examining the Table, we see that our point estimates imply that an increase in the task–level AI exposure reduces the intensity at which firms demand skills related to that task. These estimates are consistently negative across the OLS and IV specifications, and the magnitudes are largely comparable across different combinations of fixed effects. The magnitudes are also substantial: focusing on column (6) which reports the estimates from our preferred specification (IV with the full complement of fixed effects), we see that a one standard deviation increase in the task–level

AI exposure reduces the intensity at which firms demand skills related to that task by around 4.5 percent. Given that the dependent variable corresponds to a share, these estimates validate the model's implication that labor effort reallocates away from skills that are relatively more exposed to AI within an occupation towards skills that are less exposed.

3.5 Aggregate Effects

How important is the development of AI technologies for labor demand? Answering this question in the aggregate using our point estimates from the previous section is not directly feasible given the presence of calendar year fixed effects in our specifications. These fixed effects allow us to isolate the impact of AI from other economic forces during the same period that may have impacted labor markets, such as the recovery from the 2008/09 financial crisis or the Tax Cuts and Jobs Act of 2017. However, our estimates can be used to understand AI's reallocative impact—its impact on the employment shares of affected occupations.

We begin by computing the expected total net marginal effect of AI use on labor demand as a function of different job characteristics K—average occupation wages or broad occupation groups,

$$E\big[\, \widehat{\beta}\text{AI Exposure Average}^{\perp} + \widehat{\delta}\text{AI Exposure Dispersion}^{\perp} + \widehat{\gamma}\log(1+N_{f,t})^{\perp} \mid K = k\big] \tag{43}$$

Here the estimated coefficients $\widehat{\beta}$, $\widehat{\delta}$, and $\widehat{\gamma}$ are taken from the IV estimates in column (4) of Table 3. The notation $\perp$ denotes that variables have been orthogonalized with respect to all non-AI controls, and K is the characteristic of interest (salary rank or occupation broad category). Because the average of each variable has been netted out via controls, these marginal effects on employment are relative to the average employment growth, so the total effects integrated across the entire employment share–weighted distribution of K integrate to zero by construction.

We first focus on how the effects vary across occupations with different earnings levels. To do so, we compute the expected marginal effect at each salary percentile—i.e. we take k to correspond to the occupation's salary percentile and then compute (43) separately for percentile—and we plot the results from computing (43) in the top panel of Figure 4. The red line gives the expected relative marginal effect on employment from just the mean exposure component (the first term in (43)); the green line gives the effect of the reallocative component (the second term in (43)); and the yellow line shows the effect of firm productivity on labor demand (the third term in (43)). Last, the blue line shows the total net effect from all three components take together.

Examining first on the direct effect of labor substitution—the red line in the top panel of Figure 4—we see that the negative effect of mean AI exposure increasing on average as pay increases. This result echoes the fact that the probability of AI exposure is higher for higher-paid occupations in Figure 3. In terms of magnitudes, jobs that are the most highly-paid income percentiles are expected to decline by about 7% in aggregate employment share over a 5-year period due to this component of exposure, while the least highly-paid are expected to increase by about 9%.

However, this mean exposure measure abstracts from the benefits of labor reallocation and the impact of firm productivity on firm labor demand. Examining the remaining lines in the Figure, we note that these two effects serve to mediate the direct effect of labor substitution. In particular, highly-paid positions tend to be at firms which use AI intensively (yellow line), and are also exposed to larger dispersion in AI exposure (green line). These two forces almost offset the impact of direct task-level AI substitution, such that the total net effect in blue is close to zero across the bulk of the pay distribution. Interestingly, at the very top of the pay distribution we find a slightly *positive* net impact on aggregate employment, despite the relatively high exposure to labor task-substituting AI technologies. This decomposition illustrates why it can be hard to detect the impact of AI on employment across demographic categories that are differentially exposed to AI. For example, Acemoglu et al. (2022) find little to no effect of AI on aggregate employment from 2010-18, despite evidence that it affects labor demand for specific establishments.

This decomposition uses our estimates from equation (37) that exploits both within- as well as between-firm variation in AI exposure. Our conclusions remain similar if instead we focus on the within-firm estimates from equation (38). Hence we now instead implement the decomposition (43) using our within-firm specification from column 5 of Table 3. As we see in panel B of Figure 4, this within-firm decomposition highlights the relatively stronger direct substitution effect of average AI exposure, which generates within-firm employment declines among the highly-paid positions that are most exposed. The within-firm negative net effect is largest at about the 90th percentile of the salary distribution, but the direct effect entails only about a 2.3% decline in within-firm employment; including the countervailing influence of AI exposure dispersion reduces the total effect to about 1.5%. Meanwhile, the total positive effect of relatively low AI exposure at the bottom of the salary distribution is a modest 2% increase in within-firm employment. We conclude that, despite the strong negative substitution and the much higher exposure of highly-paid occupations, the net impacts of AI on both aggregate and within-firm relatively employment reallocation across the pay distribution are detectable but quantitatively small.

We next examine how these effects vary across broad occupation categories—defined at the 2-digit SOC code level. In Table 4 we report each component of the total effect of AI on employment shares across each broad occupation group. As before, these marginal effects represent predicted changes in employment shares, and therefore the employment-weighted total effects sum to 0 within each given component of exposure.

Examining Table 4, we see that several occupations groups experienced significant declines in their employment shares as a result of their exposure to AI. For instance, our estimates imply that the "Business and Financial" and "Architecture and Engineering" broad groups experience the largest AI-related employment declines on average over a 5-year period, at -1.9% and 2.6%, respectively. Interestingly, the business and financial occupations have much higher direct exposure (mean AI exposure component of -10% versus minus 6% for business and financial versus architecture and engineering respectively), but they benefit from being employed at firms with much higher average AI utilization (firm component of positive 2% versus 0.5% for architecture and engineering). The total effect is that architecture and engineering occupations have had the largest relative employment loss due to AI use.

Another interesting case are "Food preparation and serving" occupations. These occupations also experience declining employment shares because of AI, but this is driven entirely by the rarity of AI use their employers: the net employment impact of a is minus 2%, but the firm-specific component is minus 7.7%, meaning their employers do not have the ability to take advantage of AI-driven productivity improvements. Meanwhile the occupation group who has the most from AI are legal occupations, who have both benefitted from low occupational exposure to AI and have high firm utilization. Our estimates predict a net increase of 6.4% in the employment of these occupations. With this in mind, we note that our estimates are driven by the pre-Generative AI era of artificial intelligence. There is some evidence that large language models may have begun to more directly expose legal occupations in very recent years (Eloundou et al., 2023). Overall, for the most part, the aggregate effects of AI on employment by broad occupation group are again somewhat muted, albeit with larger variation than when aggregated by occupation pay levels.

Importantly, these muted broad occupation effects do however mask considerable within-group variation. To see this, we next examine the impact of AI on employment shares at a more granular level (at the 6-digit SOC level). In Figure 5, we compare the effect of AI on employment shares implied by our estimates to the actual changes in employment share experienced by these occupations during our period. In both cases, we remove the impact of the controls in column 4 of Table 3. In Panel A we focus only on the within-firm across-occupation component of AI-induced shifts in employment share—corresponding to the first two terms of equation (43). In Panel B, we focus only on the firm-level component—the last term in equation (43). Panel C plots the total effect.

Examining Figure 5, we see a strongly positive relation between the realized growth in employment shares across occupations and the AI-implied shift. Panels A and B illustrate that both the direct exposure of an occupation's tasks to AI and the average AI use of an occupation's employers are significant in accounting for the realized reallocation across occupations. Examining Panel C, we note that the AI-implied changes in employment shares can be quite significant once we narrow down our focus on specific occupations. Further, these AI-implied shifts can account for a meaningful share of the actual shifts in employment shares during this period: the regression R^2 with all components of AI exposure included is approximately equal to 14 percent; roughly 51 percent of this explanatory power is attributable to the component driven by occupation-specific task exposure, with the remainder being attributable to the average exposure of occupations to their employers' overall AI use.

Overall, we conclude that advances in Artificial Intelligence was an economically meaningful driver of shifts in employment shares across occupations over the last decade. That said, we should note that our analysis is skewed towards publicly traded firms, that tend to be larger. Given that larger firms tend to use AI more intensively, these magnitudes may overestimate the impact of AI in the broader economy.

4 Conclusion

This paper examines the impact of artificial intelligence adoption on firm dynamics and labor demand, leveraging firm-occupation level variation in AI exposure. We document three primary empirical patterns. First, AI adoption is concentrated in larger, more productive firms, which tend to have distinct growth trajectories. Instrumental variable estimates confirm that AI adoption leads to higher firm-level sales, profits, and total factor productivity. Second, at the occupational level, AI exposure is concentrated in higher-wage positions, with employment effects that depend on the dispersion of exposure across tasks. Higher average exposure reduces within-firm employment, while greater dispersion in exposure mitigates these declines by reallocating labor toward complementary tasks. Third, firm-wide AI adoption generates positive employment effects, consistent with AI-driven productivity gains increasing aggregate labor demand.

The results suggest that while AI substitutes for labor at the task level, its net employment effects are shaped by offsetting forces. Highly AI-exposed occupations experience declines in labor demand, yet within-occupation task reallocation and firm-wide AI-driven growth help sustain overall employment levels. Consistent with our conceptual framework, we find that occupations with high dispersion in AI exposure experience relatively higher employment growth, underscoring the importance of within-occupation task complementarities. Across firms, AI adoption is associated with higher employment growth in AI-intensive firms, indicating that firms integrating AI more effectively also expand their workforce.

Despite these countervailing forces, the labor market effects of AI remain nontrivial. Our estimates suggest that AI exposure accounts for roughly 14% of the variation in occupational employment growth among publicly traded firms, with half of this effect stemming from task-level substitution and the remainder from firm-wide adoption. However, at the top of the wage distribution, where AI exposure is most pronounced, we find limited net employment effects due to offsetting firm growth and reallocation within occupations. These results highlight why aggregate employment impacts of AI may be difficult to detect, even as task-level substitution is significant.

Taken together, these findings provide new evidence on the labor market implications of AI adoption. Rather than leading to broad-based job losses, AI appears to be reallocating labor across tasks and firms, with the magnitude of displacement effects contingent on the structure of AI adoption at the firm level. Future research should further investigate how firms adjust their workforce composition in response to AI, how skill demands evolve, and how AI-induced productivity gains translate into wage and employment dynamics over time.

References

Acemoglu, D. (2024, May). The simple macroeconomics of ai. Working Paper 32487, National Bureau of Economic Research.

Acemoglu, D., G. Anderson, D. Beede, C. Buffington, E. Childress, E. Dinlersoz, L. Foster, N. Goldschlag, J. Haltiwanger, Z. Kroff, P. Restrepo, and N. Zolas (2023a, May). Advanced technology adoption: Selection or causal effects? *AEA Papers and Proceedings 113*, 210–14.

Acemoglu, D., G. Anderson, D. Beede, C. Buffington, E. Childress, E. Dinlersoz, L. Foster, N. Goldschlag, J. Haltiwanger, Z. Kroff, P. Restrepo, and N. Zolas (2023b, May). Advanced technology adoption: Selection or causal effects? *AEA Papers and Proceedings 113*, 210–14.

Acemoglu, D. and D. Autor (2011). Skills, tasks and technologies: Implications for employment and earnings. In *Handbook of Labor Economics, Volume 4. Amsterdam: Elsevier-North*, pp. 1043–1171.

Acemoglu, D., D. Autor, J. Hazell, and P. Restrepo (2022). Artificial intelligence and jobs: Evidence from online vacancies. *Journal of Labor Economics 40*(S1), S293–S340.

Acemoglu, D. and P. Restrepo (2018, June). The race between man and machine: Implications of technology for growth, factor shares, and employment. *American Economic Review 108*(6), 1488–1542.

Acemoglu, D. and P. Restrepo (2021, June). Tasks, automation, and the rise in us wage inequality. Working Paper 28920, National Bureau of Economic Research.

Auer, R., D. Köpfer, and J. Sveda (2024, September). The rise of generative AI: modelling exposure, substitution and inequality effects on the US labour market. BIS Working Papers 1207, Bank for International Settlements.

Autor, D., C. Chin, A. Salomons, and B. Seegmiller (2024, 03). New frontiers: The origins and content of new work, 1940–2018*. *The Quarterly Journal of Economics 139*(3), 1399–1465.

Autor, D. H. and D. Dorn (2013, August). The growth of low-skill service jobs and the polarization of the us labor market. *American Economic Review 103*(5), 1553–97.

Autor, D. H., L. F. Katz, and M. S. Kearney (2006, May). The Polarization of the U.S. Labor Market. *American Economic Review 96*(2), 189–194.

Babina, T., A. Fedyk, A. He, and J. Hodson (2024). Artificial intelligence, firm growth, and product innovation. *Journal of Financial Economics 151*, 103745.

Babina, T., A. Fedyk, A. X. He, and J. Hodson (2023, June). Firm investments in artificial intelligence technologies and changes in workforce composition. Working Paper 31325, National Bureau of Economic Research.

Benmelech, E. and M. Zator (2022, January). Robots and firm investment. Working Paper 29676, National Bureau of Economic Research.

Bonfiglioli, A., R. Crinò, H. Fadinger, and G. Gancia (2020, November). Robot Imports and Firm-Level Outcomes. Crc tr 224 discussion paper series, University of Bonn and University of Mannheim, Germany.

Braxton, J. C. and B. Taska (2023, February). Technological change and the consequences of job loss. *American Economic Review 113*(2), 279–316.

Brynjolfsson, E., T. Mitchell, and D. Rock (2018, May). What can machines learn, and what does it mean for occupations and the economy? *AEA Papers and Proceedings 108*, 43–47.

Caunedo, J., D. Jaume, and E. Keller (2023, June). Occupational exposure to capital-embodied technical change. *American Economic Review 113*(6), 1642–85.

Dauth, W., S. Findeisen, J. Suedekum, and N. Woessner (2021, 05). The Adjustment of Labor Markets to Robots. *Journal of the European Economic Association 19*(6), 3104–3153.

Davis, S. J., J. C. Haltiwanger, and S. Schuh (1996). *Job Creation and Destruction.* MIT Press.

de Souza, G. and H. Li (2023, November). Robots, Tools, and Jobs: Evidence from Brazilian Labor Markets. Working Paper Series WP 2023-42, Federal Reserve Bank of Chicago.

Dechezleprêtre, A., D. Hémous, M. Olsen, and C. Zanella (2021). Induced automation: evidence from firm-level patent data. *University of Zurich, Department of Economics, Working Paper* (384).

Deming, D. and L. B. Kahn (2018). Skill requirements across firms and labor markets: Evidence from job postings for professionals. *Journal of Labor Economics 36*(S1), S337–S369.

Deming, D. J. and K. Noray (2020). Earnings dynamics, changing job skills, and stem careers. *The Quarterly Journal of Economics 135*(4), 1965–2005.

Eaton, J. and S. Kortum (2002). Technology, geography, and trade. *Econometrica 70*(5), 1741–1779.

Eisfeldt, A., G. Schubert, B. Taska, and M. B. Zhang (2023). Generative ai and firm values. Working paper, U.S. Census Bureau.

Eloundou, T., S. Manning, P. Mishkin, and D. Rock (2023). Gpts are gpts: An early look at the labor market impact potential of large language models.

Evans, D. S. (1987). Tests of alternative theories of firm growth. *Journal of Political Economy 95*(4), 657–674.

Felten, E. W., M. Raj, and R. Seamans (2018, May). A method to link advances in artificial intelligence to occupational abilities. *AEA Papers and Proceedings 108*, 54–57.

Gathmann, C., F. Grimm, and E. Winkler (2024, December). AI, Task Changes in Jobs, and Worker Reallocation. IZA Discussion Papers 17554, Institute of Labor Economics (IZA).

Goldsmith-Pinkham, P., I. Sorkin, and H. Swift (2020, August). Bartik instruments: What, when, why, and how. *American Economic Review 110*(8), 2586–2624.

Goos, M., A. Manning, and A. Salomons (2014, August). Explaining job polarization: Routine-biased technological change and offshoring. *American Economic Review 104*(8), 2509–26.

Graetz, G. and G. Michaels (2018, 12). Robots at Work. *The Review of Economics and Statistics 100*(5), 753–768.

Humlum, A. (2019). Robot adoption and labor market dynamics. *Princeton University*.

Humlum, A. and E. Vestergaard (2024, May). The Adoption of ChatGPT. IZA Discussion Papers 16992, Institute of Labor Economics (IZA).

Koch, M., I. Manuylov, and M. Smolka (2021, 01). Robots and Firms. *The Economic Journal 131*(638), 2553–2584.

Kogan, L., D. Papanikolaou, L. D. Schmidt, and B. Seegmiller (2023, November). Technology and labor displacement: Evidence from linking patents with worker-level data. Working Paper 31846, National Bureau of Economic Research.

Kogan, L., D. Papanikolaou, A. Seru, and N. Stoffman (2017, 03). Technological Innovation, Resource Allocation, and Growth. *The Quarterly Journal of Economics 132*(2), 665–712.

Mann, K. and L. Püttmann (2023, 05). Benign Effects of Automation: New Evidence from Patent Texts. *The Review of Economics and Statistics 105*(3), 562–579.

McElheran, K., J. F. Li, E. Brynjolfsson, Z. Kroff, E. Dinlersoz, L. S. Foster, and N. Zolas (2023, October). Ai adoption in america: Who, what, and where. Working Paper 31788, National Bureau of Economic Research.

Mikolov, T., I. Sutskever, K. Chen, G. Corrado, and J. Dean (2013). Distributed representations of words and phrases and their compositionality. *CoRR abs/1310.4546*.

Pennington, J., R. Socher, and C. D. Manning (2014). Glove: Global vectors for word representation. In *EMNLP*.

Seegmiller, B., D. Papanikolaou, and L. D. Schmidt (2023). Measuring document similarity with weighted averages of word embeddings. *Explorations in Economic History 87*(C), S0014498322000729.

Webb, M. (2020). The impact of artificial intelligence on the labor market.

İmrohoroğlu, A. and Şelale Tüzel (2014). Firm-level productivity, risk, and return. *Management Science 60*(8), 2073–2090.

Figures

Figure 1: Illustration of process for identifying AI applications from resumes and exposed occupation tasks

Example resume job description of a worker employed at JP Morgan:

> Technology delivery lead for risk and fraud forecasting models in auto, card, and home lending businesses. AI/ML model delivery in public cloud, private cloud and on prem. managing credit risk deployment services platform with continuous delivery, development and deployment of quantitative risk models that serve regulatory and credit risk assessments.

Step 1: Identify AI-related terms (if any)

"Technology delivery lead for risk and fraud forecasting models in auto, card, and home lending businesses. **AI/ML** model delivery in public cloud, private cloud and on prem. managing credit risk deployment services platform with continuous delivery, development and deployment of quantitative risk models that serve regulatory and credit risk assessments."

Step 2: Use large language models to extract the phrases likely to contain specific AI applications

"**Technology delivery lead for risk and fraud forecasting models in auto, card, and home lending businesses.** AI/ML model delivery in public cloud, private cloud and on prem. Managing credit risk deployment services platform with continuous delivery, **development and deployment of quantitative risk models that serve regulatory and credit risk assessments.**"

Step 3: Use large language models to clean the extracted AI applications

Extracted phrase: "Technology delivery lead for risk and fraud forecasting models in auto, card, and home lending businesses."

> Cleaned AI application: **"Forecast risk and fraud in various lending businesses, including auto, card, and home lending."**

Extracted phrase: "Development and deployment of quantitative risk models that serve regulatory and credit risk assessments."

> Cleaned AI application: **"Assess credit risk and provide regulatory compliance across different lines of business."**

Step 4: Use GTE sentence embeddings to measure textual similarity to identify highly exposed tasks

AI application: "Forecast risk and fraud in various lending businesses, including auto, card, and home lending."

> Most exposed O*NET occupation task by cosine similarity: **"Prepare reports that include the degree of risk involved in extending credit or lending money."** (Credit Analysts, SOC code = 132041)

AI application: "Development and deployment of quantitative risk models that serve regulatory and credit risk assessments."

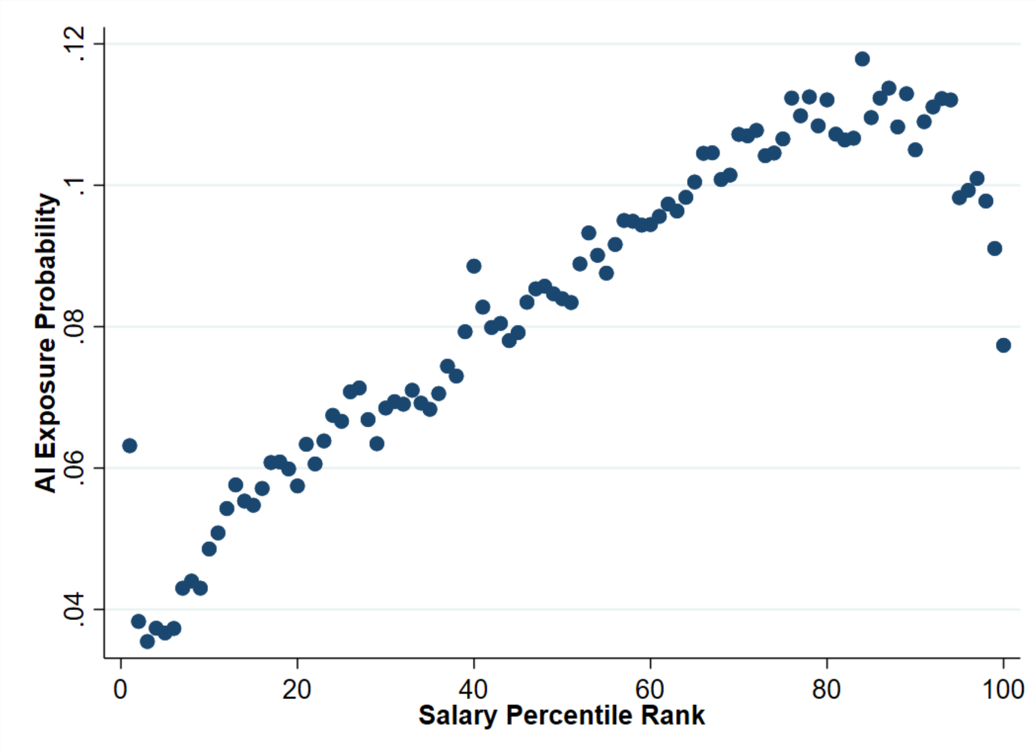
> Most exposed O*NET occupation task by cosine similarity: **"Analyze credit data and financial statements to determine the degree of risk involved in extending credit or lending money."** (Credit Analysts, SOC code = 132041)

Note: This figure shows an example of the process for identifying AI applications from online resumes and linking with exposed job tasks. See section 2 in the main text and appendix A.2 for further details.

Figure 2: Example of a pre-existing hiring network

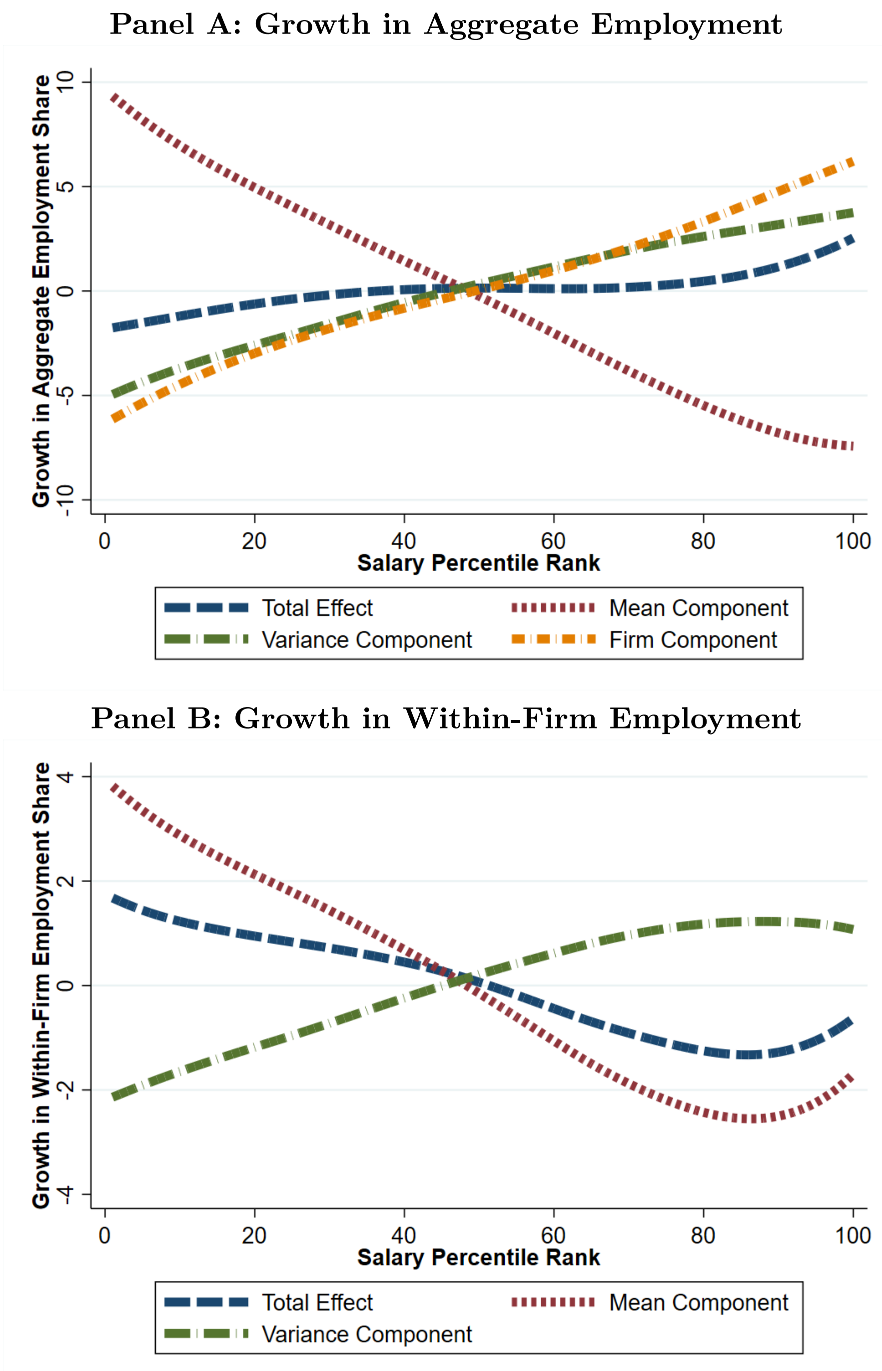
Note: This figure shows the top 10 firms in UnitedHealth's hiring network in 2018. Each percentage in parentheses corresponds to the share of UnitedHealth's total hires from 2015-2017 coming from the given firm. See section 3.1 for details.

Figure 3: AI Exposure Probability by Salary Rank



Note: This figure plots the average task-level probability of exposure to a given AI application by salary rank. We use imputed salaries for each job position to compute the ranks. See Section 2.3 for details.

Figure 4: Impact of AI on employment growth across the pay distribution



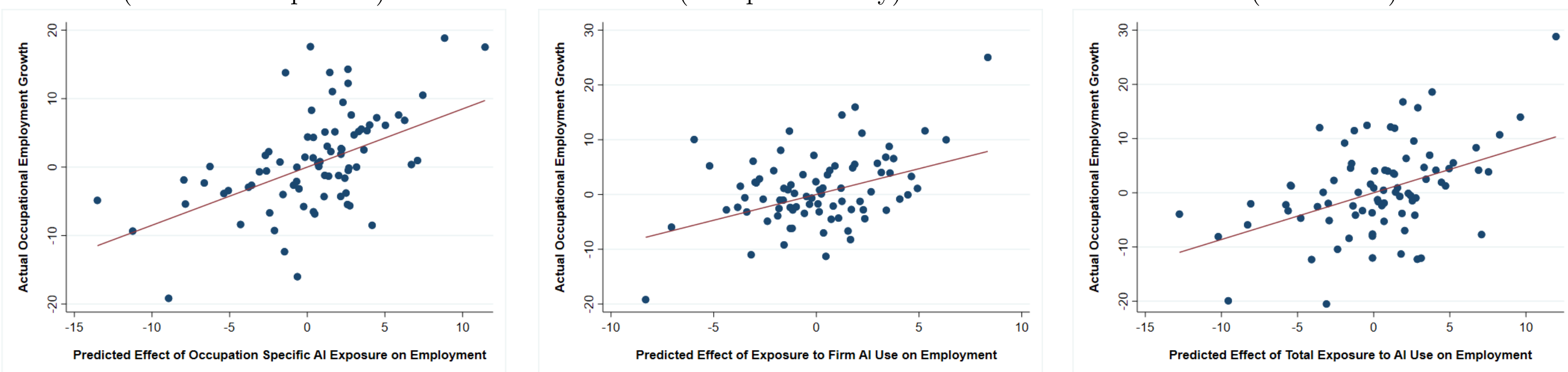
Note: This figure implements the decomposition of employment marginal effects from equation (43) in Section 3.5, where we compute the expected impact of the different components of exposure to artificial intelligence on changes in employment share at each points in the salary percentile distribution. Plots are lowess-smoothed to enhance readability. In Panel A, we plot the impacts on employment shares in the aggregate, while in Panel B, we look purely at within-firm reallocation. In red, we plot the impact of direct task-level substitution driven by our measure of average AI exposure; in green we plot the impact of across-task productivity spillovers driven by the variance of AI exposure within the occupation. In Panel A we also show the effect of firm-level AI use in yellow. The total net effect is in blue. See section 3.3 of the main text for details.

Figure 5: Actual growth in employment shares relative to AI-implied growth

A. Effect from Direct Task Exposure to AI (mean and dispersion)

B. Effect from Employers' AI Use (firm productivity)

C. Overall AI-Implied Change (total effect)



Note: This figure plots residualized binscatter plots of actual occupational employment growth against predicted occupational employment growth implied by our decomposition of employment marginal effects from equation (43) in Section 3.5. We implement the decomposition at the 6-digit SOC occupation level. In Panel A, we plot the relationship between actual occupational employment growth and the total effect of direct occupation task exposure to AI (including mean and variance of task exposure), after netting out the firm-level component. In Panel B, we do the opposite exercise by plotting the partial relationship between actual occupational employment growth and the occupational average exposure to firm-level AI use, after netting out the effects of direct task exposure. The two components taken together can explain 14 percent of realized employment growth, of which 51 percent is attributable to direct occupation exposure. See section 3.3 of the main text for details.

Tables

Table 1: Characteristics of AI-using firms

	(1) Log Sales per worker	(2) Log Sales	(3) Log Profit	(4) Log TFP	(5) Log Average Salary
log(1 + AI uses)	0.108***	0.284***	0.382***	0.117***	0.0990***
	(6.98)	(12.52)	(17.71)	(10.67)	(18.87)
N	33541	36227	33309	17034	38211
R-sq	0.345	0.644	0.614	0.183	0.427
Revelio Emp Control	X	X	X	X	X
Ind × Year FE	X	X	X	X	X

Note: This table shows regression coefficients of the logs of sales per worker, total sales, profits (defined as sales minus cost of goods sold), revenue total factor productivity, and log average Revelio salary on firm-level AI utilization log(1+Number of AI Uses)$_{f,t}$ defined in Section 2 of the main text. As controls, we include the log of total employment based on Revelio resume counts in the given year and 3-digit NAICS industry × year fixed effects. We cluster standard errors by firm and report t-statistics in parenthesis. The sample period spans 2014-2023.

Table 2: The impact of firm-level AI use on firm growth rates

Dependent variable: 100× 5-year growth rate in the firm outcome designated in each column

	OLS				IV			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	Sales	Emp	Profit	TFP	Sales	Emp	Profit	TFP
$\log(1 + \text{AI uses})$	6.06***	4.33***	6.51***	5.47***	9.47***	6.03***	8.53***	7.60***
	(4.11)	(4.01)	(4.75)	(5.87)	(4.54)	(4.10)	(4.10)	(5.94)
N	12757	13225	11652	6065	12282	12688	11246	6035
R-sq	0.13	0.12	0.12	0.23	0.080	0.054	0.028	0.18
F-stat					5567.4	5946.5	4879.6	2240.9
Controls	X	X	X	X	X	X	X	X
Ind × Year FE	X	X	X	X	X	X	X	X

Note: This table shows results from estimating Equation (36). The dependent variable is the 5-year forward growth rate in the designated firm outcome. In the last four columns, we estimate the specification using twos-stage least squares with the instrument $\log(1 + \text{Number of AI Uses})^{IV}_{f,t}$ IV defined in Section 3.1 of the main text, with corresponding IV F-statistics from the first-stage regression reported in the table. Controls include a lagged one-year growth rate and level of the dependent variable, yearly patent market value to assets from Kogan et al. (2017); the logs of total employment both based on Revelio resume counts and Compustat employment counts; and 3-digit NAICS × year fixed effects. We cluster standard errors by firm and report corresponding t-statistics in parentheses.

Table 3: AI exposure and occupational employment growth (5-year horizon)

Dependent variable: 100× 5-year growth rate in the occupation–firm employment

		OLS			IV	
	(1)	(2)	(3)	(4)	(5)	(6)
AI Exposure Average	-7.73***	-7.54***	-5.27***	-15.9***	-14.3***	-10.2***
	(-13.56)	(-12.46)	(-10.05)	(-11.53)	(-16.95)	(-11.45)
AI Exposure Variance	1.00**	1.64***	1.18***	9.92***	7.63***	7.25***
	(2.81)	(4.34)	(4.08)	(7.41)	(8.31)	(5.73)
$\log(1 + \text{AI uses})$	9.12***			15.0***		
	(15.50)			(17.30)		
N	2037346	2035179	2035179	2017810	2017168	2017168
R^2	0.13	0.54	0.60	0.11	0.017	-0.0024
F-stat (AI Exposure Average)				789.4	2752.8	1412.7
F-stat (AI Exposure Variance)				690.6	1561.4	494.9
F-stat ($\log(1 + \text{AI uses})$)				6966.5		
Controls	X	X	X	X	X	X
Year FE	X			X		
Firm × Year FE		X	X		X	X
Occ × Year FE			X			X

Note: This table shows regression estimates of Equation (17) from the main text. Columns (1), (2), and (3) correspond to the OLS estimates. Columns (4), (5), and (6) correspond to the two-stage least squares with the set of instruments described in Section 3.1 of the main text. We include associated F-statistics for each instrumented variable in the table. In addition to the designated fixed effects, all specifications include a control for the lagged one-year employment growth; specifications with only year fixed effects additionally control for yearly patent market value to assets from Kogan et al. (2017); the logs of total employment both based on both Revelio resume counts and Compustat employment counts. Observations are weighted by the yearly occupation-firm cell's share of employment. Standard errors clustered by occupation–firm are in parentheses.

Table 4: Impact of AI on relative employment growth by occupation group

	2-digit SOC	Mean Component	Variance Component	Firm Component	Total	% of Emp
Management	11	-2.27	1.55	0.78	0.057	19.0
Business and Financial	13	-10.1	6.18	2.04	-1.92	17.6
Architecture and Engineering	17	-5.96	2.82	0.51	-2.63	9.10
Science	19	1.60	-0.018	0.10	1.68	2.36
Community and Social Service	21	10.8	-5.76	0.30	5.32	0.33
Legal	23	10.0	-6.17	2.56	6.42	0.71
Education and Library	25	9.47	-5.03	0.072	4.51	1.00
Arts, Entertainment, Media	27	7.99	-4.82	2.09	5.26	5.38
Healthcare Practitioners	29	5.77	-2.63	-0.54	2.60	1.92
Healthcare Support	31	7.59	-3.95	0.42	4.06	0.47
Protective Service	33	9.37	-5.87	-1.46	2.05	0.43
Food Preparation and Serving	35	12.7	-7.02	-7.70	-1.99	2.75
Cleaning and Maintenance	37	14.5	-8.80	-3.37	2.28	0.46
Personal Care and Service	39	12.5	-6.81	-3.66	1.98	1.09
Sales and Related	41	1.47	-0.73	-1.60	-0.86	13.3
Office and Administrative	43	2.71	-2.45	0.61	0.87	10.6
Farming, Fishing, and Forestry	45	13.4	-7.76	-3.79	1.81	0.46
Construction and Extraction	47	6.41	-4.30	-0.44	1.67	2.07
Installation and Repair	49	4.03	-3.33	-0.99	-0.29	2.72
Production	51	5.80	-2.58	-2.40	0.82	3.94
Transportation	53	7.92	-4.47	-2.57	0.88	4.26

Note: This table shows results from estimating the decomposition (43) for broad 2-digit SOC occupation groups. The column "Mean Component" provides the average relative employment growth impact of the average task-level exposure to AI within the occupation group, while the "Variance Component" shows the impact of the variance in task-level exposure to AI. The "Firm-Component" column gives the employment impact of the occupation groups' average exposure to firm-level AI use. Effects are expressed relative to the aggregate average growth rate, so that the total employment-weighted effect sums to 0. See Section 3.3 of the text for further details.

Table 5: AI and Task Reallocation

Dependent Variable: 100 × 5-year Davis et al. (1996) change in share of job posting skills related to task

	OLS			IV		
	(1)	(2)	(3)	(4)	(5)	(6)
Task-level AI Exposure	-4.71***	-4.68***	-4.73***	-4.36***	-3.99***	-4.52***
	(-13.40)	(-13.91)	(-14.08)	(-9.14)	(-10.04)	(-11.15)
N	13241933	13241933	13238128	13241933	13241933	13238128
R^2	0.073	0.10	0.32	0.0073	0.0033	0.0037
F-stat				13366.9	15875.4	18755.7
Task Importance Control	X	X	X	X	X	X
Mean Task Exposure Control	X	X		X	X	
Firm × Year FE	X	X		X	X	
Occ × Year FE		X			X	
Firm × Occ × Year FE			X			X

Note: This table includes estimates of Equation (42) from the main text. The unit of analysis is at the task–firm level, and the dependent variable is the Davis et al. (1996) change in the share of job posting skills demanded that are textually linked to the give task when comparing across the firms' job postings for a given occupation in the next 5 years versus the previous 5 years. The "Task-AI Exposure" is defined in equation 40 of the main text. In Columns (1)-(3), we estimate the specification using OLS, and in Columns (4)-(6), we use the instrument defined in main text Equation (41), with associated F-statistics reported below. Standard errors are clustered by occupation-firm, with associated t-statistics reported in parentheses. Besides the designated fixed and also the average task-level AI exposure of the occupation in specifications without the full complement of firm × occupation × year fixed effects. See section 3.4 of the main text for details.

A Appendix

Section A.1 contains the derivations of the model.

A.1 Model Appendix

Here, we provide more details on the model derivation.

Optimal Hours Allocation

The optimal time allocation problem has the solution

$$h(j) = \frac{w(j)^{\frac{1}{\beta}}}{\sum_{j\in J} w(j)^{\frac{1}{\beta}}}. \tag{A.1}$$

To see this, note that each worker solves the following optimization problem, taking into account the constraint on hours,

$$\mathcal{L} = \sum_{j=1}^{J} w(j)h(j)^{1-\beta}\, dj - \lambda \left(\sum_{j=1}^{J} h(j)\, dj - 1 \right). \tag{A.2}$$

The first-order condition with respect to devoted to task j $h(j)$ is

$$(1-\beta)\, w(j)\, h(j)^{-\beta} = \lambda. \tag{A.3}$$

This leads to

$$h(j) = \left[(1-\beta)\frac{w(j)}{\lambda} \right]^{\frac{1}{\beta}}. \tag{A.4}$$

Take sum on both sides,

$$\sum_{j=1}^{J} h(j) = \sum_{j=1}^{J} \left[(1-\beta)\frac{w(j)}{\lambda} \right]^{\frac{1}{\beta}} dj = (1-\beta)^{-\frac{1}{\beta}} \lambda^{\frac{1}{\beta}} \sum_{j=1}^{J} w(j)^{\frac{1}{\beta}} = 1. \tag{A.5}$$

Thus,

$$\lambda = (1-\beta) \left(\sum_{j\in J} w(j)^{\frac{1}{\beta}} \right). \tag{A.6}$$

Apply this to (A.4) yields (A.1).

Equilibrium First-Order Conditions for Labor and Capital Demand

The firm-level cost minimization problem can be expressed as

$$\min_{Y_f(o)} \int_{\mathcal{O}} P(o,f) Y_f(o) \quad \text{s.t. } Y_f = \left(\int_{\mathcal{O}} Y_f(o)^{\frac{\chi-1}{\chi}} \right)^{\frac{\chi}{\chi-1}}. \tag{A.7}$$

Given the above, the labor demand of firm f for occupation o is equal to

$$Y(o,f) = P(o,f)^{-\chi}\, Z_f^{-\chi}\, Y_f \tag{A.8}$$

where

$$Z_f \equiv \left(\int_O P(o,f)^{1-\chi} \right)^{-\frac{1}{1-\chi}} \tag{A.9}$$

and $P(o,f)$ denotes the marginal cost firm f it pays for the output of occupation o. These prices need not be the same across firms. To simplify notation, in what follows we will suppress the firm subscripts. Firms make profits because of imperfect competition, reflecting both pricing power in product markets and monopsony power in labor markets. Denote their markup over marginal cost Z_f^{-1} by $\Theta = \frac{\theta}{\theta-1} > 1$. Since the firm has monopsony power in the labor market, its marginal cost will exceed its average cost, as we discuss further below. As a result,

$$\begin{aligned} P_f Y_f &= \Theta \int_O Y_f(o) P(o,f) \\ P_f Y_f &= \Theta \int_O P(o,f)^{1-\chi} \left(\int_O P(o,f)^{1-\chi} \right)^{\frac{\chi}{1-\chi}} Y_f \\ P_f &= \Theta \left(\int_O P(o,f)^{1-\chi} \right)^{\frac{1}{1-\chi}} = \Theta\, Z_f^{-1} \end{aligned} \tag{A.10}$$

Each firm faces the inverse demand curve

$$Y_f = P_f^{-\theta}\, P^{\theta}\, \bar{Y} \tag{A.11}$$

where

$$P \equiv \left(\int_{\mathcal{F}} P_f^{1-\theta} \right)^{\frac{1}{1-\theta}}. \tag{A.12}$$

Without loss of generality, we can normalize the aggregate price index $P = 1$, which implies

$$Y_f = P_f^{-\theta}\, \bar{Y}, \tag{A.13}$$

and given the price above, this implies

$$Y_f = \Theta^{-\theta}\, Z_f^{\theta}\, \bar{Y} \tag{A.14}$$

Now, consider the occupation's task minimization problem

$$\min_{y(j)} \sum_{j \in J} p(j)\, y(j) \qquad s.t.\;\; Y(o,f) = \left(\sum_{j \in J} y(j)^{\frac{\psi-1}{\psi}} \right)^{\frac{\psi}{\psi-1}} \tag{A.15}$$

Here $p(j)$ is the marginal cost index of producing task j output $y(j)$ after optimal input choices have been made within task j. Due to monopsony power, the marginal cost $p(j)$ will exceed the average cost of producing $y(j)$ given that the firm will internalize that hiring a marginal worker will require paying higher wages to additional, inframarginal workers. Our CES structure admits the following Hicksian demand for $y(j)$ from the FOC for problem (A.15):

$$y(j) = p(j)^{-\psi} \left[\sum_{j \in J} p(j)^{1-\psi} \right]^{\frac{\psi}{1-\psi}} Y(o, f) \tag{A.16}$$

Here, $P(o, f)$ is the marginal cost of occupation o's output.

$$P(o, f) = \left[\sum_{j \in J} p(j)^{1-\psi} \right]^{\frac{1}{1-\psi}} \tag{A.17}$$

Using the above combined with equations (A.8) and (A.14), we get

$$y(j) = \frac{1}{p(j)^{\psi}} X(o)^{\chi-\psi} \, Z_f^{\theta-\chi} \Theta^{-\theta} \, \bar{Y}. \tag{A.18}$$

where

$$X(o, f) \quad = \quad \left[\sum_{j \in J} p(j)^{1-\psi} \right]^{-\frac{1}{1-\psi}} = P(o, f)^{-1} \tag{A.19}$$

In the above, $X(o, f)$ is the productivity (the inverse of the unit cost) of productivity O and Z_f is the productivity of firm f.

The factor allocation associated with the monopsonistic cost minimization problem is isomorphic to the solution to a perfectly competitive firm which faces a wedge between the marginal cost of labor and the wage. We denote this wedge by $\mathcal{M}(j)$, which we compute below. In deriving comparative statics, we assume that the firm treats $\mathcal{M}(j)$ as constant when choosing its factor allocations for simplicity and analytical tractability. The cost minimization problem within task j is

$$\min_{l(j), k(j)} q(j) k(j) + w(j) \mathcal{M}(j) l(j) \qquad s.t. \;\; y(j) = \left[(1-\gamma_j) k(j)^{\frac{\nu-1}{\nu}} + \gamma_j \, l(j)^{\frac{\nu-1}{\nu}} \right]^{\frac{\nu}{\nu-1}} \tag{A.20}$$

Going forward, we make the re-parameterization $a_j \equiv \gamma_j^{\nu}$, $b_j \equiv (1-\gamma_j)^{\nu}$. After solving (A.20), the per-unit cost of task j equals

$$p(j) = \left(a_j \left[\mathcal{M}(j) w(j) \right]^{1-\nu} + b_j \; q(j)^{1-\nu} \right)^{\frac{1}{1-\nu}} \tag{A.21}$$

Using equation (A.21), we can rewrite (A.18) as

$$y(j) = \left(a_j \left[\mathcal{M}(j) w(j) \right]^{1-\nu} + b_j \; q(j)^{1-\nu} \right)^{\frac{-\psi}{1-\nu}} X(o, f)^{\chi-\psi} \, Z_f^{\theta-\chi} \Theta^{-\theta} \, \bar{Y}. \tag{A.22}$$

Plugging in (A.22) to the CES Hicksian demand for $k(j)$ and $l(j)$ and imposing labor market clearing gives

$$l(j) \quad = \quad \frac{a_j}{[\mathcal{M}(j)w(j)]^{\nu}} \left(a_j[\mathcal{M}(j)w(j)]^{1-\nu} + b_j\, q(j)^{1-\nu}\right)^{\frac{\nu-\psi}{1-\nu}} X(o)^{\chi-\psi}\, Z_f^{\theta-\chi} \Theta^{-\theta}\, \bar{Y}. \tag{A.23}$$

$$k(j) \quad = \quad \frac{\beta_j}{q(j)^{\nu}} \left(a_j[\mathcal{M}(j)w(j)]^{1-\nu} + b_j\, q(j)^{1-\nu}\right)^{\frac{\nu-\psi}{1-\nu}} X(o)^{\chi-\psi}\, Z_f^{\theta-\chi} \Theta^{-\theta}\, \bar{Y}. \tag{A.24}$$

Labor Market Clearing

If there are $N(o,f)$ workers in a occupation–firm pair (o,f) then the total supply of

$$L_o(j) = N(o,f)\, h^{1-\beta}(j). \tag{A.25}$$

Using the properties of the Fréchet distribution, it follows that the expected measure of workers to job o in firm f is equal to

$$N(o,f) = \underbrace{\frac{1}{\int_{f'\in\mathcal{F}} \int_{o'\in\mathcal{O}} W(o,f)^{\frac{\zeta}{1+\zeta}}}}_{\bar{\zeta}} W(o,f)^{\zeta}. \tag{A.26}$$

where $W(o,f)$ is the total earnings on the job,

$$W(o,f) \equiv \sum_{j\in J_o} h(j)^{1-b} w(j). \tag{A.27}$$

Given (A.1), the total earnings for that job are equal to

$$W(o,f) = \sum_{j\in J_o} h(j)^{1-\beta} w(j) = \frac{\sum_{j\in J} w(j)^{\frac{1}{\beta}}}{(\sum_{j\in J} w(j)^{\frac{1}{\beta}})^{1-\beta}} = \left[\sum_{j\in J} w(j)^{\frac{1}{\beta}}\right]^{\beta}. \tag{A.28}$$

Notice that as long as hours are flexible, $0 < b < 1$, then the occupation level wage is convex in the task prices. Put differently, because the worker can reallocate hours, she benefits from a mean-preserving spread in $w(j)$.

So, the total labor supply for task j is equal to

$$\begin{aligned} h(j)^{1-\beta}\, N(o,f) &= w(j)^{\frac{1-\beta}{\beta}} \left(\sum_{j\in J} w(j)^{\frac{1}{\beta}}\right)^{\beta-1} \left[\sum_{j\in J} w(j)^{\frac{1}{\beta}}\right]^{\zeta\beta} \bar{\zeta}. \\ &= w(j)^{\frac{1}{\beta}-1} \left(\sum_{j\in J} w(j)^{\frac{1}{\beta}}\right)^{\beta-1+\zeta\beta} \bar{\zeta} \end{aligned} \tag{A.29}$$

Replacing the left-hand-side of equation (A.23) with the equation for labor supply for task j

yields a system of J equations in J unkowns—the task prices $w(j)$,

$$w(j)^{\frac{1}{\beta}} \left(\sum_{j \in J_o} w(j)^{\frac{1}{\beta}} \right)^{\beta - 1 + \zeta\, \beta} \bar{\zeta} = a_j\, \mathcal{M}(j)^{-\nu}\, w(j)^{1-\nu} \left(a_j [\mathcal{M}(j) w(j)]^{1-\nu} + b_j\, q(j)^{1-\nu} \right)^{\frac{\nu-\psi}{1-\nu}} X(o,f)^{\chi-\psi}\, Z_f^{\theta-\chi} \Theta^{-\theta}\, \bar{Y}. \tag{A.30}$$

where we will show below that the labor wedge $\mathcal{M}(j)$ is constant.

Using equation (A.21), we can write

$$X(o,f) \quad = \quad \left[\sum_{j \in J} \left(a_j\, [\mathcal{M}(j)\, w(j)]^{1-\nu} + b_j\; q(j)^{1-\nu} \right)^{\frac{1-\psi}{1-\nu}} \right]^{-\frac{1}{1-\psi}} \tag{A.31}$$

Wage markdown

Firms are facing an upward labor supply curve and are monopsonists in the labor market. Next, we derive $\mathcal{M}(j)$ the wedge between the marginal cost of type j labor and the wage $w(j)$. We will show below that the labor wedge is a constant equal to $\mathcal{M}(j) = 1 + 1/\zeta$, a familiar expression in the monopsony literature which obtains when the labor supply curve has a constant elasticity. The proportional difference between marginal cost and the wage equals $\frac{1}{\zeta}$, the elasticity of the inverse labor supply curve, so the firm optimally marks down wages below marginal cost by a factor equal to $1/\mathcal{M}(j)$.

Derivation: In order to derive $\mathcal{M}(j)$, we require several building blocks. First, we need to know how the quantity of task j labor $l(j)$ changes with respect to its own price $w(j)$

$$\frac{\partial \log l(j)}{\partial \log w(j)} = \frac{\partial l(j)}{\partial w(j)} \frac{w(j)}{l(j)} = \left(\frac{1}{\beta} - 1 \right) + \left(1 - \frac{1}{\beta} + \zeta \right) h(j) = \left(\frac{1}{\beta} - 1 \right) (1 - h(j)) + \zeta h(j). \tag{A.32}$$

We also need the cross-price terms

$$\frac{\partial \log l(j)}{\partial \log w(k)} = \left(1 - \frac{1}{\beta} + \zeta \right) h(k), \tag{A.33}$$

which has a sign which depends on whether the between task substitution effect $(1 - 1/\beta < 0)$ dominates the induced increase in the number of workers from higher total wages (ζ).

To derive these equations (A.32-A.33), we work with the identity

$$\log l(j) = \log \bar{\zeta} + (1 - 1/\beta) \log w(j) + [\beta - 1 + \zeta\beta] \log \sum_{j \in J} \exp(\tfrac{1}{\beta} \log w(j)). \tag{A.34}$$

It is straightforward that differentiating the above equation yields the desired results , since

$$\frac{\partial}{\partial w(j)} \log \sum_{j \in J} \exp(\tfrac{1}{\beta} \log w(j)) = \frac{1}{\beta} \frac{w(j)^{1/\beta}}{\sum_{j' \in J} w(j)^{1/\beta}}. \tag{A.35}$$

Next, we need to understand the set of wage changes which allows the firm to increase $l(j)$ while holding the quantity of labor in all other tasks fixed. In elasticity form, the requisite wage changes are

$$\left.\frac{d\log w(k)}{d\log l(j)}\right|_{\substack{d\log l(k)=0\\ k\neq j}} = \mathbf{1}[k=j]\frac{\beta}{1-\beta} + \frac{\frac{1}{\beta}-1-\zeta}{\left(\frac{1}{\beta}-1\right)\zeta}h(j), \tag{A.36}$$

an expression we obtain by inverting the Jacobian matrix capturing the set of elasticities of task quantities with respect to task prices.

Finally, we need to understand how total costs change with each of the task-level wages

$$\frac{\partial[W(o,f)N(o,f)]}{\partial w(j)} = l(j)(1+\zeta). \tag{A.37}$$

To derive equation (A.37), we start with the fact that total wage earnings equals $\bar{\zeta}W(o,f)^{\zeta+1}$. Then by differentiating and using the definition of $l(j)$ from equation (A.29), we get

$$\frac{\partial[W(o,f)N(o,f)]}{\partial w(j)} = (1+\zeta)\underbrace{\bar{\zeta}W(o,f)^{\zeta}}_{=N(o,f)}\frac{w(j)^{\frac{1}{\beta}-1}}{\left[\sum_{j\in J}w(j)^{\frac{1}{\beta}}\right]^{\beta-1}} = (1+\zeta)l(j). \tag{A.38}$$

We can then combine these pieces (A.32, A.33, A.37) to compute marginal cost:

$$\begin{aligned}
\left.\frac{\partial[W(o,f)N(o,f)]}{\partial l(j)}\right|_{\substack{d\log l(k)=0\\ k\neq j}} &= \sum_{k=1}^{J}\frac{\partial[W(o,f)N(o,f)]}{\partial w(k)}\frac{w(k)}{l(j)}\left.\frac{d\log w(k)}{d\log l(j)}\right|_{dl(k)=0} \\
&= (1+\zeta)\left\{\frac{\beta\, w(j)}{1-\beta} + \sum_{k=1}^{J}\frac{w(k)l(k)}{l(j)}\frac{\frac{1}{\beta}-1-\zeta}{\left(\frac{1}{\beta}-1\right)\zeta}h(j)\right\}.
\end{aligned} \tag{A.39}$$

Recalling that $\mathcal{M}(j)$ is the ratio of marginal cost to the wage, and rearranging, we get that

$$\mathcal{M}(j) = \frac{(1+\zeta)}{\frac{1}{\beta}-1}\frac{\zeta[s(j)-h(j)]+(\frac{1}{\beta}-1)h(j)}{\zeta s(j)} = M \equiv 1+\frac{1}{\zeta}, \tag{A.40}$$

where $s(j) \equiv w(j)l(j)/[W(o,f)\,N(o,f)]$ is the cost share in task j. To obtain the final expression, we note that $s(j) = h(j)$.

A Log-linear Approximation

To derive an approximate solution, let us focus on the the symmetric steady state with $a_j = a$, $\beta_j = \beta$, $w(j) = w$ and $q(j) = q$. Suppose $j = 1$ gets shocked and the other tasks are not. Given the definitions of the elasticities, we can write

$$q(1) = q\,e^{-\epsilon}, \qquad q(j) = q, \qquad j \geq 2 \tag{A.41}$$

$$w(1) = w\,e^{\eta_o\,\epsilon}, \qquad w(j) = w\,e^{\eta_c\,\epsilon}, \qquad j \geq 2. \tag{A.42}$$

If we replace the above into (A.30) for $j = 1$ and $j = 2..J$, we get two equations, one for the 'shocked' task and another equation which is common to all unshocked tasks, for $j = 1$. After dividing both sides of each equation with equation (A.30) evaluated at the pre-shock equilibrium, we obtain

At $j = 1$:

$$e^{\eta_o \frac{1}{\beta}\epsilon} \left(\frac{1}{J} e^{\eta_o \frac{1}{\beta}\epsilon} + \frac{J-1}{J} e^{\eta_c \frac{1}{\beta}\epsilon} \right)^{\beta - 1 + \zeta\beta} = e^{\eta_o (1-\nu)\epsilon} \left(s_l \, e^{\eta_o (1-\nu)\epsilon} + (1 - s_l) e^{-(1-\nu)\epsilon} \right)^{\frac{\nu - \psi}{1-\nu}} \left(\frac{\tilde{X}(o,f)}{X(o,f)} \right)^{\chi - \psi} \tag{A.43}$$

at $j \neq 1$

$$e^{\eta_c \frac{1}{\beta}\epsilon} \left(\frac{1}{J} e^{\eta_o \frac{1}{\beta}\epsilon} + \frac{J-1}{J} e^{\eta_c \frac{1}{\beta}\epsilon} \right)^{\beta - 1 + \zeta\beta} = e^{\eta_c (1-\nu)\epsilon} \left(s_l \, e^{\eta_c (1-\nu)\epsilon} + 1 - s_l \right)^{\frac{\nu - \psi}{1-\nu}} \left(\frac{\tilde{X}(o,f)}{X(o,f)} \right)^{\chi - \psi} \tag{A.44}$$

where

$$\frac{\tilde{X}(o,f)}{X(o,f)} = \left[\frac{J-1}{J} \left(s_l \, e^{\eta_c (1-\nu)\epsilon} + 1 - s_l \right)^{\frac{1-\psi}{1-\nu}} + \frac{1}{J} \left(s_l \, e^{\eta_o (1-\nu)\epsilon} + (1 - s_l) e^{-(1-\nu)\epsilon} \right)^{\frac{1-\psi}{1-\nu}} \right]^{-\frac{1}{1-\psi}} \tag{A.45}$$

Taking logs of both sides (A.43) and (A.44), differentiating with respect to ϵ, evaluating it at $\epsilon = 0$ yields two linear equations in two unknowns, η_o and η_c. Define

$$s_l \equiv \frac{a\, M^{1-\nu}\, w^{1-\nu}}{a\, M^{1-\nu}\, w^{1-\nu} + b\, q^{1-\nu}}, \qquad s_k = 1 - s_l \tag{A.46}$$

is the task-level labor share and capital share of output, respectively. The solution to these equations are

$$\eta_c = -\frac{s_k}{J} \frac{(\nu - \chi)(1-\beta) - \beta\Big(\nu\,(\chi - \psi) + \zeta(\nu - \psi)\Big)}{\Big(s_k\,\nu + s_l\,\chi + \zeta\Big)\Big(1 - \beta\,(1 - s_k\,\nu - \psi\, s_l)\Big)} \tag{A.47}$$

as long as $\nu \geq 1$, the above expression is decreasing in ψ. Cross-task spillovers become positive if the within-occupation task complementarity is strong enough, that is the parameter ψ is less than

$$\bar{\psi} = \frac{((\chi + \zeta + 1)\,\nu - \chi)\,\beta + \chi - \nu}{\beta\,(\nu + \zeta)} \tag{A.48}$$

The own-task elasticity, which captures the impact of innovation $\varepsilon(j)$ on $w(j)$, is given by

$$\eta_o = -\frac{s_k}{J} \frac{(\nu - \chi)(1-\beta) + \beta\Big((J-1)(\nu - \psi)\zeta + \nu(\psi - \chi) + J\,(\nu - \psi)(s_k\,\nu + s_l\,\chi)\Big)}{\Big(s_k\,\nu + s_l\,\chi + \zeta\Big)\Big(1 - \beta\,(1 - s_k\,\nu - \psi\, s_l)\Big)}. \tag{A.49}$$

A sufficient, but not necessary, condition for η_o to be negative is that the number of tasks is sufficiently large

$$J \geq \frac{\nu + \zeta}{s_l\, \chi + s_k\, \nu + \zeta} \tag{A.50}$$

and $\nu \geq \psi$.

Last, to derive the expression for wage growth,

$$\log W_1(o, f) - \log W_0(o, f) = b \, \log \left[\sum_{j \in J} w_1(j)^{\frac{1}{b}} \right] - \log \left[\sum_{j \in J} w_0(j)^{\frac{1}{b}} \right] \tag{A.51}$$

We can consider a technology shock that affects the job (o, f) described by a vector of $\varepsilon_1 ... \varepsilon_J$. Thus, we can write

$$w_1(j) = w_0(j) e^{\eta_o\, \varepsilon(j) + \eta_c \sum_{j' \neq j} \varepsilon(j')}. \tag{A.52}$$

Plugging the above into the change in wages, and approximating around $\varepsilon(j) = 0$ for all j, we obtain Approximating the above using a second order expansion, we get that

$$\begin{aligned} \log W_1(o, f) - \log W_0(o, f) \approx & +\eta_c\, (J-1)\, \frac{1}{J} \sum_j \varepsilon(j) + \eta_o \frac{1}{J} \sum_j \varepsilon(j) + \\ & + \frac{b}{2} \left(\frac{\eta_o + \eta_c}{b} \right)^2 \frac{J-1}{J^2} \left(\sum_j \varepsilon(j)^2 - \frac{2}{J-1} \sum_j \sum_{j' > j} \varepsilon(j) \varepsilon(j') \right) \end{aligned} \tag{A.53}$$

Define

$$m(\varepsilon) = \frac{1}{J} \sum_j \varepsilon(j) \tag{A.54}$$

$$\begin{aligned} V(\varepsilon) \equiv & \frac{1}{J} \sum \left(\varepsilon(j) - m(\varepsilon) \right)^2 \\ = & \frac{J-1}{J^2} \left(\sum_j \varepsilon(j)^2 - \frac{2}{J-1} \sum_j \sum_{j' > j} \varepsilon(j) \varepsilon(j') \right) \end{aligned} \tag{A.55}$$

Replacing (A.54) and (A.55) into (A.53) yields the expression for wage growth in the text. The expression for employment growth follows directly from (A.26).

The last step consists of deriving the elasticity to a shock to firm productivity. In the symmetric steady state with $a_j = a$, $\beta_j = \beta$, and $q(j) = q$,

$$J^{\beta - 1 + \zeta\, \beta}\, w^{1+\zeta} \bar{\zeta} = a\, w^{1-\nu} \left(a\, w^{1-\nu} + b\, q^{1-\nu} \right)^{\frac{\nu - \psi}{1-\nu}}\; X(o, f)^{\chi - \psi}\, Z_f^{\theta - \chi} \Theta^{-\theta}\, \bar{Y}. \tag{A.56}$$

where

$$X(o, f) \;\; = \;\; \left[J\, \left(a\, w^{1-\nu} + b\; q^{1-\nu} \right)^{\frac{1-\psi}{1-\nu}} \right]^{-\frac{1}{1-\psi}} \tag{A.57}$$

Plugging the last to the second last equation

$$J^{\beta-1+\zeta\,\beta}\,w^{1+\zeta}\bar{\zeta}=a\,w^{1-\nu}J^{\frac{\psi-\chi}{1-\psi}}\left(a\,w^{1-\nu}+b\,q^{1-\nu}\right)^{\frac{\nu-\chi}{1-\nu}}\;Z_f^{\theta-\chi}\Theta^{-\theta}\,\bar{Y}. \tag{A.58}$$

Suppose productivity of the firm improves

$$\tilde{Z}_f=Z_f\,e^{\epsilon} \tag{A.59}$$

then task prices will change

$$w(j)=w\,e^{\eta_z\,\epsilon},\qquad\forall j \tag{A.60}$$

and (A.58) becomes

$$J^{\beta-1+\zeta\,\beta}\,w^{1+\zeta}e^{(1+\zeta)\eta_z\,\epsilon}\bar{\zeta}=a\,w^{1-\nu}e^{(1-\nu)\eta_z\,\epsilon}J^{\frac{\psi-\chi}{1-\psi}}\left(a\,w^{1-\nu}e^{(1-\nu)\eta_z\,\epsilon}+b\,q^{1-\nu}\right)^{\frac{\nu-\chi}{1-\nu}}\;Z_f^{\theta-\chi}e^{(\theta-\chi)\epsilon}\Theta^{-\theta}\,\bar{Y}. \tag{A.61}$$

Divide both sides of (A.61) with the corresponding sides of (A.58) gives

$$e^{(1+\zeta)\eta_z\,\epsilon}=e^{(1-\nu)\eta_z\,\epsilon}\left(s_l\,e^{(1-\nu)\eta_z\,\epsilon}+s_k\right)^{\frac{\nu-\chi}{1-\nu}}\;e^{(\theta-\chi)\epsilon} \tag{A.62}$$

Taking logs of both sides, differentiating with respect to ϵ, evaluating at $\epsilon=0$ and solving for η_z, we get

$$\eta_z=\frac{\theta-\chi}{s_k\,\nu+s_l\,\chi+\zeta} \tag{A.63}$$

and in terms of wage earnings

$$\begin{aligned}\log W_1(o,f)-\log W_0(o,f)&=\beta\,\log\left[J\,w^{\frac{1}{\beta}}\,e^{\frac{1}{\beta}\,\eta_z\,\epsilon}\right]-\beta\log\left[J\,w^{\frac{1}{\beta}}\right]\\&=\eta_z\,\epsilon.\end{aligned} \tag{A.64}$$

If all task prices go up by the same amount, there is no reallocation, and wages just go up by the same amount.

Task Reallocation

We now derive the elasticity of task-j hours allocation $h(j)$ to AI-driven cost changes. Applying (A.52) to (A.1) for some task j around the symmetric initial equilibrium, we have

$$
\begin{aligned}
\Delta \log h(j) &= \frac{1}{\beta}\left(\eta_o \varepsilon(j) + \eta_c \sum_{j' \neq j} \varepsilon(j')\right) - \log\left(\sum_k \exp\Big(\frac{\eta_o}{b}\varepsilon(k) + \frac{\eta_c}{\beta}\sum_{k' \neq k}\varepsilon(k')\Big)\right) \\
&\approx \frac{1}{\beta}\left[\eta_o \varepsilon(j) + \eta_c \sum_{j' \neq j} \varepsilon(j') - \frac{1}{J}\Big(\sum_k \eta_o \varepsilon(k) + \eta_c \sum_k \sum_{k' \neq k} \varepsilon(k')\Big)\right] \\
&= \frac{1}{\beta}\left[\eta_o\left(\varepsilon(j) - m(\varepsilon)\right) + \eta_c\left(Jm(\varepsilon) - \varepsilon(j) - (J-1)m(\varepsilon)\right)\right]
\end{aligned}
\tag{A.65}
$$

which simplifies to

$$
\Delta \log h(j) \approx \frac{\eta_o - \eta_c}{\beta}\left(\varepsilon(j) - m(\varepsilon)\right) \tag{A.66}
$$

Thus task-specific hours reallocation is proportional to the task-level AI capital cost shock relative to the average shock across all tasks within the occupation.

A.2 Extracting Firm-Level AI Applications

To identify potential AI applications at the firm level, we first impose a filter on the text in workers' job description. After converting job descriptions to lower case, require that the description includes at least one of the following strings: “artificial intelligence”; “machine learning”; ” ml ”; “ ai ”; “deep learning”; “deep-learning”; “neural net”; “neural-net”; “ nlp ”; “natural language processing”; “computer vision”; “large language model”; or “ llm ”. We further require AI positions to come from jobs with 2-digit SOC code between 11 and 19 (professional occupations). Upon reading many examples, we find that AI-tagged positions coming from these occupations are nearly exclusively direct implementers of AI, while this is occasionally not the case for the non-professional occupations.

This results in 547,329 distinct job positions which describe implementing artificial intelligence in at least one application. We consider the position to be active at a firm in a specific year if the position is current for at least a 6 month within the given year. We next apply a series of filters using large language models to read these descriptions of AI positions to extract and clean the phrases which describe specific ways in which AI is being applied. To do this, we use the Llama 3.1 70B model created by Meta. We access the model using an API provided by DeepInfra. We set the temperature parameter to zero in all prompts in order minimize any potential variability in responses to the exact same query.

Step-1 LLM Filter: Identifying and cleaning AI-related phrases

Our first-step LLM filter extracts the specific raw phrases in a job description which describe using AI, as well as an LLM-generated summary of the AI application. The prompt instructs the LLM to follow a four-step process in order to guide its "reasoning". The steps are as follows: 1), filtering out the tasks in the task which are unrelated to applications of AI (including discarding descriptions of hardware related to AI rather than the specific use of AI); 2), generate a list of applications identified from the first step; 3), audit answers to ensure that the AI application is clearly specified; and 4), reread the original text to make sure no AI applications were missed in the original reading. Finally, the LLM is asked to report to the user the key applications filtered from the text; the original raw text that generated the specific key application; and finally, the final answer which is the cleaned AI applications.

Our specific LLM prompt for this step is as follows:

*Your current task is to review the following descriptions of job duties being performed by employees of the same company and summarize each of the applications of AI that you see being performed. The goal is to produce an itemized list, where each item corresponds with a different use case for artificial intelligence methods being described. For each application, please describe, in a few sentences based ONLY on the resume descriptions, what functions AI tools are being applied to perform (it is important not to make predictions unless a use case is described in the text). Your answers should be focused on which tasks these AI tools are being used to perform, rather than on which tools are being used. In other words, I only want you to summarize instances in which these employees describe using AI to perform a specific function or solve a particular problem. I am looking for descriptions of the tasks and functions that *the AI tools themselves are performing*, rather than just the responsibilities or activities of the employees who are working with those tools.*

To organize your efforts, I suggest you follow a four-step process. In the first step, please filter out descriptions of tasks which are unrelated to applications of artificial intelligence. If a description does not refer to how an artificial intelligence method is being used (e.g., because it describes development of hardware or other infrastructure related to AI deployment), please disregard the information. In the second step, produce your temporary itemized list from the filtered text. Now let's start the third step: Think aloud. Please audit your answers according to the original text. Sometimes, a task is clearly AI-related, but the specific application is not really specified. An example would be an employee mentioning that they are maintaining data infrastructure or deploying algorithms without saying anything about which data they are using or what the purpose of the underlying algorithms are. When reviewing your preliminary set of bullets, feel free to discard items which fall into this category of not specifying an actual application. For fourth step, please provide your final answer to improve your previous answers. Before finalizing your answer, please also reread the original body of text and identify any additional applications, if any, which were not included in the original list. Extract key applications from the following text document. Please output ONLY as a JSON list (Do not include "“' and anything else). The JSON should represent a table with three columns:

(1) The first column, labeled 'Key Application', should contain concise summaries or key insights extracted from the text.

(2) The second column, labeled 'Raw Excerpt', should include the corresponding raw excerpts from the text that support each key point.

(3) The third column, labeled 'Final Answer', should include your final answer.

< INSERT JOB DESCRIPTION HERE >

END PROMPT

This prompt does not require that a given position can only use AI for one purpose. Accordingly, out of the 547,329 distinct AI positions, this first prompt identifies 1,324,884 distinct applications of AI.

Step-2 LLM Filter: Removing uninformative text

In the second step we feed in the LLM's final response (the third output from the step-1 query) as input for the query. Responses from the first-step query typically take the form "AI tools are being used to..." or "using NLP to..." followed by the actual application. Because we don't want our textual representations of documents to be biased by these generic and uninformative phrases about the particular AI techniques–rather we want to highlight the specific application, not the particular tool being used to accomplish the application–we devise a prompt designed to filter such language from the text. The prompt allows for deleting an AI application entirely if the description of its use is still too vague to offer a clearly-defined specific application. After following this step, we have 1,115,982 filtered AI applications remaining. The second prompt is as follows:

The excerpt below describes how an artificial intelligence technology is being applied. Assume that it is already known that the excerpt refers to a use of artificial intelligence; the reader only wants to know the specific final application. Therefore, all references to any type of AI tool (e.g. natural language processing, machine learning, computer vision, generative AI, or any specific AI/ML algorithm) are redundant and should be stripped from the text. If the text only contains reference to an AI tool and without a clearly specified application, you should return 'N/A' when you filter the text.

For reference, here are a few examples of correctly applied filters:

-'AI tools are being used to measure text similarity in educational settings using NLP' should become 'Measure text similarity in educational settings'

-'Machine learning is being applied to perform tasks related to database analysis and firmware/software development for embedded environments' should become 'Perform tasks related to database analysis and firmware/software development for embedded environments'

-'AI-powered chatbots are being used to provide customers with quick solutions and answers using

natural language processing capabilities.' should become 'Provide customers with quick solutions and answers.'

-'Analyzing customer reviews using NLP to understand customer needs and wants' should become 'Analyze customer reviews to understand customer needs and wants'

-'AI tool is being used to deploy computer vision model' should become 'N/A", because computer vision models themselves are an AI tool, and the exact use of computer vision is not specified.'

With this in mind, please filter the following excerpt describing an AI application. < STEP 1 LLM OUTPUT HERE >

END PROMPT

Step 3 LLM Filter: Small refinements on step 2

Upon inspection of the LLM output in the second step, we found a few specific phrases which were more likely to be associated with some remaining uninformative text that occasionally bypassed the filter. Accordingly, for the final step we first identify a small subset the AI-related applications with the specific keywords *'data analysis', 'text analysis', 'predictive analytics', 'visualization', 'predictive analysis'.* There are around 35000 such applications, which we pass to an LLM for cleaning. Because there are only 35,000 texts in this step, we use the more expensive but higher-performing GPT-4o model using the OpenAI API. The prompt is: *The excerpt below describes how an artificial intelligence technology is being applied. Please determine if ithe application is very specific. If yes, please summarize the application (without outputing anything else). All references to any type of AI tool (e.g. natural language processing, machine learning, computer vision, generative AI, or any specific AI/ML algorithm) are redundant and should be stripped from the text. Otherwise, respond 'N/A'. Here are some examples:*

-'Predictive Analytics' should be 'N/A' as it is very broad;

-'Data Visualization' should be 'N/A' as it is very broad;

-'AI-driven NFT Collection Visualization' should be kept as it is a very specific application.

-'Perform exploratory data analysis for invoice anomalies' should be 'invoice anomalies'

-'Provide self-service data access and custom visualization interfaces for the oceanic team' should be 'custom visualization interfaces for the oceanic team' as this is a specific application.

With this in mind, please filter the application: <INSERT FILTERED APPLICATION HERE>

END PROMPT

This final filter removes an additional 7600 AI applications from our final set.

Appendix Figures and Tables

Table A.1: Top and bottom 25 occupations by average AI exposure

25 Most Exposed Occupations	25 Least Exposed Occupations
Occupation	Occupation
Market Research Analysts and Marketing Specialists	Tire Builders
Management Analysts	Terrazzo Workers and Finishers
Logisticians	Tire Repairers and Changers
Computer Hardware Engineers	Tree Trimmers and Pruners
Financial Specialists	Bartenders
Computer and Information Systems Managers	Helpers–Carpenters
Sales Engineers	Dishwashers
Financial Risk Specialists	Food Preparation Workers
Transportation, Storage, and Distribution Managers	Maids and Housekeeping Cleaners
Industrial Engineers	Aircraft Service Attendants
Life, Physical, and Social Science Technicians	Animal Trainers
Aerospace Engineers	Actors
Materials Engineers	Ophthalmic Laboratory Technicians
Sales Managers	Gambling Dealers
Sales Representatives of Services	Cooks, Private Household
Credit Analysts	Janitors and Cleaners
Cost Estimators	Childcare Workers
Advertising and Promotions Managers	Food Servers, Nonrestaurant
Marketing Managers	Mechanical Door Repairers
Chemical Engineers	Cooks, Restaurant
Electrical Engineers	Judicial Law Clerks
Purchasing Agents	Insurance Appraisers, Auto Damage
Purchasing Managers	Makeup Artists, Theatrical and Performance
Production, Planning, and Expediting Clerks	Flight Attendants
Bioengineers and Biomedical Engineers	Home Health Aides

Note: This table details the top and bottom 25 occupations ranked by average AI exposure, as determined by the measurement process described in Section 2.3. These rankings are based on the computed AI exposure scores, which leverage task-level similarities between AI applications and occupational descriptions. See Section 2.3 for more details.