Job Transformation, Specialization, and the Labor Market Effects of AI*

Lukas B. Freund Lukas F. Mann

This version: February 10, 2026

First version: August 16, 2025

Please click here for the latest version.

Abstract

A central effect of automation, both historically and today, is to shift the task content of jobs—what we call *job transformation*. We develop a task-based model of job transformation and use it to understand who wins and loses from AI. Occupations bundle multiple tasks. Workers' skills vary across tasks and they sort into occupations by comparative advantage. When a task is automated, the remaining tasks gain in relative importance. This generates wage effects that depend on workers' full skill profiles, not just their occupation's exposure. To quantify these effects, we measure the task content of jobs and estimate the distribution of task-specific skills. We project the wage effects of automation by large language models (LLMs) and characterize winners and losers across three dimensions: occupational exposure, skills, and the wage distribution. First, on average, wage effects are mildly positive at low exposure but sharply negative at high exposure. These effects vary across workers even conditional on occupation and this dispersion increases with exposure. Second, AI raises the value of social and non-routine manual skills while reducing the return to analytical skills. Third, AI is mildly progressive, benefiting low-earners more than high-earners. Together, our results suggest that job transformation is central to the distributional consequences of AI.

*Freund: Boston College, lukas.beat.freund@gmail.com, www.lukasfreund.com; Mann: Arizona State University, lfmann@asu.edu, www.lukasmann.com.

We thank David Autor, Maarten De Ridder, John Grigsby, Kyle Herkenhoff, Joachim Hubmer (discussant), Anders Humlum, Larry Katz, Huiyu Li, Ilse Lindenlaub (discussant), Ellen McGrattan, Simon Mongey, Sergio Ocampo Díaz, Pascual Restrepo, Benjamin Schoefer, Ayşegül Şahin, Joe Stiglitz, Mike Webb and Nathan Zorzi for helpful conversations and comments. We also thank seminar and conference audiences at Columbia, Minneapolis Fed, Boston University, Harvard Kennedy School, Harvard, IMF, Columbia Business School, and BC, as well as at Midwest Macro 2025, SED 2025, BSE Summer Forum 2025, SITE 2025, BC Macro-AI Workshop 2025, NBER EFG Fall Meeting 2025, Philadelphia Junior Macro Workshop 2025, and the NBER Labor Studies Fall Meeting 2025 for useful comments and suggestions. Alli Tucker and Claude provided outstanding research assistance. A conference draft of this paper was circulated with the title "For Whom the Bot Tolls: Specialization and the Earnings Effects of AI."

1 Introduction

Recent advances in artificial intelligence (AI) raise the prospect of machines taking over an increasing number of tasks. What will be the labor market consequences? History and a rich empirical literature (e.g., Autor *et al.*, 2003; Spitz-Oener, 2006; Atalay *et al.*, 2020) suggest that a central effect of automation is to shift the task content within jobs—what we call *job transformation*. During the industrial revolution, power looms shifted weavers' duties from physical labor toward monitoring and trouble-shooting multiple looms (Bessen, 2012). In the late 20th century, CNC tools moved machinists from routine tasks like positioning parts to non-routine tasks like programming digital processes (Bartel *et al.*, 2007). The same process is unfolding today: legal associates spend more time with clients as AI takes care of document review; radiologists concentrate on patient consultation and care coordination as AI handles routine image analysis; and software engineers focus on program design as AI tools churn out code.[1]

Notwithstanding the seeming importance of job transformation, quantifying its labor market effects is difficult. How jobs are transformed depends on which specific tasks will be automated; and, crucially, inferring the effects on wages requires knowledge of workers' entire portfolios of task-specific skills, which are typically unobserved. Facing these measurement challenges, state-of-the-art models of automation generally abstract from job transformation entirely (e.g., Acemoglu and Restrepo, 2018a,b).

In this paper, we construct a formal model of job transformation, estimate it, and project how automation by large language models (LLMs) will affect wages for heterogeneously skilled workers. Our findings characterize the winners and losers of AI-induced automation from three complementary perspectives: occupational exposure, skills, and the wage distribution. First, the average wage effects of occupational exposure to AI are non-monotonic. Wages are roughly unchanged in occupation with no exposure, rise by around 4% at moderate exposure, and drop by as much as 35% in the most exposed occupations. These averages mask, however, that within any occupation there are winners and losers and such dispersion increases with exposure. Second, AI raises the return to social and non-routine manual skills, while reducing the return to analytical skills. Third, the overall effects of AI are mildly progressive: average wage gains for low-wage workers (3%) exceed those for high-wage workers (1%), reflecting a decline in the relative return to skills concentrated at the top of the wage distribution. None of these results arises without job transformation. Together, they show job transformation is crucial for understanding the labor market consequences of AI.

[1]See Appendix C.1 for details on the historical vignettes. For present-day examples, see `https://vivekhaldar.com/articles/when-compilers-were-the--ai--that-scared-programmers`, Financial Times (Dec 04, 2025), and Financial Times (Dec 18, 2025). More systematic studies likewise point to large-scale task shifts within jobs (Bonney *et al.*, 2024).

Theory. In the first part of the paper, we formulate a general equilibrium model of automation-driven job transformation, building on canonical task-based theory.[2] As usual in the automation literature, production involves tasks, which are assigned to humans or machines. Distinctively, task aggregation occurs at the occupational level, with different occupations attaching heterogeneous weights to tasks. A standard, final-goods sector aggregates occupation-specific outputs, capturing general equilibrium price effects. Each worker has a portfolio of task-specific skills, i.e., their productivity may vary across tasks. We integrate this production structure with Roy (1951)-style occupational choice based on comparative advantage.

A key ingredient is *task bundling*: to produce output, a worker must perform all tasks in an occupation, including tasks they are relatively less skilled at. We do not model the sources of bundling but view it as eminently plausible—most economists, say, are more productive than the average worker at data analysis or math but no more productive at emailing, yet emailing remains part of their job—and later discipline the degree of bundling by measuring the task mix of jobs. Task bundling significantly affects how automation works. When one of several tasks is reassigned from labor to machines, workers in affected occupations spend less time on that task (none if fully automated), while time on all other bundled tasks increases proportionately. This is *job transformation*.

The model yields an intuitive characterization of wages and their response to automation. Wage levels reflect both absolute advantage (i.e., average skill) and comparative advantage (i.e., alignment between skill specialization and occupational task requirements). Following automation, individual wages change through the standard displacement and productivity effects, whose net effect appears as an occupation-level shifter, and through job transformation: wages respond to the interaction between changing task weights and workers' multidimensional skills. Workers more skilled at automated tasks than at non-automated tasks lose, those relatively less skilled at automated tasks gain. Consider lawyers: when document processing is automated, a lawyer specialized in this task loses, whereas a colleague with a comparative advantage in client engagement gains. While occupational automation exposure affects the potential magnitude of job transformation effects, their sign and size depends on individual, multi-dimensional skills.

Measuring the skill distribution. In the second part of the paper, we estimate the distribution of multi-dimensional skills, leveraging the model's structure. The identification logic is as follows. Consider two occupations, economists and software engineers, where both code, but economists also write. Identification comes from two sources: wage comparisons and occupational choices. If we observe a worker in both occupations, their wage as a software engineer reveals their coding skill, letting us infer their writing skill from their economist wage. If we only observe the worker as an economist, this reveals their coding skills fall below the threshold for choosing software

[2] See, in particular, Autor *et al.* (2003); Acemoglu and Autor (2011); Acemoglu and Restrepo (2018b).

engineering. We formalize this intuition as a maximum-likelihood estimation problem to recover the means and variance-covariance structure of skills alongside other structural parameters.

We implement the estimation with publicly available data, the National Longitudinal Survey of Youth (NLSY) 1979, and measures of occupational task weights which we construct using natural language processing (NLP) techniques. We cluster approximately 19,000 detailed occupation-specific tasks from the Occupational Information Network (O*NET) into 38 tractable and interpretable task categories. Since, in our theory, occupational task weights reflect optimal time allocation, we recover these weights by using an LLM to measure time shares across tasks for each occupation.[3] We measure these task weights for two periods: pre-2000 and post-2000.

We extensively validate the estimated model. First, the estimated skill correlations, aggregated to the level of typical classifications, are consistent with the empirical literature. Second, in steady state, the model matches several salient empirical moments. For example, in both data and model, workers move to jobs with task requirements similar to their origin occupation, consistent with skills being task-specific but hard to reconcile with general or occupation-specific skills (Gathmann and Schönberg, 2010). Third, and importantly, we study "routine-biased technological change" (RBTC) in the late 20$^{\text{th}}$ century through the lens of the model. Our model captures RBTC through changing task inputs, both within occupations and due to shifting employment shares, aligned with existing evidence (Autor *et al.*, 2003; Atalay *et al.*, 2020). We then replicate two influential empirical studies inside our structural model and show that its predictions are consistent with their findings: RBTC gives rise to job polarization (Goos *et al.*, 2009; Autor and Dorn, 2013) and a rising return to social skills (Deming, 2017).

The effects of LLM automation. In the third part of the paper, we use the estimated model to project how automation by LLMs will affect wages for heterogeneous workers. Our measurement strategy maps model tasks to existing measures of automation exposure, allowing us to pin down which specific tasks are affected. We draw on Eloundou *et al.*'s (2023) data, which identify several information-processing tasks—common in white-collar roles such as financial analysts—as most exposed to LLMs. We consider the scenario where the five most exposed tasks are automated and characterize who wins and who loses along three complementary dimensions: occupational exposure, skills, and the wage distribution.

First, the wage effects of occupational exposure to AI are non-monotonic and highly dispersed, even among workers in the same occupation. On average, moderate exposure benefits workers (individuals in occupations with around 10% exposure gain 4%); high exposure harms them (those in the most exposed occupations lose around 35% on average). These non-monotonic effects arise from job transformation interacting with selection. As workers sort into occupations by comparative advantage, incumbents of highly exposed occupations tend to be

[3] We extensively validate this LLM-based approach, including through comparisons to worker time diary data.

specialized in automatable tasks. When exposure is moderate, however, automation affects more peripheral tasks, freeing workers to concentrate on their comparative advantages. Yet these averages mask large heterogeneity within occupations: two workers in the same occupation can experience wage changes of opposite sign. On the one hand, some workers are "trapped" and incur large losses. Their wage drops not only in their origin occupation but also in their most likely outside options, which undergo a similar transformation. On the other hand, two groups of workers gain. First, those in exposed occupations who excel at customer-facing and coordination tasks stay and experience wage gains, as work rebalances toward their strengths. Second, some workers switch *into* highly exposed occupations and experience large gains. For them, automation removed tasks that previously represented a skill-based entry barrier.

Second, AI raises the return to social and non-routine manual skills (e.g., public speaking or repairing equipment, respectively) and, in a break with the past, reduces the return to analytical skills (e.g., analyzing data). These effects disappear absent job transformation. Workers with high analytical skills are thus over-represented among those who lose from the AI shock, while those with high non-routine manual skills are over-represented among winners.

Third, AI is mildly progressive, benefiting low-earners more than high-earners on average. Low-wage workers gain around 3% on average, a figure that drops to around 1% at the top. Analytical skills are concentrated at the top of the wage distribution, so automation that reduces their return compresses wages, while non-routine manual skills, which gain in value, are more equally distributed across workers. Our model thus highlights a novel channel through which AI has a progressive effect: the reshaping of returns to multi-dimensional skills through job transformation.

Literature. Our paper contributes to, and integrates, three strands of the literature on, respectively, the theory of task-based production; the measurement of skills; and the quantitative analysis of AI's labor market consequences.

First, we contribute to the literature on task-based production by developing a formal model of job transformation.[4] In comparison to models with economy- or sector-level technologies (e.g., Acemoglu and Restrepo, 2022), our model centers on a job-level production technology that bundles together multiple tasks. This distinctively enables us to capture how automation affects wages through job transformation—a mechanism that is not captured in the standard model.[5] We furthermore integrate the task-based theory of production with Roy-style occupational choice, allowing us to capture occupational reallocation as an empirically salient adjustment

[4] See, among others, Autor *et al.* (2003); Acemoglu and Autor (2011); Acemoglu and Restrepo (2018b); Ocampo Díaz (2022); Moll *et al.* (2022); Freund (2025); Restrepo (2024).

[5] An additional difference is that, in our approach, positive productivity effects accrue only to exposed occupations, because automated tasks are not bundled together with every other task as in existing models.

margin (Dauth *et al.*, 2021; Boustan *et al.*, 2022).[6] Our primary contribution lies in measurement: The labor share is no longer a sufficient statistic for displacement, so we show how to leverage measures of task-level exposure; and wage effects require knowledge of the distribution of task-specific skills, which we estimate. Our emphasis on task bundling is shared with Autor and Thompson (2025)[7], but the papers are otherwise different in methodology and focus: Autor and Thompson (2025) use their model with a novel reduced-form approach to resolve the historical puzzle of why routine task automation often raised wages in routine-intensive occupations despite employment declines. We provide a structural, quantitative analysis of earnings effects from ongoing AI automation.

Second, we contribute to the literature on multi-dimensional skills.[8] Guvenen *et al.* (2020), Lise and Postel-Vinay (2020) and Baley *et al.* (2022) all use military test scores to measure a small number of broad skill types, and assume a parametric relationship between these measures and workers' task-level productivity. In contrast, we directly estimate the distribution of task-level productivity using our model. Our approach offers three advantages. First, we are able to remain agnostic on the relationship between test scores and task-level productivity. Second, our methodology can be applied to any large-scale worker dataset with information on occupations and wages, without requiring (rare) data on test scores. Third, we estimate skill distributions in potentially high-dimensional task spaces rather than being restricted to low-dimensional categories like cognitive, manual, and interpersonal skills. This allows us to connect skills to granular tasks, as considered in the literature on automation exposure.[9] Overall, our paper closes the gap between two strands of research. On the one hand, the literature on multi-dimensional skills has thought carefully about skill measurement and sorting but relies on abstract notions of technological change. On the other hand, research on tasks—both the theoretical literature and empirical studies documenting shifts in task requirements within jobs over time—highlights how the demand for specific tasks is shaped by automation.[10] We close this gap by estimating the

[6]The literature on occupational choice is vast. Recent contributions include, e.g., Dix-Carneiro (2014); Hsieh *et al.* (2019); Traiberman (2019), as well as Humlum (2019); del Rio-Chanona *et al.* (2021); Bocquet (2022); Fan (2025); Böhm *et al.* (2025); Grigsby and Zorzi (2025).

[7]On bundling also see Heckman and Scheinkman (1987), Edmond and Mongey (2021), Hernnäs (2023) and Choné *et al.* (2025).

[8]In terms of theory, Lindenlaub (2017) likewise studies multidimensional matching between workers and jobs and how technological change shapes it. While Lindenlaub (2017) focuses on shifts in complementarity between skills and production requirements, we adopt a task-based approach to study automation. Our model also resembles Lazear's (2009) skills-weights approach, treating skills not as inherently specific to a single production unit—firms in Lazear (2009), occupations in ours—but recognizing that different units attach different weights to different skills.

[9]Grigsby (2023), while pursuing a different question, likewise infers the multidimensional skill distribution from occupational choices and wages. The most important of several differences in methodology is that in Grigsby's (2023) approach, a task corresponds to a group of occupations, whereas we conceptualize occupations as bundles of tasks and estimate the distribution of these granular skills. This distinction between occupations and tasks is essential for studying the consequences of job transformation.

[10]Woessmann (2024, p.4) summarizes the gap in the literature this paper helps fill: "[Although] worker skills

distribution of skills at the level of granular tasks, which allows quantifying the earnings effects of job transformation.

Third, our paper contributes to a burgeoning literature evaluating the labor market consequences of AI. One influential strand empirically quantifies task exposure to new technologies (Webb, 2019; Kogan *et al.*, 2023; Felten *et al.*, 2018, 2021; Brynjolfsson *et al.*, 2018; Eloundou *et al.*, 2023). These exposure measures alone cannot predict earnings consequences. Our paper complements this work by offering a structural approach to map task exposure measures to individual-level labor market outcomes. Our findings call for caution in interpreting exposure, as similarly exposed individuals may experience very different earnings effects.

Methodologically, our work belongs to a second strand that uses structural models.[11] Althoff and Reichardt (2025) use a generalized Lise and Postel-Vinay (2020) model to assess the labor market impact of AI. Compared to our paper, they consider several channels beyond automation, which they quantify relying on LLM-generated forecasts for AI productivity gains and task simplification; their results highlight the latter channel. By contrast, our paper is most clearly distinguished in its focus on job transformation. We directly estimate the distribution of task-level skills and demonstrate that automation-driven job transformation is a significant channel along multiple dimensions. Hampole *et al.* (2025) use CV and job posting data to construct firm- and time-varying measures of exposure to existing machine learning/AI techniques, which allows them to study heterogeneity across firms, a margin we are silent on. Like ours, their model features occupations comprising multiple tasks. We make two distinct contributions. First, whereas in their model skills are ex-ante identical across workers and tasks, our model revolves around workers with heterogeneous portfolios of multi-dimensional skills, which we empirically discipline. Indeed, we find that workers, even in the same occupation, may fare very differently depending on their skills. Second, while Hampole *et al.* (2025) study an earlier wave of machine learning/AI techniques, our model links to forward-looking task exposure measures. Overall, our approach enables us to construct forward-looking projections quantifying the labor market consequences of AI—central to policymakers' concerns.

---

motivate the entire task-based approach to how labor markets adjust to technological change, the consideration of multidimensional tasks has not been matched by multidimensional measurement of skills on the empirical side. While the tasks required in different jobs are richly described, worker skills are still mostly proxied rudimentarily by educational degree."

[11] Beyond the papers discussed in the main text, recent quantitative models of AI effects include (Fan and Restrepo, 2025; Lashkari *et al.*, 2025; Chequer *et al.*, 2025; Ide and Talamas, 2025). The broader literature on AI is vast, comprising surveys characterizing adoption (e.g., Bick *et al.*, 2024; Humlum and Vestergaard, 2025b); models of growth and R&D automation (Aghion *et al.*, 2017; Jones, 2022; Jones and Tonetti, 2026); and numerous RCT studies that causally identify the productivity effects of generative AI adoption in narrowly defined contexts.

2 Theoretical Framework

In this section we set out the theoretical environment (Section 2.1), derive optimality conditions and define the equilibrium (Section 2.2). We then define automation in the context of the model and characterize its effects on wages (Section 2.3).

2.1 Environment

Time is discrete and runs forever. The economy is populated by a unit mass of infinitely-lived workers. There are two types of goods: a final good, which is consumed by workers and serves as the numeraire, and occupation-specific output, which serves as an input to produce the final good. There is a representative final good aggregator that purchases occupation-specific output to produce the final good. All markets are perfectly competitive.

Workers. Before the onset of time, each worker draws and observes their skill vector $s_i \in \mathbb{R}^{n_{\text{skill}}}$, where $s_i \sim \mathcal{N}(\bar{s}, \Sigma_s)$. This skill vector remains fixed forever.[12] In each period t, a worker draws two shocks: a productivity shock $\varepsilon_{i,t} \sim \mathcal{N}(\delta_t, \varsigma^2)$, and a vector of occupation-specific preference shocks $u_{i,\cdot,t} \in \mathbb{R}^{n_{\text{occ}}}$, $u_{i,\cdot,t} \sim \text{Gumbel}(0, \nu)$.

Final good aggregator. A perfectly competitive final goods aggregator produces a homogeneous numeraire good Y_t by combining occupation-specific outputs $\{Y_{o,t}\}_{o \in \mathcal{O}}$ via a constant elasticity of substitution (CES) production function with elasticity σ:

$$Y_t = \left(\sum_{o \in \mathcal{O}} \omega_o^{1/\sigma} Y_{o,t}^{(\sigma-1)/\sigma} \right)^{\sigma/(\sigma-1)}$$

where $\{\omega_o\}_{o \in \mathcal{O}}$ are output weights. The final goods aggregator takes occupation-specific output prices $\{P_{o,t}\}_{o \in \mathcal{O}}$ as given and chooses demands $\{Y_{o,t}\}_{o \in \mathcal{O}}$ to minimize costs. We normalize the final good price to unity.

Production. Production occurs across n_{occ} occupations indexed by $o \in \mathcal{O}$, each producing a distinct output sold at price $P_{o,t}$. Production in occupation o requires that a series of tasks $\tau \in \mathcal{T}$ be carried out; what distinguishes occupations from each other are the weights $\{\alpha_{o,\tau}\}_{\tau \in \mathcal{T}}$ attached to these tasks, with $\sum_{\mathcal{T}} \alpha_{o,\tau} = 1 \, \forall o \in \mathcal{O}$. Concretely, the amount of output in an occupation o job is determined by a Cobb-Douglas aggregator with occupation-specific weights

[12] The assumption that skills are time-invariant is driven by computational constraints in the estimation of skills.

$\alpha_{o,\tau}$. Hence, the output of a worker i in occupation o is

$$Y_{i,o,t} = \prod_{\tau \in \mathcal{T}} X_{i,o,\tau,t}^{\alpha_{o,\tau}} \tag{1}$$

where $X_{i,o,\tau,t}$ is the amount of task τ used in production.[13] We interpret these tasks as concrete work steps that need to be performed in a given occupation, such as analyzing business data, moving materials, delivering instruction, etc. A task can be produced using (i) the worker's time or (ii) machine capital. Machine capital has a productivity $\exp(z_\tau)$ at task τ and can be rented from an infinitely elastic capital market at exogenous rate r.[14] We denote the set of tasks produced with human labor as $\mathcal{T}_l$ and the set of tasks produced with machine capital as $\mathcal{T}_m$. For now, we treat these sets as exogenous and assume only that they do depend neither on the specific occupation nor on the skill of any individual worker. For the purposes of formalizing automation, Section 2.3.1 discusses a set of additional assumptions under which $(\mathcal{T}_l, \mathcal{T}_m)$ can be endogenized.

A competitive firm sector sets (log) wages $w_{i,o,t}$ as a function of occupational output prices $P_{o,t}$ on the one hand, and a worker's skill $s_{i,\tau}$ and idiosyncratic shock $\varepsilon_{i,t}$ on the other. The firm freely allocates the worker's unit measure of labor across tasks in $\mathcal{T}_l$, employing effective labor $\exp\left(s_{i,\tau} + \varepsilon_{i,t}\right) \cdot \ell_{i,o,\tau,t}$ to produce task τ.[15] For any task $\tau \in \mathcal{T}_m$, the firm chooses what quantity of capital $M_{i,o,\tau,t}$ to rent. The firm thus optimizes output subject to the constraints

$$\sum_{\tau \in \mathcal{T}_l} \ell_{i,o,\tau,t} = 1$$

$$X_{i,o,\tau,t} = \begin{cases} \exp\left(s_{i,\tau} + \varepsilon_{i,t}\right) \cdot \ell_{i,o,\tau,t} & \text{if } \tau \in \mathcal{T}_l \\ \exp\left(z_\tau\right) \cdot M_{i,o,\tau,t} & \text{if } \tau \in \mathcal{T}_m \end{cases}$$

Occupational choice. In every period t, each worker chooses an occupation to work in. Given their skill vector s_i and productivity shock $\varepsilon_{i,t}$, they fully anticipate their earnings conditional on entering occupation $o \in \mathcal{O}$. We assume that in any period t, the worker chooses the occupation

[13] The unit elasticity of substitution across bundled tasks implicit in equation (1) represents a common baseline in the literature (e.g. Acemoglu and Restrepo, 2022, pp. 1986) and carries some important advantages. First, it provides a transparent way for measuring $\{\alpha_{o,\tau}\}_{o \in \mathcal{O}}$, as described in Section 3.2, because task shares are invariant to shifts in task-specific skill. Second, it confers significant tractability when estimating the skill distribution by producing a log-linear wage equation. Third, because under Cobb-Douglas a productivity-enhancing automation shock does not mechanically increase the relative demand for human-performed tasks (and hence wages), this assumption transparently isolates wage effects arising from the interaction of shifting task weights and skill specialization.

[14] An infinitely elastic capital supply will tend to raise average wages following the adoption of a new automation technology (Caselli and Manning, 2019), relative to the case of a fixed capital stock (Acemoglu and Restrepo, 2018b, Section I). Our focus lies on the distributional effects.

[15] For ease of notation, we suppress skills for machine tasks from the vector of human skills, i.e., $|\mathcal{T}_l| = n_{\text{skills}}$.

yielding the highest utility given their individual vector of occupation-specific wages and preference shocks $u_{i,\cdot,t}$.[16] We further assume that each worker has log utility over their consumption of the numeraire, which equals their wage. Thus, the worker's occupational choice $\hat{o}_{i,t}$ is a function of log wages:

$$\hat{o}_{i,t} = \text{argmax}_o w_{i,o,t} + u_{i,o,t}. \tag{2}$$

The total amount of occupational output produced is given by an aggregate of the worker's choice probabilities and the occupation-specific output they produce:

$$Y_{o,t} = \int P(\hat{o} = o | w_{i,\cdot,t}) \cdot Y_{i,o,t} \, d\Gamma(i).$$

where Γ denotes the distribution of types in the population.

2.2 Optimality conditions and equilibrium

We next characterize optimality conditions, derive formulas for equilibrium wages and occupational choice, and then define an equilibrium.

Firm optimality and output. The firm's problem is

$$\max_{\ell_{i,o,\tau,t}, M_{i,o,\tau,t}} P_{o,t} Y_{i,o,t}(\{\ell_{i,o,\tau,t}\}_{\tau \in \mathcal{T}_l}, \{M_{i,o,\tau,t}\}_{\tau \in \mathcal{T}_m}) - \exp(w_{i,o,t}) \cdot 1 - r \sum_{\tau \in \mathcal{T}_m} M_{i,o,\tau,t}$$
$$\text{s.t.} \sum_{\tau \in \mathcal{T}_l} \ell_{i,o,\tau,t} = 1$$

where

$$Y_{i,o,t}(\{\ell_{i,o,\tau,t}\}_{\tau \in \mathcal{T}_l}, \{M_{i,o,\tau,t}\}_{\tau \in \mathcal{T}_m}) = \prod_{\tau \in \mathcal{T}_l} \left(\exp(s_{i,\tau} + \varepsilon_{i,t}) \cdot \ell_{i,o,\tau,t}\right)^{\alpha_{o,\tau}} \prod_{\tau \in \mathcal{T}_m} \left(\exp(z_\tau) \cdot M_{i,o,\tau,t}\right)^{\alpha_{o,\tau}}.$$

[16] We introduce no exogenous occupational switching frictions, so any persistence in occupational choices arises endogenously from the interaction of task-level skill specialization and occupational differences in task loadings.

Defining $M_{i,o,t} := \sum_{\tau\in\mathcal{T}_m} M_{i,o,\tau,t}$ and taking first order conditions yields

$$\begin{aligned} \ell_{i,o,\tau,t} &= \frac{\alpha_{o,\tau}}{\sum_{\tau'\in\mathcal{T}_l}\alpha_{o,\tau'}}\,\forall\tau\in\mathcal{T}_l, \\ M_{i,o,\tau,t} &= \frac{\alpha_{o,\tau}}{\sum_{\tau'\in\mathcal{T}_m}\alpha_{o,\tau'}} M_{i,o,t}\,\forall\tau\in\mathcal{T}_m, \\ \left(\sum_{\tau\in\mathcal{T}_m}\alpha_{o,\tau}\right)&\frac{P_{o,t}Y_{i,o,t}}{M_{i,o,t}} = r. \end{aligned} \tag{3}$$

which implies that log output equals

$$\begin{aligned} y_{i,o,t} = \log Y_{i,o,t} = &\sum_{\tau\in\mathcal{T}_\ell}\frac{\alpha_{o,\tau}}{LS_o}s_{i,\tau} + \varepsilon_{i,t} + \sum_{\tau\in\mathcal{T}}\frac{\alpha_{o,\tau}}{LS_o}\log\alpha_{o,\tau} \\ &+ \sum_{\tau\in\mathcal{T}_m}\frac{\alpha_{o,\tau}}{LS_o}(z_\tau - \log r + \log P_{o,t}) - \log LS_o. \end{aligned}$$

Wages. Zero profits imply

$$\exp(w_{i,o,t}) = LS_o\,P_{o,t}\,Y_{i,o,t}, \qquad w_{i,o,t} = \log(LS_o) + \log P_{o,t} + y_{i,o,t},$$

such that the log wage of individual i in occupation o given their skill vector $s_{i,\cdot}$ and their productivity shock $\varepsilon_{i,t}$ is

$$w_{i,o,t} = \mu_{o,t} + \sum_{\mathcal{T}_l}\frac{\alpha_{o,\tau}}{LS_o}\cdot s_{i,\tau} + \varepsilon_{i,t} \tag{4}$$

where $\mathrm{LS}_o = \sum_{\tau\in\mathcal{T}_l}\alpha_{o,\tau}$ denotes the labor share in occupation o and the occupation-specific term is

$$\mu_{o,t} = \frac{1}{LS_o}\log P_{o,t} + \sum_{\tau\in\mathcal{T}}\frac{\alpha_{o,\tau}}{\sum_{\tau\in\mathcal{T}_l}\alpha_{o,\tau}}\log\left(\alpha_{o,\tau}\right) + \left(\sum_{\tau\in\mathcal{T}_m}\frac{\alpha_{o,\tau}}{\sum_{\tau\in\mathcal{T}_l}\alpha_{o,\tau}}\left(z_\tau - \log r\right)\right).$$

To gain intuition, we can can write the wage for worker i in occupation o as

$$w_{i,o,t} = \mu_{o,t} + \underbrace{\frac{1}{n_{\text{skill}}}\sum_{\mathcal{T}_l}s_{i,\tau}}_{\text{scalar absolute advantage}} + \mathrm{Cov}\left(n_{\text{skill}}\cdot\frac{\alpha_{o,\cdot}}{LS_o}, \underbrace{s_{i,\cdot} - \frac{1}{n_{\text{skill}}}\sum_{\mathcal{T}_l}s_{i,\tau}}_{\text{specialization vector}}\right) + \varepsilon_{i,t}, \tag{5}$$

where the covariance is taken with respect to equal weights over $\tau \in \mathcal{T}_l$, with $n$skill $= |\mathcal{T}_l|$. Thus, worker i's wage in o depends on both *absolute* advantage, captured by the unweighted average skill, and on *comparative* advantage, i.e., how much the worker specializes in the skills that are important for that occupation, as captured by the covariance term.

For future reference, it is useful to define the following matrix, which contains the relative weights across tasks assigned to labor for each occupation.

Remark 1 (Task-weight matrix.)**.** *The matrix A, defined as*

$$A = \begin{pmatrix} \frac{\alpha_{1,1}}{LS_1} & \frac{\alpha_{1,2}}{LS_1} & \cdots & \frac{\alpha_{1,n_{skill}}}{LS_1} \\ \vdots & \vdots & \ddots & \vdots \\ \frac{\alpha_{n_{occ},1}}{LS_{n_{occ}}} & \frac{\alpha_{n_{occ},2}}{LS_{n_{occ}}} & \cdots & \frac{\alpha_{n_{occ},n_{skill}}}{LS_{n_{occ}}} \end{pmatrix} \in \mathbb{R}^{n_{occ} \times n_{skill}} \tag{6}$$

summarizes the relative weights attached to each task $\tau \in \mathcal{T}_l$ across occupations $o \in \mathcal{O}$. The row vector $A_o := A_{o,\cdot}$ contains the relative task weights for occupation o.

To make this more tangible, consider the job of financial analysts and suppose this occupation comprises four tasks. One, numerical calculations, is taken care of by a machine, while three others are performed by workers and carry equal weight: creating financial models, writing reports to guide investment decisions, and communicating insights with clients. In this example, the row of A corresponding to financial analysts comprises three entries equal to $1/3$ and 0s otherwise.

Using the A notation, we can write the vector of wages for a worker with skill vector s as

$$w_{i,\cdot,t} = \mu_{\cdot,t} + As_i + \varepsilon_{i,t} \in \mathbb{R}^{n_{\text{occ}}}$$

Occupational choice. Given the utility maximization problem in equation (2), the probability that individual i chooses occupation o conditional on their wage vector $w_{i,\cdot}$ is

$$P(\hat{o} = o | w_{i,\cdot,t}) = \frac{\exp(w_{i,o,t}/\nu)}{\sum_{o'} \exp(w_{i,o',t}/\nu)} \tag{7}$$

Finally, we can define an equilibrium.

Remark 2 (Equilibrium)**.** *An equilibrium is a vector of occupational and final good output $(Y_{\cdot,t}, Y_t)$, a distribution $\Gamma(i)$, occupation choices $\hat{o}_{i,t}$, log wages $\{w_{i,o,t}\}$, log skills s_i, idiosyncratic productivity shocks $\varepsilon_{i,t}$ that are functions of i, and a set of prices $\{P_{o,t}\}_{o \in \mathcal{O}}$, such that: (i) equation* (4) *holds at any point in the distribution (firms make zero profits); (ii) the marginal distribution of occupations conditional on wages follows equation* (7) *(workers optimize); (iii) the final goods* *aggregator optimizes, yielding occupation-level demands* $Y_{o,t} = \omega_o P_{o,t}^{-\sigma} Y_t$ *for all* $o \in \mathcal{O}$*; (iv) occupation-level output markets clear:* $Y_{o,t} = \int \mathbb{I}\{\hat{o}_{i,t} = o\} Y_{i,o,t} d\Gamma(i)$ *for all* $o \in \mathcal{O}$*; (v) the final good aggregator makes zero profits:* $Y_t = \sum_{o \in \mathcal{O}} P_{o,t} Y_{o,t}$*; and (vi) the unconditional marginal distributions of skills* s_i *and occupational shocks* $\varepsilon_{i,t}$ *follow* $\mathcal{N}(\bar{s}, \Sigma_s)$ *and* $\mathcal{N}(\delta_t, \varsigma^2 I)$*, respectively.*

2.3 The wage effects of automation in theory

What happens when a particular task $\tau^\star$ is automated? We now formalize automation, then characterize the induced wage change as a function of skills. We allow for arbitrarily large shocks with potentially non-linear effects rather than relying on first-order perturbation methods, which may not capture a shock's transformative nature.

2.3.1 Automation in the model

An automation shock is a one-time, permanent change of $z_{\tau^\star}$ that leads to the reassignment of task $\tau^\star$ from labor to machines. Formally, we endogenize the task assignment $(\mathcal{T}_l, \mathcal{T}_m)$ and make it dependent on the underlying machine productivity z_τ at every task τ. Appendix A.2 outlines the set of assumptions we introduce to this end and derives a minimum-machine productivity threshold $\bar{z}_{\tau^\star}$ above which automation optimally occurs in equilibrium. The qualitative analyses in this Section 2.3 hold for any value of machine productivity $\{z_{\tau^\star} : z_{\tau^\star} \geq \bar{z}_{\tau^\star}\}$. For our quantitative analyses in Section 4 we will need to take a stand on the exact value of $z_{\tau^\star}$.[17]

Letting the prime symbol denote variables after an automation shock in period $t^\star$, the new task sets following automation are

$$\mathcal{T}_l' = \mathcal{T}_l \setminus \{\tau^\star\}, \quad \mathcal{T}_m' = \mathcal{T}_m \cup \{\tau^\star\}.$$

Associated with the automation shock is a change in the occupational task weight matrix A, whereby automation reduces the weight on the automated task to zero and increases the weight on all other entries proportional to their weight. This change in A represents *job transformation*:

$$\begin{aligned} A_o' - A_o &= \begin{pmatrix} \frac{\alpha_{o,1}}{LS_o'} \cdot \frac{\alpha_{o,\tau^\star}}{LS_o} & \frac{\alpha_{o,2}}{LS_o'} \cdot \frac{\alpha_{o,\tau^\star}}{LS_o} & \dots & -\frac{\alpha_{o,\tau^\star}}{LS_o} & \dots \end{pmatrix} \\ &= \begin{pmatrix} \frac{\alpha_{o,1}}{LS_o'} & \frac{\alpha_{o,2}}{LS_o'} & \dots & -1 & \dots \end{pmatrix} \cdot \frac{\alpha_{o,\tau^\star}}{LS_o} \end{aligned}$$

[17] In models of automation which do not explicitly feature task bundling or occupational choice, such as Acemoglu and Restrepo (2018b) and Acemoglu and Restrepo (2022), tasks can be ordered by the relative productivity of humans to machines. The threshold at which automation occurs can then be written as the point at which this productivity ratio equals the ratio of wages to capital costs. The introduction of occupational choice and skill heterogeneity in our setting complicates this simple rule. To maintain tractability, we maintain the assumption that the automation choice is common across occupations and workers of different skills.

Thus, in our earlier example of financial analysts, if a new technology becomes available that allows the employer to entirely automate the writing of reports, this will raise the relative weight on the remaining two tasks performed by labor, financial modeling and client interaction, to $1/2$ each.

Partial automation. Our modeling of automation nests the case where a skill becomes *partially* automated; that is, only a fraction $\zeta_{\tau^\star} \in [0,1]$ can be automated, while the remaining fraction must be performed by human labor. We discuss this formally in Appendix A.1.

Wage effects. The automation shock leads to a change in the potential wage for a worker i in any occupation o:

$$\begin{aligned}\Delta w_{i,o,t} &= w'_{i,o,t} - w_{i,o,t-1} = \Delta\mu_{o,t} + (A'_o - A_o)s_i + \Delta\varepsilon_{i,t} \\ &= \Delta\mu_{o,t} + \underbrace{\overbrace{\frac{\alpha_{o,\tau^\star}}{LS_o}}^{\text{occ. exposure}} \cdot \overbrace{\left(\sum_{\mathcal{T}_l \setminus \tau^\star} \frac{\alpha_{o,\tau}}{LS_o - \alpha_{o,\tau^\star}} s_{i,\tau} - s_{i,\tau^\star}\right)}^{\text{worker specialization}}}_{\text{job transformation effects}} + \Delta\varepsilon_{i,t}\end{aligned} \tag{8}$$

where

$$\Delta\mu_{o,t} = \underbrace{\frac{\alpha_{o,\tau^\star}}{LS_o - \alpha_{o,\tau^\star}} \left(z_{\tau^\star} - \log r + \mu_{o,t-1}\right)}_{\text{productivity \& displacement effects}} + \left(\underbrace{\frac{\log P'_{o,t}}{LS_o - \alpha_{o,\tau^\star}} - \frac{\log P_{o,t-1}}{LS_o}}_{\text{GE effects}}\right)$$

Equation (8) captures two important terms. First, workers are more likely to see increases in their origin-occupation wage when $\Delta\mu_{o,t}$ is large. This effect captures the effects of automation in the canonical task-based model: negative displacement effects exert downward pressure on wages, positive productivity effects push wages up, and general equilibrium effects operate through changes in occupation-level output prices.

Second, the change in wages is driven by the interaction of shifting task weights and the worker's task-specific skills. We refer to the effect captured by this second term as *job transformation effect.* The effect of job transformation is ambiguous and itself depends on two terms. First, the *magnitude* of job transformation is in part governed by an occupation's *exposure* to automation, as captured by $\frac{\alpha_{o,\tau^\star}}{LS_o}$. Job transformation effects are generally larger in jobs that are more exposed to the automation shock; that is, occupations in which workers spend more time on the automated task before the shock. Second, the skill set of an individual worker governs both the magnitude and *sign* of the individual-level effects that a worker experiences from job transformation. Workers are more likely to benefit if they are relatively unskilled in the automated task relative to the remaining tasks, where the latter are weighted by the occupation-specific loadings.[18] However, workers may also be harmed by job transformation, which is more likely to occur when they are relatively skilled in the newly automated task.

To illustrate, return to the earlier example of financial analysts and consider two different workers. One of them excels at writing, is passable at financial modeling but lacks charisma when interacting with clients. Their colleague is equally proficient at modeling, excels in handling clients but struggles with writing. Equation (8) shows that the two analysts experience diverging wage changes as their job is transformed when report writing is automated. Both now spend more time on modeling and client interaction, but while the analyst adept at client interaction sees their productivity and earnings increase, their less gregarious colleague is likely to see a dip in theirs.

Finally, foreshadowing our quantitative results, we note that selection is likely to push toward negative wage effects for high-exposure occupations. By equations (5) and (7), "specialist" occupations that heavily load on one task tend to attract workers who are strongly specialized in this task. When such a task gets automated, workers in such occupations thus tend to lose.

2.3.2 The important role of task bundling

This is a good place to underline the important role of task bundling in our analysis: Job transformation effects distinctively arise when multiple tasks are performed concurrently within the same occupation.

Remark 3 (Task bundling.)**.** *An occupation features* ***task-bundling*** *if*

$$|\{\tau \in \mathcal{T}_l : \alpha_{o,\tau} > 0\}| > 1.$$

Conversely, the economy features a ***no-bundling property*** *if no occupation features task bundling:*

$$|\{\tau \in \mathcal{T}_l : \alpha_{o,\tau} > 0\}| = 1 \quad \forall o \in \mathcal{O}.$$

In a no-bundling economy, there exists an assignment function $g : \mathcal{O} \to \mathcal{T}$ that pins down the

[18]This mechanism mirrors the mechanism Freund (2025) documents in the context of team production among specialized workers: Your productivity is enhanced by a coworker—whether human or artificial—insofar as their presence enables you to focus on those task you are best at; the magnitude of this complementarity effect is increasing in the degree of skill specialization.

unique task required in any given occupation.[19] In this case, the wage equation reduces to

$$w_{i,o,t} = \mu_{o,t} + A_{o,g(o)} s_{i,g(o)} + \varepsilon_{i,t}. \tag{9}$$

In a no-bundling economy, workers in an occupation o subject to automation thus experience wage changes that are solely driven by changes in the occupation-specific shifter, i.e. $\Delta\mu_{o,t}$. The wage changes are, thus, driven by the well-understood balance between negative displacement effects, associated with a declining labor share, and positive productivity effects, driven by $\bar{z}_{\tau^\star}$.[20] Crucially, workers do not experience any effects from a changing task mix of their occupation. Moreover, conditional on staying in the same occupation, all workers in an occupation experience the same wage change. In contrast, under task bundling, individual wages change also because automation shifts the task content of their occupation, as described in equation (8).

3 Theory Meets Data

We now take the model to data. We begin by describing our estimation methodology (Section 3.1), its empirical implementation (Section 3.2), and validate the method in Monte Carlo exercises (Section 3.3). We then present the estimation results (Section 3.4) alongside an extensive verification of the model's empirical fit, both in steady state (Section 3.5) and for a historical episode of job transformation (Section 3.6).

3.1 Estimation methodology

Estimating the model parameters requires three data inputs: (i) a panel of workers, indexed by i, for whom both occupational choices $\hat{o}_{i,t}$ and wages $w_{i,\hat{o}_{i,t},t}$ are observed over time; (ii) measures of the occupational task weight matrix A, as defined in Remark 1, over time; and (iii) occupation-level labor shares LS_o. While (i) and (iii) are straightforward, (ii) is more involved and we explain how to construct A in Section 3.2.

Conditional on observing (i)-(iii), we make the additional assumption that the model is in one of two steady states throughout our estimation window, characterized by two different A

[19]A special case of this is the case where $A = I$, under which our model nests the standard Roy model. Note that occupations having the same A_o is not sufficient for them to be perfect substitutes from a worker's perspective, as they may involve labor shares or different machine tasks with differing productivities.

[20]For a detailed review see Acemoglu *et al.* (2025). A subtle difference in the operation of the positive productivity effects compared to the canonical model is worth noting. For example, in Acemoglu and Restrepo (2022, cf. equations (6) and (13)), the productivity effect raises the wages of *all* workers. What underlies this feature is the assumption that substitution across all tasks is governed by a uniform elasticity parameter. In contrast, in our model, production occurs at the level of occupations, so automation carries no positive productivity effects for occupations that do not utilize the automated task. However, spillovers may occur through the movement of occupational prices.

matrices with other parameters held fixed. The A matrix changes in the year 2000. Under these conditions, we can estimate the model with maximum likelihood. As the notation is involved, we relegate all details of the estimation and its implementation to Appendix B.1, where we also discuss our solutions to several challenges arising from the dimensionality of the parameter space, the constraints imposed by GE, and the presence of unobserved data. Here, we give a brief summary:

(1) We deal with the dimensionality of the parameter space and the presence of unobservable data by maximizing a stochastic approximation of the likelihood function instead of the full multidimensional integral.

(2) We exploit techniques widely used in training neural networks to speed up optimization: auto-differentiation computes gradients of the (approximate) likelihood and stochastic gradient descent further improves performance via batching.

(3) We use the implicit function theorem to augment the gradient, allowing us to account for parameter constraints imposed by GE.

Intuitively, we identify model parameters and prices $(\nu, \varsigma, \bar{s}, \Sigma_s, \vec{P}, \vec{P}')$ from a large sample of workers' revealed occupational choices and wages over time. Maximum likelihood inverts the mapping our model implies from the underlying distribution of skills to the joint distribution of occupational choices and wages across many periods. This approach does not require estimating any individual worker's skill set exactly, but it does yield posteriors on any given worker's skill set conditional on their observed data.

3.2 Data & measurement

As noted above, estimation requires three data inputs: a worker panel, the occupational task weight matrix A, and occupational labor shares. We use the National Longitudinal Survey of Youth 1979 (NLSY79) as our worker panel; construct the task weight matrix by applying natural language processing (NLP) tools and large language models (LLMs) to O*NET data; and construct occupational labor shares using data from the Bureau of Economic Analysis (BEA). We describe each in turn.

NLSY. The NLSY 1979 tracks 6,033 workers' occupations and wages between 1979 and 2018, comprising 110,618 observations. We construct an annual panel of each individual's primary job (if any), selecting among multiple jobs by weekly hours worked. We restrict the sample to worker-years with weekly hours exceeding 35 hours (full-time"), as our model does not include an hours margin. Wages are deflated using the CPI (1982–1984=100). We drop individuals in the military sample and the minority oversample. Following Lise and Postel-Vinay (2020), we create a harmonized occupational classification at the SOC-2000 level using crosswalks from Sanders (2012). We use the "minor groups" of occupations and restrict the sample to occupations with an employment share of at least 0.3% to facilitate the estimation of occupation-level objects, leaving us with 61 occupations.

Tasks & occupational task weights. We construct the occupational task weight matrix A defined in Remark 1 in two steps. First, we cluster approximately 19,000 detailed tasks with similar skill requirements using NLP. Second, we measure the weights different occupations attach to these task clusters using LLMs. Appendices B.2 and B.3 offer further details on both steps.

We start from 18,796 detailed, occupation-specific ("micro") tasks from O*NET. This starting point is natural: many technology-specific automation exposure measures reference this list, enabling us to leverage them in our analysis of automation shocks. We group micro tasks with similar skill requirements into clusters that serve as the empirical analogue to $\mathcal{T}_l$. The clustering involves a multi-stage NLP pipeline: We extract relevant features from the task statements, generate embeddings for these features, apply a hierarchical density-based clustering algorithm (HDBSCAN, McInnes *et al.* (2017)) to the embeddings, and, finally, use an LLM to create summary labels and concise descriptions for each cluster based on the micro tasks it contains. This yields 38 clusters with interpretable labels.[21]

Two examples illustrate the clustering. "Performing detailed manual tasks" groups micro tasks such as "lubricate moving parts on gate-crossing" and "smooth rough spots on walls using sandpaper"; its description emphasizes precise, hands-on operations that rely on manual dexterity and attention to detail. Meanwhile, "processing and analyzing records" groups tasks such as "prepare reports of activities, evaluations, recommendations, or decisions." and "prepare, examine, or analyze accounting records, financial statements, or other financial reports"; its description emphasizes numerical reasoning and analytical skills.

Theory guides the second step: measuring A. Equation (3) implies that each entry of the A matrix corresponds to the optimally chosen share of time allocated to task τ in occupation o. We measure these shares by prompting OpenAI's o3-mini-high to construct time allocation diaries for each occupation across our 38 task categories. We repeat this for pre-2000 and post-2000, yielding two period-specific A matrices. Appendix B.3 details our LLM prompts.

The resulting A matrix is visualized in Figure 1. It confirms substantial task bundling, and has intuitive properties. Only two occupations have a single task comprising more than half of incumbents' time, and in fewer than 30% of occupations does a single task account for more than a quarter of total time. Individual entries are likewise intuitive: "performing detailed

[21] Unlike k-means, where k is a user input, HDBSCAN automatically determines the number of clusters through a hierarchical approach based on cluster stability.

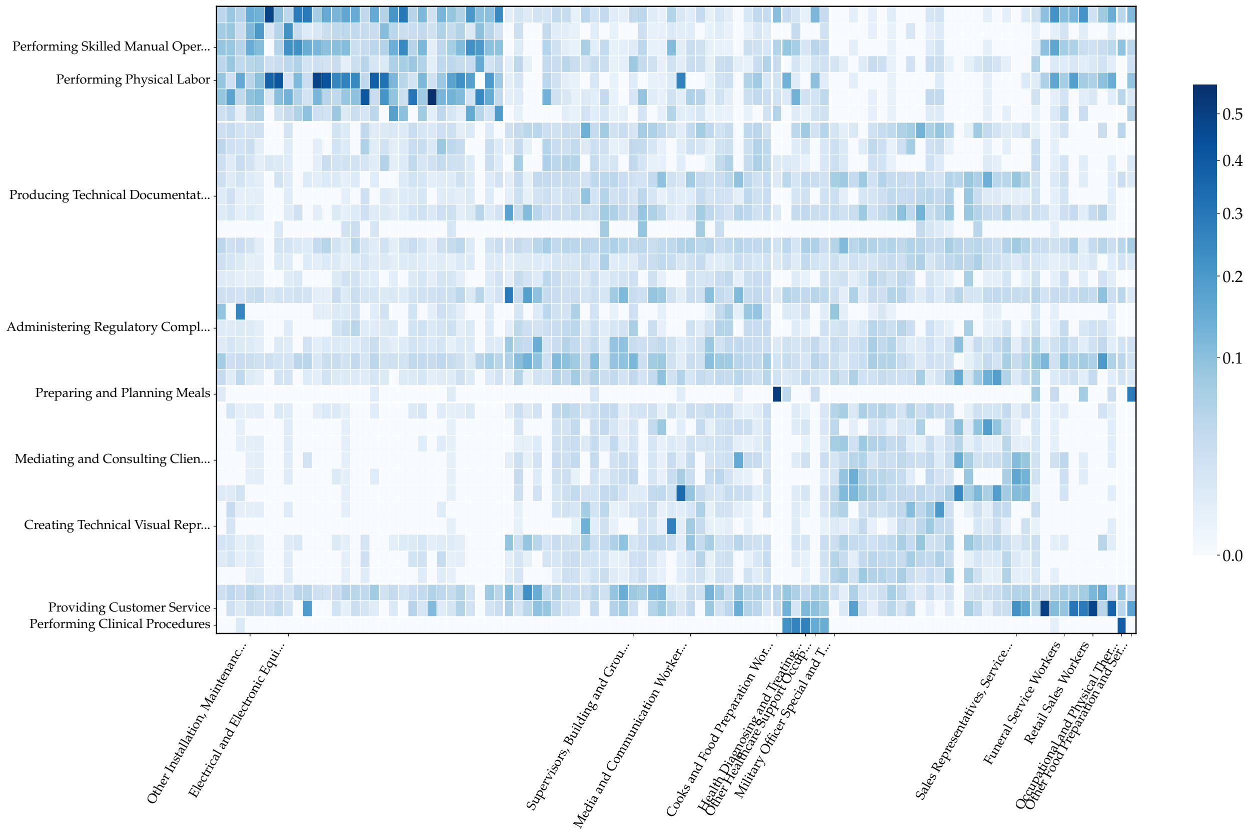


Figure 1: Task weight matrix

Notes. This figure shows the measured A matrix for the post-2000 years; each cell value corresponds to $\frac{\alpha_{o,\tau}}{\sum_{\tau \in \mathcal{T}_l} \alpha_{o,\tau}}$. To aid visualization, the matrix is reordered using a spectral co-clustering algorithm, and example tasks and occupations are highlighted.

manual tasks" appears as a prominent task across both service sector occupations, like "food and beverage serving workers" and manufacturing roles, like "assemblers and fabricators." However, the service-sector occupations additionally emphasize tasks like "providing customer service" while the manufacturing-oriented jobs involve more technical tasks like "operating, calibrating, and inspecting equipment."

This LLM-based approach to measuring A is flexible and easy to adapt to different task taxonomies or occupational classifications. But how reliable are the measurements? We conduct a battery of validation exercises that jointly corroborate the approach.[22] Most importantly, we compare LLM-generated task weights at the occupation-cluster level to the average importance rating that O*NET assigns to the micro tasks within each cluster. While O*NET importance ratings do not directly map onto our A matrix entries, unlike time shares, they are strongly

[22] Our use of LLMs here resembles deploying a vast pool of research assistants with unlimited time to gather diverse data and exercise judgment in constructing cardinal time shares—rather than consulting a metaphorical oracle to answer questions no human could answer.

correlated with our baseline measures. Appendix B.3.2 details these validation exercises.

Occupation-level labor shares. We construct LS_o by combining industry-level data on value-added and wage payments from the BEA-BLS Integrated Industry-level Production Accounts with data on wage payments at the industry-occupation level from the BLS Occupational Employment and Wage Statistics (OEWS) Tables. As value-added and hence labor shares are generally defined and measured at the industry-level, we apportion industry-level value-added to occupations based on their share in an industry's total wage bill. Appendix B.4 provides more details.

Elasticity of substitution. Finally, we set the occupational elasticity of substitution $\sigma = 2$ following Burstein *et al.* (2019), who estimate $\sigma \in \{1.81, 2.10\}$.

3.3 Validation: Monte-Carlo exercises

To verify that our estimation approach robustly identifies the parameters, we conduct a Monte Carlo exercise before turning to the main results. The exercise proceeds in three steps: First, we estimate parameters by applying our methodology to the data described in Section 3.2. Second, we generate artificial data from the estimated model under these parameters. Third, we apply our methodology to the artificial data and compare the resulting estimates with those from step one. If our method correctly recovers the data-generating parameters, the two sets of estimates should align closely.

The results corroborate our methodology: estimated parameters align well with the data-generating process ("dgp") in the simulation. Figure 2 illustrates this comparison, with each panel showing one set of estimated parameters. We split the skill covariance matrix into its correlation component (C_s, omitting the diagonal of ones in the figure) and its vector of standard deviations S_s; that is, we decompose it according to $\Sigma_s = \text{diagm}(S_s) \cdot C_s \cdot \text{diagm}(S_s)$. The remaining two panels show the estimated and data generating parameters for the vector of mean log skills, $\bar{s}$, and the parameters (ν, ς), respectively. The fit is generally good; in particular, we capture the large number of parameters governing bilateral skill correlations quite well.

3.4 Parameter estimates

For the scalar parameters, we estimate $\nu = 0.20$ and $\varsigma = 0.30$. The estimate of ν implies that reducing prospective wages in a given occupation by 1% lowers the odds of choosing it by about 5.1% since $\frac{1}{\nu} \approx 5.1$. $\varsigma = 0.30$ indicates that a one-standard-deviation occupation-specific random productivity shock can raise or lower wages by about 30% in a given year.[23]

[23]We take this to be a relatively large value. This estimate arises because the data features substantial period-to-period wage variation even for individuals who stay in their occupation. However, this is of no consequence for our

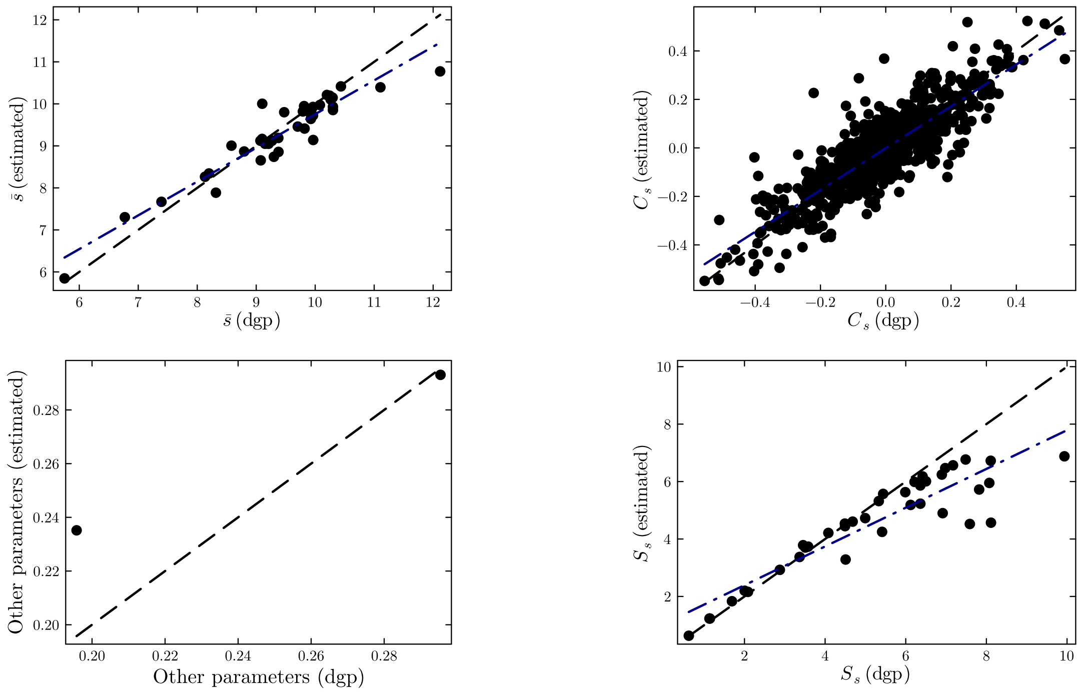


Figure 2: Monte-Carlo exercise: data generating parameters and their estimates

Notes. The horizontal axis displays parameter values used to generate artificial data. The vertical axis displays corresponding estimated values. "Other parameters" refers to the tuple (ν, ς). The black dashed line is the 45 degree line. The blue dash-dotted line is the line of best fit.

We next examine the mean and dispersion of skills. Figure 3 reports the estimated mean skills $\bar{s}$ with errors bands indicating their standard deviation (S_s where $\Sigma_s = \text{diagm}(S_s) \cdot C_s \cdot \text{diagm}(S_s)$). The values of $\bar{s}$ indicate tasks that are more or less productive on average in the population. The figure reveals a notable pattern: manual tasks such as "performing physical labor" or "performing detailed manual tasks" are associated with skills that are much less dispersed than more analytical tasks such as "coordinating strategic initiatives" or "analyzing and optimizing systems". We return to this pattern shortly.

Finally, we consider the estimated correlation matrix C_s of all pairwise task-specific skills (see Appendix Figure B.5 for an illustration). To gauge whether the estimated correlations are plausible, we compare them to evidence in the empirical literature, typically based on military enrollment test scores (Deming, 2017; Lise and Postel-Vinay, 2020; Girsberger *et al.*, 2022). This comparison requires aligning our task categorization with that used in the literature. We follow Autor *et al.* (2003) in distinguishing between five broader categories: non-routine analytical (NRA), non-routine interactive (NRI or "social"), non-routine manual (NRM), routine-cognitive (RC), and routine-manual (RM). Using an LLM, we assign each granular task τ to one of these

quantitative exercises below, since the realization of $\varepsilon_{i,t}$ does not have an impact on occupational choices or the relative occupation-specific wage prospects of workers.

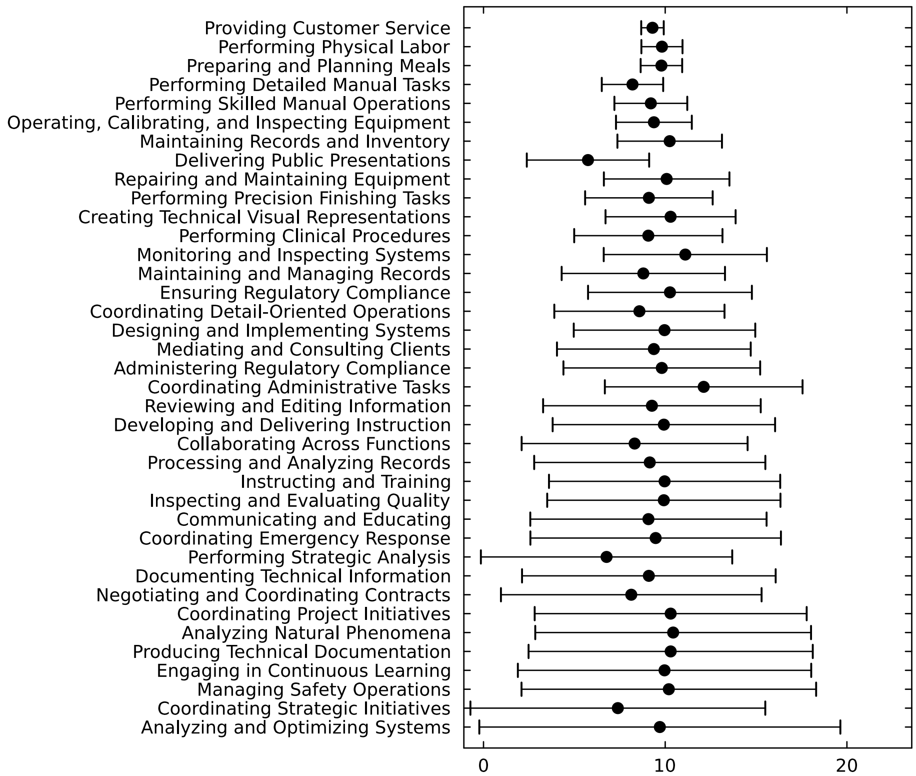
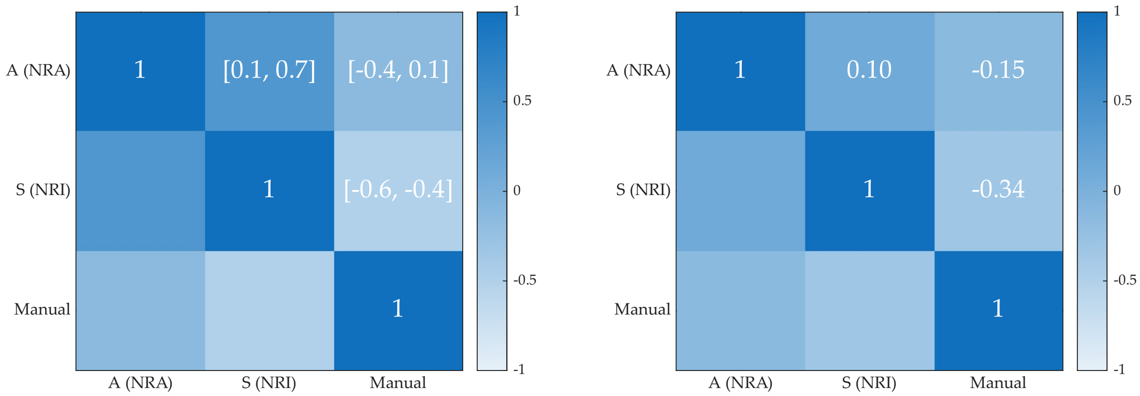


Figure 3: Skill means and dispersion

Notes. Estimated skill means $\bar{s}_\tau$ are indicated by dots. Error bands cover two standard deviations of estimated skill dispersion $S_{s,\tau}$.

five categories c. (The same categorization will resurface repeatedly going forward.) We then construct for each simulated worker i a skill index for each category c by averaging across tasks τ. This step is comparable to how the empirical literature constructs a skill index by averaging across test scores for multiple questions. Concretely, as in Deming (2017), we standardize each task-specific log skill $s_{i\tau}$ in the population, i.e., $\tilde{s}_{i\tau} = \frac{s_{i\tau}-\bar{s}_\tau}{\sigma_{s_\tau}}$; then average standardized skills within category c, $\bar{s}^{\text{raw}}_{ic} = \frac{1}{|\mathcal{T}_c|}\sum_{\tau\in\mathcal{T}_c}\tilde{s}_{i\tau}$; and, finally, re-standardize the index $S_{ic} = \frac{\bar{s}^{\text{raw}}_{ic}-\overline{\bar{s}^{\text{raw}}_{c}}}{\sigma_{\bar{s}^{\text{raw}}_{c}}}$.

Figure 4 reports the estimated correlations across analytical, social, and manual skill categories along with our best-effort reading of the empirical literature (also see Benzell and Myers (2026)). While the literature contains a fairly wide range for each correlation, our estimates match three important patterns. First, analytical and social skills are positively correlated. Second, analytical and manual skills have a negative to weakly positive correlation. Third, social and manual skills consistently show the most negative correlation.



(a) Literature **(b)** Model

Figure 4: Comparison of estimated skill correlations with the empirical literature

Notes. Social skills correspond to NRI, analytical skills to NRA, manual skills to the average of NRM and RM.

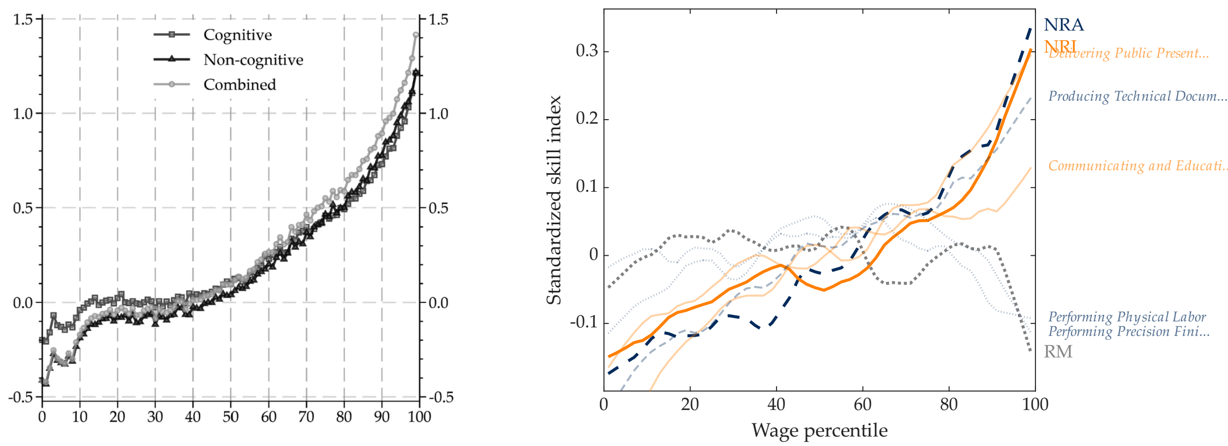
3.5 Steady-state model properties and validation

We now describe the steady-state moments implied by the estimated parameters and compare them to the data. All model moments are based on a simulation of 50,000 workers. Here we consider the model under the post-2000 A matrix; the following section evaluates how the model performs given *changes* in A.

Skills along the wage distribution. How do skills vary along the wage distribution? As an empirical benchmark, Bratsberg *et al.* (2025) draw on large-scale administrative data for men in Finland and Norway to establish a convex relationship between earnings rank and cognitive ability, steepening at the top (Figure 5a). A similar pattern holds for non-cognitive/personality traits. Figure 5b shows that our model matches these convex relationships, though it somewhat understates the steepening quantitatively. Unlike analytical and social skills, *manual* skills are relatively flat along the distribution—a contrast we return to when discussing the distributional effects of AI.

Occupational employment shares and wages. The estimated model matches several occupation-level outcomes well. First, it almost perfectly matches occupational employment shares (Appendix Figure B.8a), notably without requiring occupation-level amenity shifters. Second, the correlation between average occupational wages in the data and the model is 0.59 (Appendix Figure B.8b). Third, turning to wage dispersion, the standard deviation of log wages is 0.54 in the data and 0.57 in the model.

Occupational choice. Workers in the model sort into occupations on the basis of task-level comparative advantage, choosing jobs that emphasize tasks where they possess relative skill advantages. How do the resulting patterns of occupational switching compare to the data?



(a) Data

(b) Model

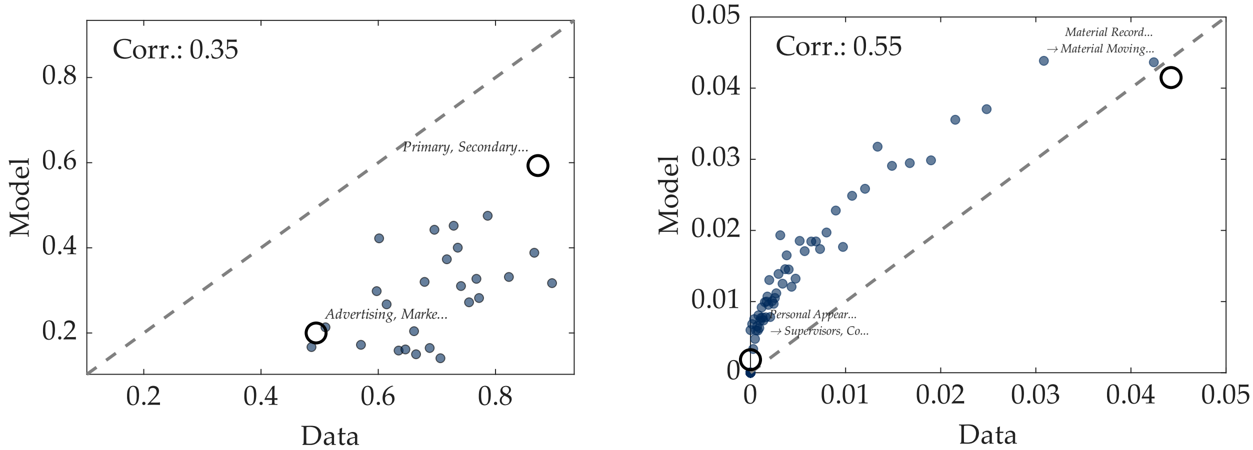
Figure 5: Skills along the wage distribution

Notes. The left panel is Figure A4 (panel B) from Bratsberg *et al.* (2025). The right panel is the model counterpart.

Figure 6 compares model-implied occupational transition matrix entries to the NLSY. The model generates positive correlations for both staying probabilities (diagonal elements, 0.35) and switching probabilities (off-diagonal elements, 0.55). However, it notably under-predicts occupational persistence: the average annual staying probability is 31%, well below the 70% in the NLSY. The likely reason is that transitions in the model are driven purely by comparative advantage and preference shocks. We could close this gap by incorporating exogenous switching costs or frictions, but our approach transparently reveals how much the task-based model can explain endogenously.

An important test of our task-based model concerns its distinctive predictions for the *direction* of occupational switches. We replicate the empirical analysis in Gathmann and Schönberg (2010), who show, using German panel data, that workers are more likely to move to occupations with similar task requirements. This pattern cannot be rationalized by models with one-dimensional (general) or occupation-specific skills, but follows naturally from task-specific skills. We compare the realized distribution of between-occupation distances in task space to a random-mobility benchmark where mobility is governed solely by the relative size of destination occupations. Distance between occupations o and o' is one minus the angular separation of their task-weight vectors, A_o and $A_{o'}$. Figure 7a shows that, in the NLSY, observed distances concentrate more heavily at short distances than does the random-mobility benchmark, consistent with Gathmann and Schönberg (2010). (The peak at long distances in the benchmark reflects that many occupation pairs are far apart in task space.) Crucially, as Figure 7b shows, the model replicates this pattern.

The model also matches evidence that specialization generates persistence in occupational choice (Kambourov and Manovskii, 2008; Geel *et al.*, 2011). Workers with more specialized

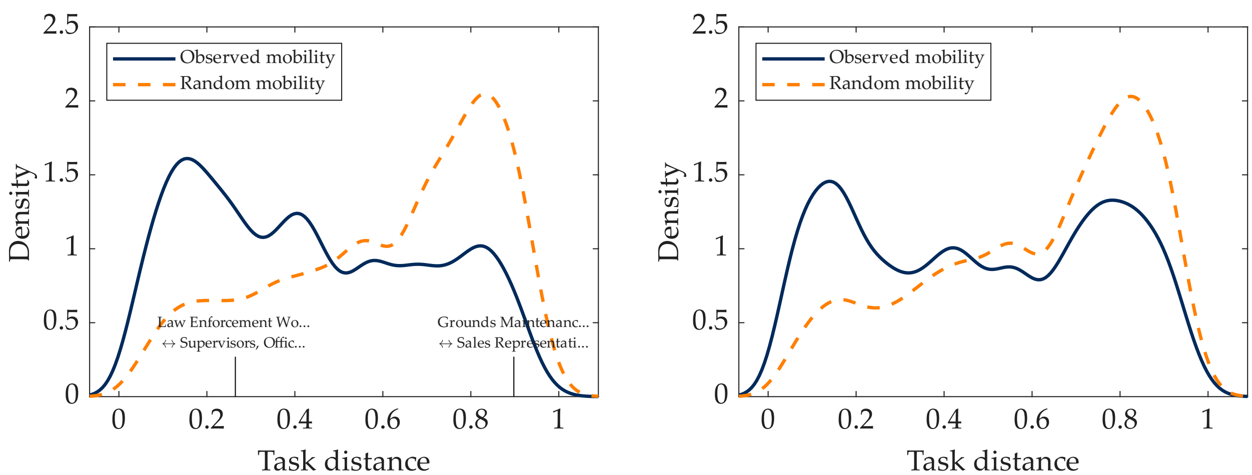


(a) Staying probabilities

(b) Switching probabilities

Figure 6: Occupational transition matrix: model vs. data

Notes. This figure compares the model-generated entries of the transition matrix to those derived from the NLSY. The left panel is a binned scatter plot of diagonal entries, the right panel a binned scatter of off-diagonal entries.



(a) Data

(b) Model

Figure 7: Task-specific skills shape the direction of moves in data & model

Notes. Both panels plot the observed density of distances conditional on switching occupation (solid line) and that under a random-mobility benchmark (dashed line). The left panel is based on the NLSY, the right panel on model-generated data.

skills move less frequently. We measure skill specialization as the within-worker coefficient of variation of skills. Appendix Figure B.9 shows that greater specialization is associated with a lower probability of switching occupations.

3.6 Historical validation: the case of RBTC

The moments considered so far characterize the model's steady state. Since our goal is to analyze *changes* in A (i.e., job transformation), it is important to also verify the model's performance in this context. To this end, we study a historical episode of job transformation: "routine-biased technological change" (RBTC), the tendency for technology to replace labor specifically in routine tasks (e.g., Autor *et al.*, 2003; Goos *et al.*, 2009, 2014; Autor and Dorn, 2013). We first verify that our model, and specifically the measured A matrices, capture RBTC, then examine the consequences of RBTC for employment and wages by replicating two widely cited empirical findings inside our structural model.

RBTC in the data & model. There are two ways of documenting RBTC. First, and most directly, RBTC manifests as a change in the A matrix, i.e., changes in task mix *within* occupations. Figure 8 displays these changes from the pre-2000 to the post-2000 period, aggregating granular tasks into the five aforementioned broad categories widely used in the RBTC literature. The rising importance of non-routine tasks, especially interactive and analytical, and the corresponding decline in routine tasks, is clearly visible. These shifts are consistent with systematic empirical studies (Atalay *et al.* (2020, e.g., Table 3); Spitz-Oener (2006, Table 5)) demonstrating the central role of within-occupation changes in explaining the economy-wide shift from routine to non-routine analytic and social tasks.

Second, our model is also consistent with earlier work establishing large changes in the economy's task use through shifts in employment shares *between* occupations. Following Autor *et al.* (2003, Fig. 1), one way to illustrate this channel is to rank demographic groups by their intensity in each task (category), assign percentiles based on their position in the employment distribution, freeze these percentile assignments, and track how the employment-weighted average shifts as workers reallocate. Figure 9 shows the data from Atalay *et al.* (2020) on the left and the model counterpart on the right. While the timelines do not match exactly, the trends are sufficiently monotone to permit comparison. Despite using entirely different data sources for both occupational task intensities and employment shares, the model matches the empirical trends remarkably well.

Employment polarization. An influential literature argues that RBTC drove the simultaneous growth in employment in both the highest-pay and lowest-pay occupations, offset by declining employment in the middle of the distribution, observed in the late 20$^{\text{th}}$ and early 21$^{\text{st}}$ century.

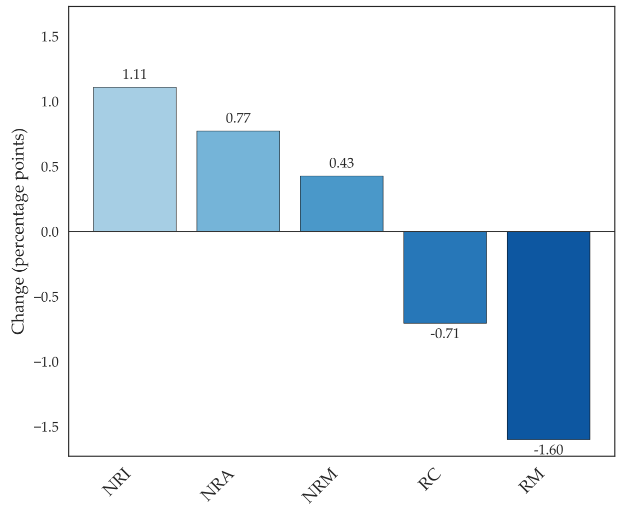
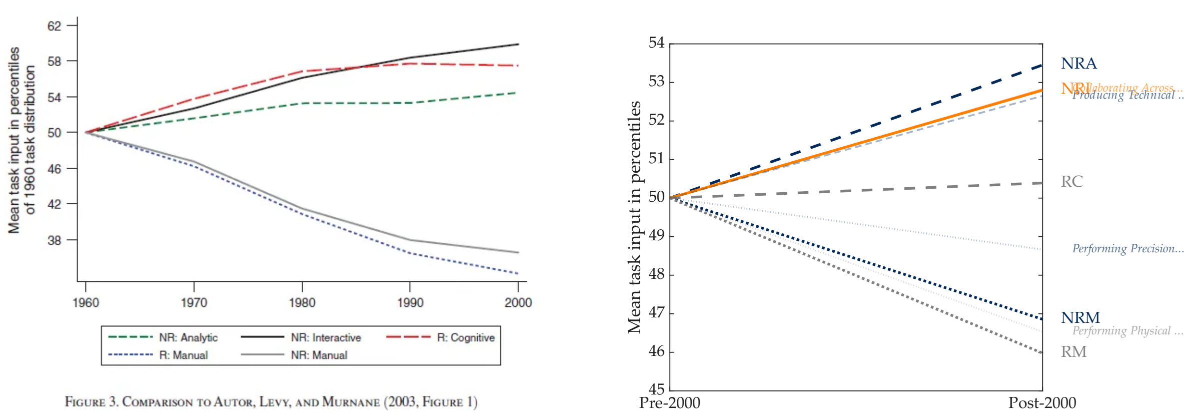


Figure 8: RBTC as captured by the changing A matrix

Notes. The granular tasks are assigned to one of the following five categories: non-routine analytical (NRA), non-routine interactive (NRI or "social"), non-routine manual (NRM), routine-cognitive (RC), and routine-manual (RM). The percentage changes are unweighted averages across occupations and granular tasks within each category.



(a) Data (Atalay *et al.*, 2020) **(b)** Model

Figure 9: RBTC in data & model

Notes. The left panel shows Atalay *et al.* (2020, Fig. 3), which is based on task measures constructed off of job ads. The authors compute percentiles of demographic groups' task averages based on their 1977 task content, then compute the mean employment-weighted percentile for each year between 1960 and 2000, taking 1960 employment shares as the baseline. The right panel performs a comparable exercise using model data, comparing the post-2000 and the pre-2000 steady state of the model.

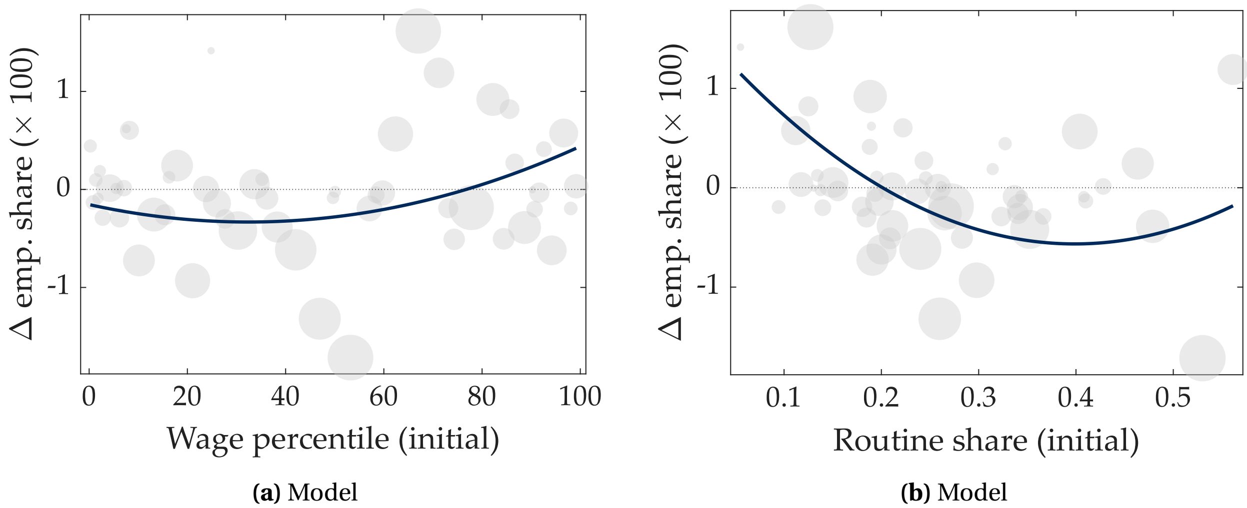


Figure 10: RBTC-driven employment polarization in the model

Notes. This figure shows the change in occupational employment shares between the pre-2000 and the post-2000 steady state of the model. "Routine" tasks comprise model tasks assigned to either the routine-cognitive or the routine-manual category. The solid line indicates the quadratic best-fit regression line.

Appendix Figure B.10 shows the evidence on this "polarization" pattern from Goos *et al.* (2009) and Autor and Dorn (2013). Figure 10b verifies that the model replicates this non-monotone growth of employment by initial wage percentile (left panel), which is indeed driven by the rise and decline of low-routine and high-routine jobs, respectively.

Shifting skill returns. Turning to wages, Deming (2017) influentially argues that the labor market return to social skills was greater post-2000 than pre-2000, offering RBTC as one plausible driver. Figure 11 displays Deming's (2017) estimates for the change in log wage associated with a one standard deviation increase in social skills, based on NLSY data, as orange cross markers. Plus markers display estimates from Edin *et al.* (2022), who use high-quality Swedish administrative data. Both papers find an increased return to social skills, though the levels differ. In blue are the estimated returns to analytical skills, which Deming (2017) finds to have decreased, whereas Edin *et al.* (2022) find an increase, though followed by a mild decline in the 2000s.

We use our model-based skills estimates and wage equation to construct a direct counterpart to these empirical estimates. First, we construct normalized skill indices as averages across task-specific skills, as described already in Section 3.4 and exactly following Deming (2017, see esp. Sections III.B and III.C), who averages over multiple survey questions. Denoting each task category by c, we then estimate the wage return as the OLS coefficient β_c in the regression

$$w_{it} = \gamma_t + \sum_{c \in C} \beta_c S_{ic} + \epsilon_{it}, \tag{10}$$

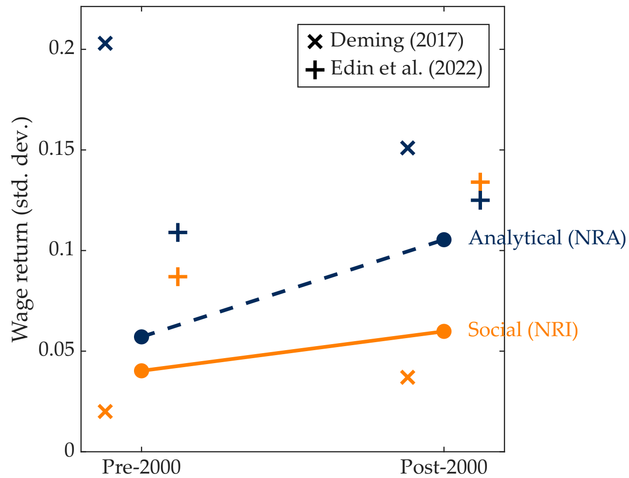


Figure 11: Historical changes in skill returns

Notes. The vertical axis indicates the change in log wage associated with a one standard deviation increase in the respective skill measure. The estimates indicated for Deming (2017) correspond to those reported in Table IV, where pre-2000 corresponds to the estimates for the NLSY79 and post-2000 to those for the NLSY97. The estimates indicated for Edin *et al.* (2022) are averages over the year-specific returns displayed in their Figure 1.

which is identified off of cross-sectional variation in skills and wages.

Figure 11 shows, as a solid orange line, that our model implies a rise in the return to social skills (i.e., skills in NRI tasks), rising from around 3% to 5.5% and, thus, falling squarely between Deming's (2017) and Edin *et al.*'s (2022) estimates. The model also implies a rising return to analytical skills, differing from Deming (2017) but consistent with Edin *et al.* (2022).

In summary, the estimated model fits the data well in terms of both steady-state moments and the historical effects of routine-biased technological change.

4 The Labor Market Effects of LLMs

This section presents the main results of the paper: We use the estimated model to project the labor market effects of automation due to large language models (LLMs) and document the important role played by job transformation. Section 4.1 explains how we leverage task exposure measures to pin down which model tasks LLMs can automate. Section 4.2 then presents three sets of results, each offering a complementary perspective on the winners and losers of AI-induced automation: in terms of exposure; in terms of skills; and in terms of one's initial position in the wage distribution.

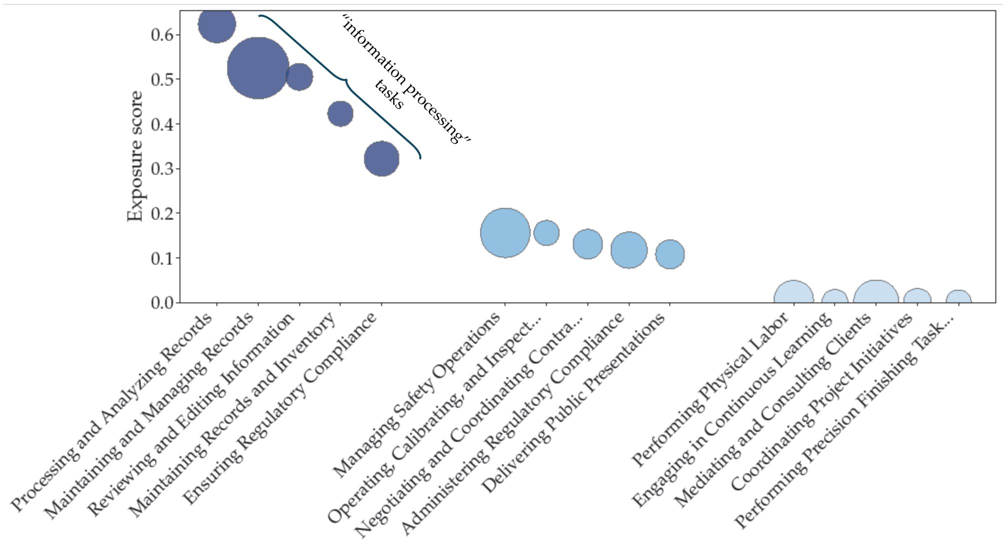


Figure 12: Eloundou *et al.* (2023) exposure scores aggregated to task clusters

Notes. This chart shows the exposure of the task clusters to LLM automation based on the share of detailed tasks they contain which are rated as fully or almost fully automatable by Eloundou *et al.* (2023). The size of each bubble indicates the number of detailed O*NET tasks contained in each cluster.

4.1 Construction of automation shocks

To quantify automation-induced earnings effects for technologies currently being rolled out or will be adopted in the future, we need to know which tasks the technology automates. Unlike backward-looking studies, we cannot rely on labor share changes to measure automation in industries or occupations. Moreover, even if such changes could be constructed, they would not reveal which specific tasks within occupations face automation, as required for an analysis of job transformation effects.

Mapping to task exposure measures. Our methodology allows us to overcome this challenging by providing a direct mapping between our model tasks—constructed from detailed O*NET tasks—and empirical measures of technology-specific, task-level automation exposure that a very active literature constructs from a variety of data sources, ranging from patents (Webb, 2019) and capability-specific AI benchmarks (Felten *et al.*, 2021) to expert and machine judgment (Eloundou *et al.*, 2023). Our model can link directly to any exposure measure at the detailed O*NET task level.

Motivated by the rapid diffusion of large language models (LLMs) with increasingly advanced capabilities (Bick *et al.*, 2024), we focus on LLMs. To identify which tasks are most likely automated through LLMs, we draw on Eloundou *et al.* (2023), who quantify the exposure to LLM automation for each O*NET detailed task using human labeling and GPT-4 classifications. We aggregate their scores to our task clusters by measuring, for each cluster, the share of micro tasks it contains that Eloundou *et al.* (2023) classify as either fully or almost fully automatable. Figure 12 shows the resulting exposure scores for our task clusters, ordered by descending exposure. The most exposed categories are "Processing and Analyzing Records" and "Maintaining and Managing Records," followed by "Reviewing and Editing Information" and "Producing Technical Documentation." The first two clusters include detailed tasks such as "Compute payment schedules" and "Prepare reports showing places of departure and destination, passenger ticket numbers, [...]" or "Maintain and update human resources documents, such as organizational charts, [...]" and "Organize archival records and develop classification systems to facilitate access to archival materials," respectively.

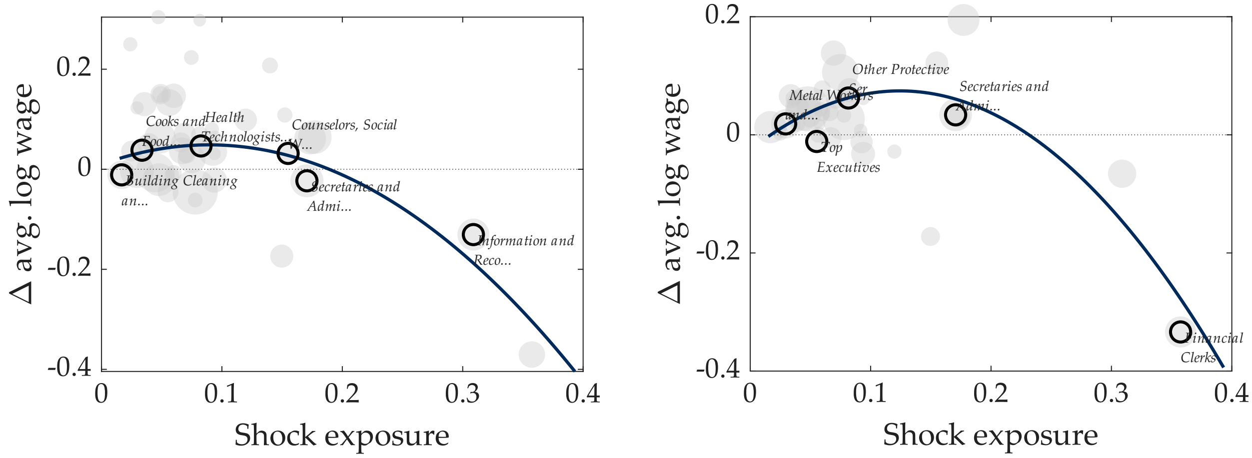
To provide some context for these LLM exposure scores, we compare them to exposure scores for industrial robots taken from Webb (2019). Webb (2019) constructs task automation exposure measures using the overlap between job task descriptions and patent text.[24] This methodology points to "Performing Detailed Manual Tasks" as the most robot-exposed task cluster—comprising detailed tasks such as "Lubricate moving parts" and "Remove excess materials or impurities from objects, using air hoses or grinding machines"—followed by "Performing Physical Labor," which includes tasks like "Hammer out bulges or bends in metal workpieces" and "Dump refuse or recyclable materials at disposal sites." Appendix C.2 provides more details. In short, different automation technologies affect very distinct sets of tasks.

Automation scenario. The scenario we consider is that the five tasks most exposed to LLMs are automated. More specifically, we suppose that these five tasks are partially automated, where for each task the fraction automated, $\zeta_{\tau^\star}$, corresponds to the score indicated in Figure 12. Importantly, we suppose that machine productivity in each automated task, $z_{\tau^\star}$, is at its automation threshold, defined in Section 2.3.1. This means that $z_{\tau^\star}$ is *just* high enough to make the automation of task $\tau^\star$ optimal, so the scenario we are considering can be interpreted as a lower bound on average productivity and wage effects.

4.2 Results

Who wins and who loses from AI-induced job transformation? We organize our answer along three dimensions. First, exposure: does working in a more exposed occupation predict whether a worker wins or loses? Second, skills: which task-specific skills help workers weather the shock? Third, income: does the shock benefit low-earners or high-earners more? Throughout, we focus on *individual*-level effects and show *occupation*-level outcomes in Appendix C.3.1.

[24] Many thanks to Mike Webb for sharing the exposure scores at the detailed-task level.



(a) All workers　　**(b)** Stayers

Figure 13: Average wage effects of AI exposure are non-monotonic

Notes. The vertical axis indicates the average change in log wages between the year following the shock, t, and the preceding period, $t-1$. Each dot represents an occupation o and averages across the outcomes for all individuals who chose o in year $t-1$. The horizontal axis indicates the LLM exposure of occupation o. Dot sizes correspond to pre-shock employment shares. The solid line is the quadratic, weighted line of best fit.

4.2.1 Exposure

We begin with a simple question: when an occupation is more exposed to AI, are workers in that occupation more likely to lose from the shock?

We classify individuals by their pre-shock occupation and plot the average change in log wages against the effective exposure of the pre-shock occupation, measured as the sum of relative task weights on automated tasks, i.e., $\sum_\tau \zeta_{\tau^\star} A_{o,\tau}$. We initially consider all individuals regardless of their post-shock occupation; later we distinguish between stayers and leavers.

Average effects are non-monotonic. Our first finding is that the *average* individual-level wage effects of exposure are non-monotonic: moderate exposure is associated with wage gains on average, but high exposure tends to lead to large losses. Consider Figure 13a. Workers in occupations with effectively no exposure at all tend to experience only small, though positive, wage changes.[25] At moderate exposure (around a tenth of tasks automated) the average wage change is positive, around 4%, and considerably more for some occupations. At high exposure, however, workers tend to experience heavy losses, on the order of 15% to nearly 40%. Note that

[25]The near-zero effect for occupations with effectively no exposure, and thus the non-monotonicity in effects, represents a notable difference to what the model in, say, Acemoglu and Restrepo (2022), would predict, and specifically in the operation of positive productivity effects. In Acemoglu and Restrepo (2022, cf. equations (6) and (13)), the productivity effect raises the wages of *all* workers. What underlies this feature is the assumption that substitution across all tasks is governed by a uniform elasticity parameter. In contrast, in our model, production occurs at the level of occupations, so, in partial equilibrium at least, automation carries no positive productivity effects for occupations that do not utilize the automated task. General-equilibrium effects can of course alter this result but, in our quantitative analysis, do not appear to do so to a quantitatively significant degree.

these are *individual*-level outcomes. Appendix Figure C.3 shows occupation-level outcomes, i.e., outcomes that are partly driven by changes in skill composition following the shock.

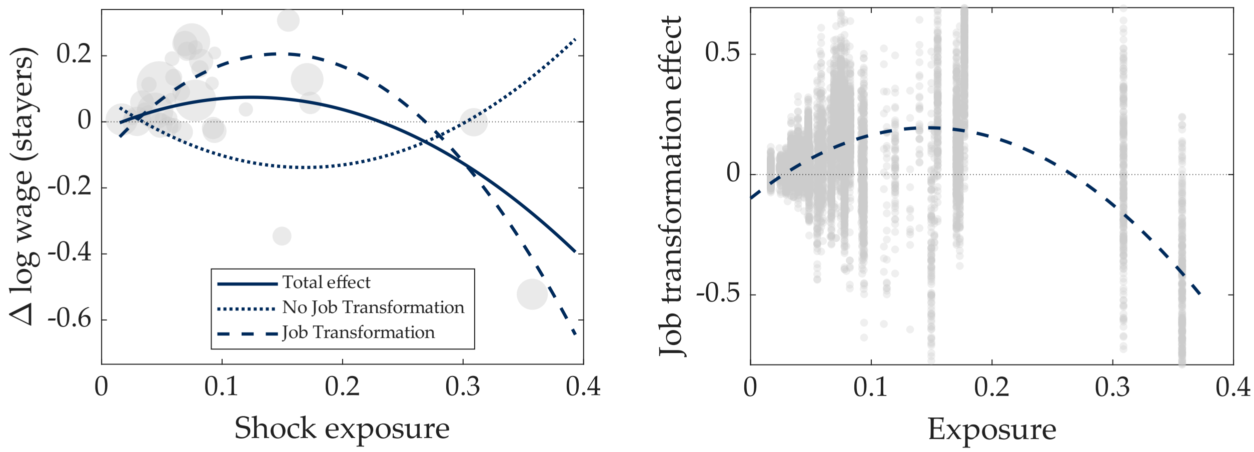
Though reliable evidence on the wage effects of LLM automation is still scarce, we note that our findings are consistent with—and rationalize—the empirical study by Eisfeldt *et al.* (2023), who show that occupations whose "core" tasks are automatable by LLMs experienced declines in labor demand following the release of ChatGPT, whereas occupations whose "supplemental" tasks were automatable saw positive effects on employment and wages.

Importantly, it is job transformation that gives rise to this distinctive, inverted-U shape. To see this, we zoom in on stayers, i.e., individuals who chose the same occupation o before and after the shock, which allows us to abstract from wage changes arising from occupational switching. This narrowing of focus does not alter the main result: as Figure 13b shows, the average-effect curve for stayers is very similar to that for all incumbents. Figure 14a decomposes the total effect for stayers. The dotted line labeled "No Job Transformation" represents a counterfactual economy in which the automation shock alters machine productivity and labor shares as in the baseline yet holds occupational task weights A fixed. The shock thus induces the same productivity and displacement effects as in the baseline (as well as the corresponding GE effects), but does not induce job transformation. Absent job transformation, the inverted-U shape disappears. Instead, the job transformation term (the dashed line) generates the overall shape.

What lies behind this result? The selection forces discussed in Section 2.3 are central. As workers select into occupations by comparative advantage, those choosing occupation o tend to have especially high skills in tasks used more intensively in o. Incumbents of highly exposed occupations, which rely heavily on automatable tasks, therefore tend to be specialized in precisely those tasks and lose from the shock on average. Conversely, if an occupation is only moderately exposed, workers spend a smaller share of their time on automatable tasks and tend to be more specialized in other tasks, which are more central to their occupation. Here, automation frees them to spend more time on what they are best at, leading to productivity and wage gains.

Exposure generates *potential* for change. The preceding paragraphs describe the *average* experience of incumbent workers, but a central feature of our quantitative estimates is that that outcomes are *heterogeneous*, even conditional on workers' initial occupation. Figure 15a plots the kernel density of wage changes for occupational stayers. While the average wage change experienced by stayers in heavily exposed occupations is lower than that incurred by workers in less exposed occupations, there is a great deal of individual-level variation in both tails.

Behind this individual-level variation in wage changes lies, once again, job transformation. This follows immediately from equation (8), noting that $\Delta\mu_{o,t}$ is identical for every stayer within an occupation. Figure 14b plots the individual-level variation by plotting the wage changes due to job transformation for stayers against occupational exposure. Dispersion grows in exposure. This



(a) Decomposition

(b) Individual job-transformation effects

Figure 14: Job-transformation effects drive both shape and dispersion

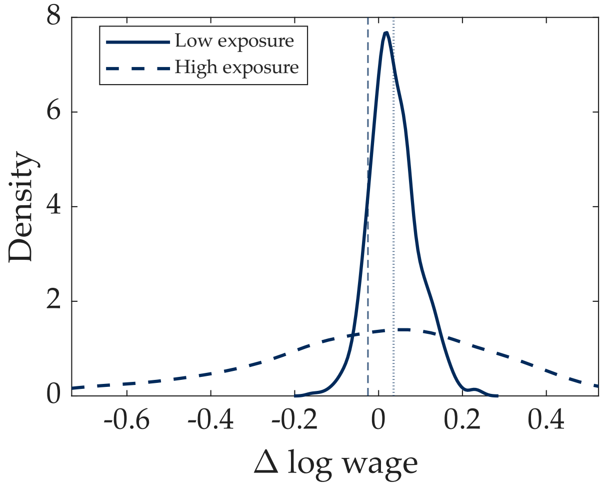
Notes. The left panel decomposes the average wage changes by exposure from Figure 13b. The right panel displays the variation in individual-level wage changes due to job transformation by occupational exposure.

result reflects the logic of equation (8): Exposure governs the magnitude of job transformation effects and thus captures the *potential for change* in a worker's earnings, but for any two workers in a given exposed occupation the wage effects of job transformation may be quite different and could be positive or negative.

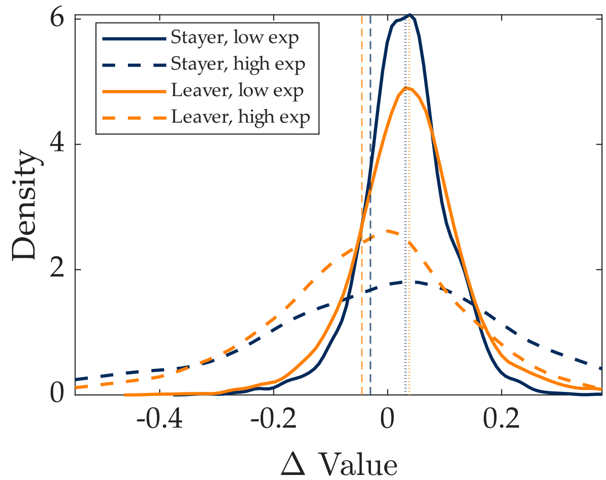
Outcomes are just as heterogeneous for leavers as for stayers. Since wage changes for leavers, unlike those of stayers, are partly driven by idiosyncratic preference shocks, we compare groups using the change in the *expected utility value* $V_{i,t}$ that a worker enjoys given their wage prospects across all occupations, before and after the shock. By the properties of the Gumbel distribution and equation (2), this value is given by $V_{i,t} = \nu \log\left(\sum_{o \in O} \exp\left(\frac{w_{i,o,t}}{\nu}\right)\right)$. This measure evaluates outcomes independently of a worker's current occupation. Figure 15b shows the distribution of value changes for four groups: stayers and leavers in low- and high-exposure occupations, respectively. Dispersion is greater for leavers in high-exposure than low-exposure occupations, paralleling the pattern for stayers.

Some individuals are "trapped" and incur large losses. Figure 15b also highlights that some workers lose severely. To see why, observe that when a task is automated, it simultaneously affects all occupations that use it intensively, not just one. Consider a worker skilled in the automated task whose potential wage drops in their current, task-intensive occupation. In principle, the ability to switch occupations serves an insurance function.[26] But the simultaneous transformation of multiple jobs limits the value of this insurance. As Figure 7 showed, workers

[26]Computing each leaver's counterfactual wage had they stayed in their origin occupation, holding prices fixed, shows that the average leaver's wage would be 5.67% lower had they not switched.



(a) Stayers by exposure



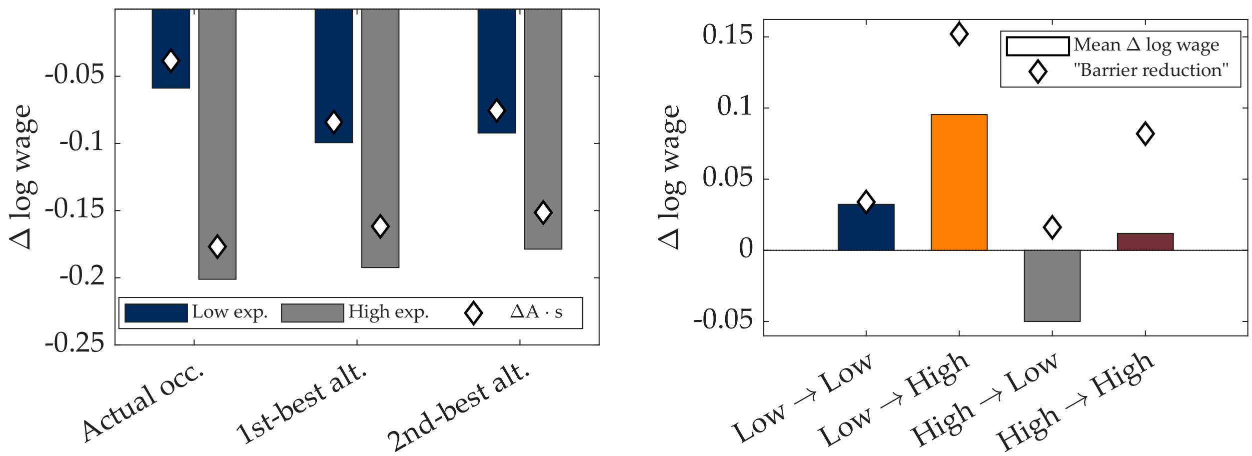
(b) Value change for stayers and leavers

Figure 15: Distribution of individual changes by subgroups

Notes. The left panel plots the kernel density of log wage changes for stayers, distinguishing among low- and high-exposure occupations, defined as worker-weighted bottom and top quartiles. The right panel distinguishes between occupational leavers and stayers, as well as exposure, and plots the distribution of value changes, as defined in the main text.

tend to move to occupations with similar task requirements, where their skills transfer most readily. But occupations with task requirements similar to the worker's origin occupation are also transformed similarly. Figure 16a illustrates this "trap": zooming in on stayers with negative wage changes, it displays the wage change in both their actual occupation and their first- and second-best alternatives (pre-shock). The wage declines similarly in all three occupations. This pattern applies to all stayers, but especially those in high-exposure occupations. The correlated decline in potential wages across likely occupations is driven by job transformation, as shown by the diamond markers, which represent the average change in potential wages associated with job transformation, i.e., $\Delta A_o \cdot s_i$. Simply put, automation can leave exposed workers with nowhere to go.

Workers switching into high-exposure occupations tend to gain. Some of the biggest winners are leavers who were previously deterred from highly exposed occupations by skill barriers in now-automated tasks. Following automation, these workers switch into the transformed roles. To illustrate, we split occupations into below- and above-median exposure and compute the average wage change for leavers across the four combinations of origin and destination occupation. Figure 16b shows that the largest gains accrue to workers moving from low-exposure into high-exposure occupations. These occupations undergo substantial changes in task content, with a diminished role for information-processing tasks, and thus attract workers whose skills align



(a) Stayers' deteriorating outside options

(b) Switching into high-exposure occs. yields gain

Figure 16: Job transformation drives heterogeneity in worker effects

Notes. The left panel considers the sub-sample of stayers experiencing negative wage changes, distinguishing between those in below-median and above-median exposure occupations. It shows the potential wage change in both the actual occupation and the two next-best alternatives (based on pre-shock choice probabilities). The diamond markers show the change in the potential wage due to ΔA. The right panel considers the sub-sample of leavers, splitting them into four groups based on origin- and destination-occupation exposure. The "barrier reduction" effect describes the change in the potential wage in the destination occupation due to ΔA.

better with the new task profile.[27] We formalize this as a "barrier reduction" effect: $\Delta A_{\hat{o}'_i} \cdot s_i$, where $\hat{o}'_i$ is the occupation chosen after the shock. This term captures how much more attractive the destination occupation has become for a given worker due to job transformation. Its magnitude is by far the largest for workers moving from low- to high-exposure (transformed) jobs. For example, a frequent transition is from health practitioners to health technicians. Post-automation, the latter occupation involves fewer information-processing tasks, like maintaining, processing and analyzing data, while the weight on hands-on technical tasks, like operating, calibrating, and inspecting equipment, has increased. Newly minted health technologists, previously deterred by a lack of information-processing skills, experience wage gains as their new job has rebalanced toward their strengths.

4.2.2 Skills

The preceding analysis reveals large and heterogeneous wage changes, particularly in high-exposure occupations. What characteristics predict who wins and who loses? Our second set of results provides an answer in terms of the skills they possess. We first show that AI-induced job

[27]This change in labor supply to exposed occupations has parallels with the idea in Autor and Thompson (2025) that a drop in expertise requirements enlarges the set of potential workers who can perform the occupation's remaining tasks. In our model, the shift in supply arises not because overall expertise requirements change but because specific task requirements are altered. Hosseini Maasoum and Lichtinger (2026) document empirical evidence consistent with these labor-supply effects.

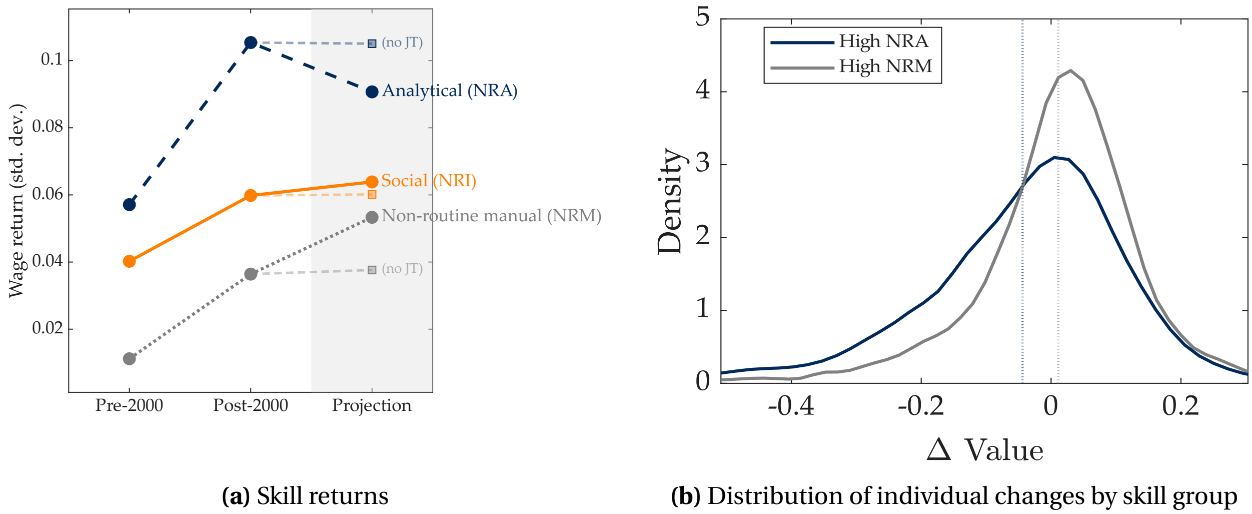


Figure 17: Automation effects on skill returns

Notes. The left panel shows the model-estimated return to different categories of task-specific skills as well as the projected return in the automation scenario. The transparent lines show returns for the hypothetical scenario of holding A fixed. The right panel plots the change in value for individuals in the top quartile of NRA and NRM skills, respectively.

transformation alters the return to different types of skills. We then show that specialization in analytical versus non-routine manual skills is a strong predictor of whether a worker is likely to win or lose from the AI shock.

Figure 17a projects the returns to aggregated skill types after the automation shock, extending the historical analysis from Section 3.6. Three findings stand out. First, the returns to social skills continue to increase as a result of the shock. Workers with strong skills in "negotiation" or "delivering public presentations", for instance, are expected to benefit. Second, the returns to cognitive and analytical skills, especially those in data analysis, are projected to decrease. Third, the return to non-routine manual skills—like the operation, calibration and maintenance of equipment or precision measurement—increases. These are the sorts of skills required in the "skilled trades": hands-on occupations requiring specific technical knowledge and training. Job transformation is central to this result, too. As the transparent lines in Figure 17a indicate, the "no job transformation" counterfactual, which holds A fixed, leads to near-zero changes in skill returns.

These changes arise in large part because LLMs chiefly automate information-processing tasks classified as "analytical." Skills in such tasks are widely dispersed (Figure 3); as a result, these skills command high compensation and are over-represented at the top of the wage distribution in the pre-automation steady state (see Figure 5b). After the shock, workers who excel at such tasks see the returns to their skills decrease, as machines take over a portion of analytical tasks. Meanwhile, those whose comparative advantage lies in social and non-routine manual tasks benefit, as they no longer need to perform the automated tasks, which were never their strength.

As an important corollary, skills help predict who loses from AI-induced job transformation and who wins. Workers with high analytical skills are over-represented among losers, while those with high non-routine manual skills are over-represented among winners. Figure 17b illustrates: it plots the kernel density of value changes for workers with above median non-routine analytical skills and for workers with above median non-routine manual skills. There is a large left tail of workers in the former group who sees value losses of 30-40%. In contrast, workers in the latter group win on average and are unlikely to experience losses in excess of about 20%.

4.2.3 Distribution

Our third set of results addresses a question that has attracted considerable attention: Do high-wage or low-wage workers tend to experience larger gains or losses? In other words, is the LLM-driven automation shock progressive or regressive? While some have argued that, similar to past automation shocks, AI will exacerbate wage inequality (Acemoglu and Restrepo, 2022), others have suggested that AI might in fact "rebuild the middle class."[28]

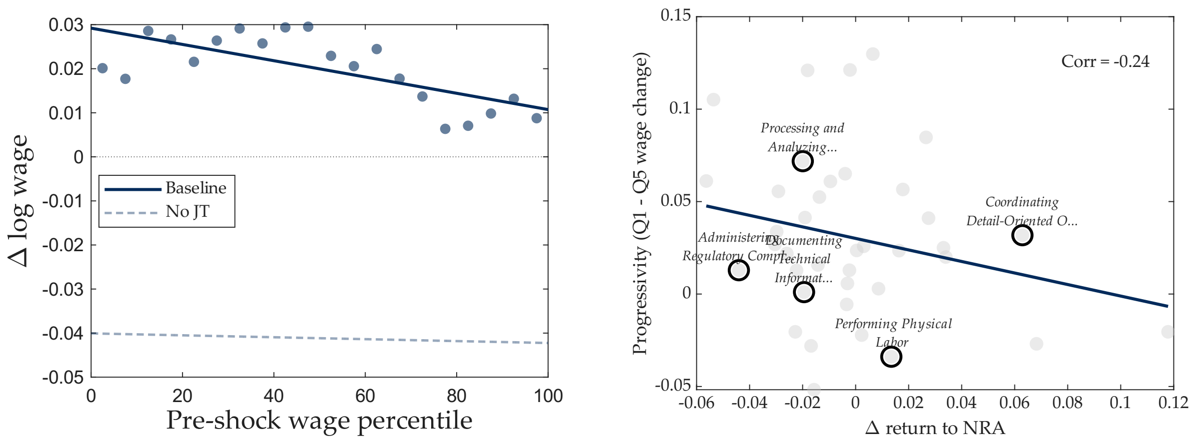
Figure 18a bins workers by their initial position in the wage distribution and shows the mean wage change by percentile. The model predicts that the wage effects of AI are mildly progressive: those who earned less pre-shock gain around 3%, on average, a figure that drops to around 1% at the top. Once again, job-transformation effects are crucial: As the dashed line indicates, the shock's progressivity largely disappears when A is held fixed.

The figure also reveals that AI-induced job transformation raises wages *across* the distribution: on average, workers are more productive in the remaining tasks than in those that were automated.

To understand why, we use the model to individually automate each task and study the effects of each shock. Figure 18b measures, on the vertical axis, the degree of progressivity, measured as the difference in wage gains in the bottom quintile compared to the top quintile. On the horizontal, we plot the change in the return to analytical skills. The figure documents that automation shocks lowering the return to analytical skills tend to be progressive. This is in part because such skills are more prevalent at the top of the wage distribution, as illustrated in Figure 5b.

In short, AI is a mildly progressive force in our model because it transforms jobs in a manner that reduces the returns to skills most prevalent at the top of the wage distribution, and raises the returns to skills more equally distributed across workers.

[28] See, for instance, Webb (2019) or David Autor in https://www.noemamag.com/how-ai-could-help-rebuild-the-middle-class.



(a) AI is mildly progressive

(b) Comparing shocks

Figure 18: LLM automation is mildly progressive due to job-transformation effects

Notes. The left panel plots the average log wage change by pre-shock wage percentile in both the model and the hypothetical scenario of holding A fixed. In the right panel, each dot represents a different automation scenario whereby a single task is fully automated.

Discussion. Two remarks contextualizing this result are in order. First, in a model of multi-dimensional skills, whether a shock is progressive or regressive is a distinct question from how it affects the overall dispersion of wages. Workers can achieve identical initial wages via different skill combinations, and automation differentially rewards these combinations. Appendix A.3 provides details.[29] We focus on the question of progressivity, as it is arguably of greater interest to society and policymakers experiencing the shock. Second, AI will inevitably affect the income distribution through many channels. For instance, our analysis assumes automation is universally adopted across firms, whereas some recent work suggests that there may be substantial differences in reality (Bonney *et al.*, 2024; Hubmer and Restrepo, 2025). We also do not engage with the important question of how capital income is distributed (Moll *et al.*, 2022). Our model instead highlights a distinct and novel, but natural, mechanism that emerges from the interplay between the multi-dimensionality of skills and the effects of automation on the demand for different tasks.

4.3 Discussion

We close this section by briefly revisiting important simplifications of our model and noting what limitations arise from them as well as highlighting avenues for future research.

First, the model features no switching costs or frictions. This attributes all non-random

[29]In very recent work, Benzell and Myers (2026) elegantly characterize the role of skill correlations in shaping the inequality effects of new technologies.

mobility to skill heterogeneity and understates occupational persistence, which may translate into excessive reallocation following shocks. As noted before, we could incorporate exogenous switching barriers into the model, but our baseline approach is more transparent. Relatedly, we assume that skills are time-invariant. This assumption facilitates estimation and analytical characterization. The interpretation of our results thus hinges on how long the distribution of skills in the population is expected to remain stable. Future work investigating the speed of skill adjustment in response to technology shocks would help illuminate this question.[30]

Second, we do not model heterogeneity along observable dimensions such as age or gender, isolating the role of skill heterogeneity. Exploring how job transformation effects vary with demographic characteristics is a natural direction for future work.

Third, following automation, the weight on non-automated tasks increases proportionately to their pre-shock occupation-specific weight. This, or a similar, assumption is indispensable for a forward-looking analysis. Future research tracking how workers actually reallocate time following AI-induced automation would be very valuable.

Fourth, the model does not incorporate a non-employment margin, as our focus is on wage outcomes. Automation-induced job loss is, of course, an important consideration, especially in the context of large shocks. The model could easily be extended to include non-employment as one option in workers' choice problem, with a payoff less sensitive to skills than employment. The literature suggests that a careful analysis of non-employment would require modeling the interplay of technological and cyclical forces (Jaimovich and Siu, 2020). A rigorous analysis of how AI-induced job transformation and non-employment interact is an interesting direction for future work.

5 Conclusion

If history offers any guidance, the rise of AI will lead to a substantial transformation of jobs. The central contribution of this paper is to propose a formal, task-based model of job transformation, develop the measurement tools to estimate it, and quantify the wage effects of AI.

One way to interpret our paper is that it formalizes widely shared, though loose, intuitions about the effects of automation.[31] From this perspective, a key insight of the paper is that the

[30] Adão *et al.* (2024) show that the speed of adjustment itself may endogenous to the nature of the technological shock, depending on whether innovations require skills that are more or less distinctive from those prevalent in the population. The adjustment to the ICT transition, for example, is shown to have been slow, relying on the entry of new cohorts.

[31] As one concrete example, John Burn-Murdoch in an article in the Financial Times (Jan 09, 2026) entitled "How to AI-proof your job" cites Deming's (2017) research on social skills and speculates: "as the writing of functions and formulas goes the way of hammering metal, quants and coders can reframe their professional identities around being the creative problem solvers, ideas people and project managers they always were".

mechanism underlying these intuitions is job transformation. This underlines the importance of modeling production at the level of jobs that bundle tasks, which the canonical task model abstracts from. We quantitatively demonstrate the important distributional consequences of AI-induced job transformation.

On the practical side, our findings carry two implications. First, the current discourse on the labor market consequences of AI often centers on whether entire jobs are eliminated (Frey and Osborne, 2017; Susskind, 2020), with evidence on both sides (e.g., Brynjolfsson *et al.*, 2025; Humlum and Vestergaard, 2025a). This is one important lens, especially in the long run. However, the absence of widespread job elimination should not be taken to mean that AI lacks major labor market effects. Automation can be disruptive for individual workers even when it does not eliminate their jobs, as the nature and requirements of their work are transformed, with significant consequences for wages. Second, our analysis clarifies how to interpret occupational exposure measures: average effects of exposure are non-monotonic, and workers exposed to the same shock in the same occupation may fare differently depending on their relative specialization. Exposure measures capture which tasks specific technologies affect and indicate the *potential* for change; deriving implications for worker outcomes requires a structural model that maps exposure into wages.

Our framework has several properties that, we hope, will make it useful for follow-up work. Because the model links directly to task-exposure measures, including forward-looking ones, it can project labor market effects without waiting for ex-post data—a feature especially useful for policy. While our quantitative application focused on LLMs, the framework readily accommodates different automation shocks, from self-driving vehicles to humanoid robots. Data requirements are modest, as worker-level panel data are widely available, so the framework can be applied across different countries and automation scenarios.

References

Acemoglu, D. and Autor, D. (2011). Skills, Tasks and Technologies: Implications for Employment and Earnings. In *Handbook of Labor Economics*, volume 4, pages 1043–1171. Elsevier.

Acemoglu, D. and Restrepo, P. (2018a). Artificial Intelligence, Automation and Work.

Acemoglu, D. and Restrepo, P. (2018b). 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 Restrepo, P. (2022). Tasks, Automation, and the Rise in U.S. Wage Inequality. *Econometrica*, **90**(5), 1973–2016.

Acemoglu, D., Kong, F., and Restrepo, P. (2025). Tasks At Work: Comparative Advantage, Technology and Labor Demand. *mimeo*.

Adão, R., Beraja, M., and Pandalai-Nayar, N. (2024). Fast and Slow Technological Transitions. *Journal of Political Economy Macroeconomics*, **2**(2), 183–227.

Aghion, P., Jones, B. F., and Jones, C. I. (2017). Artificial Intelligence and Economic Growth.

Althoff, L. and Reichardt, H. (2025). AI and Comparative Advantage. *mimeo*.

Atalay, E., Phongthiengtham, P., Sotelo, S., and Tannenbaum, D. (2020). The Evolution of Work in the United States. *American Economic Journal: Applied Economics*, **12**(2), 1–34.

Autor, D. and Thompson, N. (2025). Expertise.

Autor, D. H. and Dorn, D. (2013). The Growth of Low-Skill Service Jobs and the Polarization of the US Labor Market. *American Economic Review*, **103**(5), 1553–1597.

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

Baley, I., Figueiredo, A., and Ulbricht, R. (2022). Mismatch Cycles. *Journal of Political Economy*, **130**(11), 2943–2984.

Bartel, A., Ichniowski, C., and Shaw, K. (2007). How Does Information Technology Affect Productivity? Plant-Level Comparisons of Product Innovation, Process Improvement, and Worker Skills. *The Quarterly Journal of Economics*, **122**(4), 1721–1758.

Benzell, S. G. and Myers, K. R. (2026). Automation Experiments and Inequality.

Bessen, J. (2012). More Machines, Better Machines... or Better Workers? *The Journal of Economic History*, **72**(1), 44–74.

Bessen, J. E. (2011). Was Mechanization De-Skilling? The Origins of Task-Biased Technical Change. *SSRN Electronic Journal*.

Bick, A., Blandin, A., and Deming, D. J. (2024). The Rapid Adoption of Generative AI.

Bocquet, L. (2022). The Network Origin of Slow Labor Reallocation. *Working Papers*, (halshs-03703862).

Böhm, M., Etheridge, B., and Irastorza-Fadrique, A. (2025). The Impact of Labour Demand Shocks When Occupational Labour Supplies are Heterogeneous.

Bonney, K., Breaux, C., Buffington, C., Dinlersoz, E., Foster, L., Goldschlag, N., Haltiwanger, J., Kroff, Z., and Savage, K. (2024). The impact of AI on the workforce: Tasks versus jobs? *Economics Letters*, **244**, 111971.

Boustan, L. P., Choi, J., and Clingingsmith, D. (2022). Computerized Machine Tools and the Transformation of US Manufacturing.

Bratsberg, B., Rogeberg, O., and Terviö, M. (2025). Steeper at the top: Cognitive ability and earnings in Finland and Norway. *European Sociological Review*, **41**(3), 329–342.

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

Brynjolfsson, E., Chandar, B., and Chen, R. (2025). Canaries in the Coal Mine? Six Facts about the Recent Employment Effects of Artificial Intelligence. *Working Paper*.

Burstein, A., Morales, E., and Vogel, J. (2019). Changes in Between-Group Inequality: Computers, Occupations, and International Trade. *American Economic Journal: Macroeconomics*, **11**(2), 348–400.

Caselli, F. and Manning, A. (2019). Robot Arithmetic: New Technology and Wages. *American Economic Review: Insights*, **1**(1), 1–12.

Chequer, M., Herkenhoff, K., Papanikolaou, D., Schmidt, L., and Seegmiller, B. (2025). An Eddie Lazear Model of AI and Labor Markets: Theory Meets Resume Data. *mimeo*.

Choné, P., Kramarz, F., and Skans, O. N. (2025). Wage Markdowns and the Sorting of Workers to Firms when Skills are Bundled. *CEP Discussion Papers*, (DP20618).

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

del Rio-Chanona, R. M., Mealy, P., Beguerisse-Díaz, M., Lafond, F., and Farmer, J. D. (2021). Occupational mobility and automation: A data-driven network model. *Journal of The Royal Society Interface*, **18**(174), 20200898.

Deming, D. J. (2017). The Growing Importance of Social Skills in the Labor Market. *The Quarterly Journal of Economics*, **132**(4), 1593–1640.

Dix-Carneiro, R. (2014). Trade Liberalization and Labor Market Dynamics. *Econometrica*, **82**(3), 825–885.

Edin, P.-A., Fredriksson, P., Nybom, M., and Öckert, B. (2022). The Rising Return to Noncognitive Skill. *American Economic Journal: Applied Economics*, **14**(2), 78–100.

Edmond, C. and Mongey, S. (2021). Unbundling Labor. *mimeo*.

Eisfeldt, A. L., Schubert, G., and Zhang, M. B. (2023). Generative AI and Firm Values.

Eloundou, T., Manning, S., Mishkin, P., and Rock, D. (2023). GPTs are GPTs: An Early Look at the Labor Market Impact Potential of Large Language Models.

Fan, T. (2025). The Labor Market Incidence of New Technologies.

Fan, T. and Restrepo, P. (2025). Partial Automation. *mimeo*.

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

Felten, E. W., Raj, M., and Seamans, R. (2018). A Method to Link Advances in Artificial Intelligence to Occupational Abilities. *AEA Papers and Proceedings*, **108**, 54–57.

Freund, L. B. (2025). Superstar Teams. *JIWP Number: 2235*.

Frey, C. B. and Osborne, M. A. (2017). The future of employment: How susceptible are jobs to computerisation? *Technological Forecasting and Social Change*, **114**, 254–280.

Gathmann, C. and Schönberg, U. (2010). How General Is Human Capital? A Task-Based Approach. *Journal of Labor Economics*, **28**(1), 1–49.

Geel, R., Mure, J., and Backes-Gellner, U. (2011). Specificity of occupational training and occupational mobility: An empirical study based on Lazear's skill-weights approach. *Education Economics*, **19**(5), 519–535.

Girsberger, E. M., Koomen, M., and Krapf, M. (2022). Interpersonal, cognitive, and manual skills: How do they shape employment and wages? *Labour Economics*, **78**, 102235.

Goos, M., Manning, A., and Salomons, A. (2009). Job Polarization in Europe. *The American Economic Review*, **99**(2), 58–63.

Goos, M., Manning, A., and Salomons, A. (2014). Explaining Job Polarization: Routine-Biased Technological Change and Offshoring. *American Economic Review*, **104**(8), 2509–2526.

Grigsby, J. (2023). Skill Heterogeneity and Aggregate Labor Market Dynamics. *mimeo*.

Grigsby, J. and Zorzi, N. (2025). The Labor Market Consequences of a Rapid Climate Transition. *mimeo*.

Guvenen, F., Kuruscu, B., Tanaka, S., and Wiczer, D. (2020). Multidimensional Skill Mismatch. *American Economic Journal: Macroeconomics*, **12**(1), 210–244.

Hampole, M., Papanikolaou, D., Schmidt, L. D., and Seegmiller, B. (2025). Artificial Intelligence and the Labor Market.

Heckman, J. and Scheinkman, J. (1987). The Importance of Bundling in a Gorman-Lancaster Model of Earnings. *The Review of Economic Studies*, **54**(2), 243–255.

Hernnäs, S. (2023). Automation when skills are bundled. Technical Report 2023:2, IFAU - Institute for Evaluation of Labour Market and Education Policy.

Hosseini Maasoum, S. M. and Lichtinger, G. (2026). Generative AI and Occupational Entry Barriers: The Labor-Supply Channel of Technological Change.

Hsieh, C.-T., Hurst, E., Jones, C. I., and Klenow, P. J. (2019). The Allocation of Talent and U.S. Economic Growth. *Econometrica*, **87**(5), 1439–1474.

Hubmer, J. and Restrepo, P. (2025). Not a Typical Firm: Capital–Labor Substitution and Firms' Labor Shares. *American Economic Journal: Macroeconomics*.

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

Humlum, A. and Vestergaard, E. (2025a). Large Language Models, Small Labor Market Effects.

Humlum, A. and Vestergaard, E. (2025b). The unequal adoption of ChatGPT exacerbates existing inequalities among workers. *Proceedings of the National Academy of Sciences*, **122**(1), e2414972121.

Ide, E. and Talamas, E. (2025). Artificial Intelligence in the Knowledge Economy. *Journal of Political Economy*.

Jaimovich, N. and Siu, H. E. (2020). Job Polarization and Jobless Recoveries. *The Review of Economics and Statistics*, **102**(1), 129–147.

Jones, C. I. (2022). The Past and Future of Economic Growth: A Semi-Endogenous Perspective. *Annual Review of Economics*, **14**(Volume 14, 2022), 125–152.

Jones, C. I. and Tonetti, C. (2026). Past Automation and Future A.I.: How Weak Links Tame the Growth Explosion. *mimeo*.

Kambourov, G. and Manovskii, I. (2008). Rising Occupational and Industry Mobility in the United States: 1968-97. *International Economic Review*, **49**(1), 41–79.

Kogan, L., Papanikolaou, D., Schmidt, L. D., and Seegmiller, B. (2023). Technology and Labor Displacement: Evidence from Linking Patents with Worker-Level Data.

Lashkari, D., Qiu, C., Li, W., and Thompson, N. (2025). AI, Scale, and Skills: A Quantitative Task-Based Theory of Automation. *mimeo*.

Lazear, E. P. (2009). Firm-Specific Human Capital: A Skill-Weights Approach. *Journal of Political Economy*, **117**(5), 914–940.

Lindenlaub, I. (2017). Sorting Multidimensional Types: Theory and Application. *The Review of Economic Studies*, **84**(2), 718–789.

Lise, J. and Postel-Vinay, F. (2020). Multidimensional Skills, Sorting, and Human Capital Accumulation. *American Economic Review*, **110**(8), 2328–2376.

McInnes, L., Healy, J., and Astels, S. (2017). Hdbscan: Hierarchical density based clustering. *Journal of Open Source Software*, **2**(11), 205.

Moll, B., Rachel, L., and Restrepo, P. (2022). Uneven Growth: Automation's Impact on Income and Wealth Inequality. *Econometrica*, **90**(6), 2645–2683.

Ocampo Díaz, S. (2022). A task-based theory of occupations with multidimensional heterogeneity. Technical report, mimeo.

Restrepo, P. (2024). Automation: Theory, Evidence, and Outlook. *Annual Review of Economics*, **16**(Volume 16, 2024), 1–25.

Roy, A. D. (1951). Some Thoughts on the Distribution of Earnings. *Oxford Economic Papers*, **3**(2), 135–146.

Sanders, C. (2012). Skill Uncertainty, Skill Accumulation, and Occupational Choice. *2012 Meeting Papers*, (633).

Spitz-Oener, A. (2006). Technical Change, Job Tasks, and Rising Educational Demands: Looking outside the Wage Structure. *Journal of Labor Economics*, **24**(2), 235–270.

Susskind, D. (2020). *A World Without Work: Technology, Automation, and How We Should Respond*. Metropolitan Books, New York, N.Y.

Traiberman, S. (2019). Occupations and Import Competition: Evidence from Denmark. *American Economic Review*, **109**(12), 4260–4301.

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

Woessmann, L. (2024). Skills and Earnings: A Multidimensional Perspective on Human Capital.

Online Appendix

This appendix contains supplemental material. Any references to sections, equations, figures, or tables that are not preceded by a capital letter refer to the main article.

A Appendix: Theory

A.1 Partial automation

To formalize the idea of partial automation, note that we can re-write the pre-automation production technology of a firm employing worker i in occupation o at time t as

$$\begin{aligned} Y_{i,o,t} &= \Gamma \prod_{\tau \in \mathcal{T}_m} \left(X^{\text{machines}}_{i,o,\tau,t}\right)^{\alpha_{o,\tau}} \prod_{\tau \in \mathcal{T}_\ell \setminus \{\tau^\star\}} \left(X^{\text{labor}}_{i,o,\tau,t}\right)^{\alpha_{o,\tau}} \\ &\quad \times \left((1-\zeta_{\tau^\star}) X^{\text{labor}}_{i,o,\tau^\star,t}\right)^{(1-\zeta_{\tau^\star})\alpha_{o,\tau^\star}} \left(\zeta_{\tau^\star} X^{\text{labor}}_{i,o,\tau^\star,t}\right)^{\zeta_{\tau^\star}\alpha_{o,\tau^\star}}, \end{aligned}$$

where

$$\Gamma = (1-\zeta_{\tau^\star})^{-\alpha_{o,\tau^\star}(1-\zeta_{\tau^\star})} \zeta_{\tau^\star}^{-\alpha_{o,\tau^\star}\zeta_{\tau^\star}}$$

and superscripts indicate whether a task input X is produced using labor or machines. This formulation essentially "splits" the task $\tau^\star$ into two parts: an automatable and a non-automatable component. Automation means that the automatable task is reassigned from labor to machines. The post-automation production technology thus becomes

$$Y_{i,o,t} = \Gamma \left[\prod_{\tau \in \mathcal{T}_\ell \setminus \{\tau^\star\}} \left(X^{\text{labor}}_{i,o,\tau,t}\right)^{\alpha_{o,\tau}} \cdot \left(X^{\text{labor}}_{i,o,\tau^\star,t}\right)^{(1-\zeta_{\tau^\star})\alpha_{o,\tau^\star}} \right] \cdot \left[\left(X^{\text{machines}}_{i,o,\tau^\star,t}\right)^{\zeta_{\tau^\star}\alpha_{o,\tau^\star}} \cdot \prod_{\tau \in \mathcal{T}_m} \left(X^{\text{machines}}_{i,o,\tau,t}\right)^{\alpha_{o,\tau}} \right].$$

A.2 Endogenizing $(\mathcal{T}_l, \mathcal{T}_m)$ & the automation threshold $\bar{z}_{\tau^\star}$

Section 2.1 treats the assignment of production tasks to labor and machines, $(\mathcal{T}_l, \mathcal{T}_m)$, as exogenous. We now discuss a set of additional assumptions that allow us to endogenize these sets as firms' choices. This allows us to determine, for any task $\tau^\star$ an "automation threshold" $\bar{z}_{\tau^\star}$ that triggers the optimal automation of this task. For our quantitative exercises, we assume that the productivity of any automated task $\tau^\star$ equals $\bar{z}_{\tau^\star}$.

Entrepreneurs. There is a large mass of entrepreneurs. In every period, every worker randomly matches with $N \geq 2$ entrepreneurs. Before the occupation and skill are revealed to the entrepreneur, the entrepreneur makes an automation decision. That is, they decide the set of tasks that are produced with human labor, $\mathcal{T}_l$, and the set of tasks done by machines, $\mathcal{T}_m$. After automation decisions are taken, the occupation o and the worker's characteristics $(s_{i,\cdot}, \varepsilon_{i,t})$ are revealed. Wages are then set via Bertrand competition. Lastly, the winning entrepreneur forms a match with the worker and optimally allocates the worker's time to human tasks and machine capital to machine tasks.

Automation choice. Given some vector $\{z_\tau\}_{\tau\in\mathcal{T}}$, we define an optimal automation choice as task sets $(\mathcal{T}_l, \mathcal{T}_m)$ such that no entrepreneur finds it optimal to deviate from this task assignment. Note that the wage paid to a given worker is independent of the automation choice from the perspective of an individual firm considering a deviation. Thus, for any task τ, the condition that no firm finds it optimal to deviate from the assignment $(\mathcal{T}_l, \mathcal{T}_m)$ can be written as

$$\begin{aligned} &\int \left(\max_{m'} P_o Y_o'(M') - \exp(w(s,o,\varepsilon)) - rM' \right) dF(s|o)dG(\varepsilon)d\Lambda(o) \\ \leq &\int \left(\max_{m} P_o Y(m) - \exp(w(s,o,\varepsilon)) - rM \right) dF(s|o)dG(\varepsilon)d\Lambda(o) \end{aligned} \tag{A.1}$$

where Y' denotes the production function under a given alternative choice of task sets $(\mathcal{T}_l', \mathcal{T}_m')$, and Λ, G, and F denote the distributions of occupational choices, idiosyncratic shocks ε (which are independent of occupational choices), and skills s conditional on occupational choices, respectively. The task assignment $(\mathcal{T}_l, \mathcal{T}_m)$ is thus optimal if and only if, for any alternative task assignments $(\mathcal{T}_l', \mathcal{T}_m')$, equation (A.1) is satisfied.

Using the expectations operator in place of integrals and substituting in optimality conditions, we can also write equation (A.1) as

$$\begin{aligned} &\mathbb{E}_{(s|o),\varepsilon,o}\left[\exp\left(\mu_o' + \sum_{\tau\in\mathcal{T}_l'} \frac{\alpha_{o,\tau}}{\sum_{\tau\in\mathcal{T}_l'}\alpha_{o,\tau}} s_{i,\tau} + \varepsilon_{i,t}\right)\right] \\ \leq\ &\mathbb{E}_{(s|o),\varepsilon,o}\left[\exp\left(\mu_o + \sum_{\tau\in\mathcal{T}_l} \frac{\alpha_{o,\tau}}{\sum_{\tau\in\mathcal{T}_l}\alpha_{o,\tau}} s_{i,\tau} + \varepsilon_{i,t}\right)\right], \end{aligned} \tag{A.2}$$

where μ_o' holds occupational prices fixed. This establishes that the optimal automation threshold is the one that would leave average wages constant if occupational choices were held constant.

It can be shown that it is always possible to find values of $\{z_\tau\}_{\tau\in\mathcal{T}}$ that justify a given initial task assignment $(\mathcal{T}_l, \mathcal{T}_m)$ as optimal. For each task τ, we define the automation threshold $\bar{z}_\tau$ as the point at which Equation (A.2) holds with equality. That is, holding occupational choices constant, the average wage in the economy stays constant whether or not task τ is automated. It can be verified that, given an initially optimal assignment, there is such a threshold value $\bar{z}_\tau$ for any task.

Equilibrium with endogenous automation. An equilibrium with endogenous automation is defined as a tuple of automation choices $(\mathcal{T}_l, \mathcal{T}_m)$, a vector of occupational and final good output $(Y_{\cdot,t}, Y_t)$, a distribution $\Gamma(i)$, occupation choices $\hat{o}_{i,t}$, log wages $\{w_{i,o,t}\}$, log skills s_i, idiosyncratic productivity shocks $\varepsilon_{i,t}$ that are functions of i, and a set of prices $\{P_{o,t}\}_{o\in\mathcal{O}}$, such that: (i) equation (A.1) holds for any alternative choice of task sets $(\mathcal{T}_l', \mathcal{T}_m')$; (ii) equation (4) holds at any point in the distribution (firms make zero profits); (iii) the marginal distribution of occupations conditional on wages follows equation (7) (workers optimize); (iv) the final goods aggregator optimizes, yielding occupation-level demands $Y_{o,t} = \omega_o P_{o,t}^{-\sigma} Y_t$ for all $o \in \mathcal{O}$; (v) occupation-level output markets clear: $Y_{o,t} = \int \mathbb{I}\{\hat{o}_{i,t} = o\} Y_{i,o,t} d\Gamma(i)$ for all $o \in \mathcal{O}$; (vi) the final good aggregator makes zero profits: $Y_t = \sum_{o\in\mathcal{O}} P_{o,t} Y_{o,t}$; and (vii) the unconditional marginal distributions of skills s_i and occupational shocks $\varepsilon_{i,t}$ follow $\mathcal{N}(\bar{s}, \Sigma_s)$ and $\mathcal{N}(\delta_t, \varsigma^2 I)$, respectively.

Automation of multiple tasks. When multiple tasks are being automated, as in Section 4.1, we construct the $\bar{z}_{\tau^\star}$ thresholds by supposing tasks are automated sequentially in order of exposure; that is, we first calculate the automation threshold for partial automation of the most exposed task, then, starting from an equilibrium in which the most exposed task has been partially automated, we proceed to calculate the automation threshold for the second most exposed task, and so on.

A.3 Automation, progressivity, and inequality when skills are multi-dimensional

This appendix shows that in a model of multi-dimensional skills, a shock leading those with lower initial wages to gain more (or lose less) need not reduce the dispersion of log wages.

Let w_i^{pre} denote worker i's average log wage pre-automation and w_i^{post} post-automation. Define the wage change $\Delta w_i = w_i^{post} - w_i^{pre}$. We say a shock is *progressive* if $\mathrm{Cov}(\Delta w_i, w_i^{pre}) < 0$. Inequality increases if $\mathrm{Var}(w_i^{post}) > \mathrm{Var}(w_i^{pre})$.

We can decompose the change in the variance of log wages by noting that, by definition,

$$\mathrm{Var}(w_i^{post}) = \mathrm{Var}(w_i^{pre} + \Delta w_i) = \mathrm{Var}(w_i^{pre}) + \mathrm{Var}(\Delta w_i) + 2\mathrm{Cov}(w_i^{pre}, \Delta w_i).$$

Rearranging:

$$\underbrace{\mathrm{Var}(w_i^{post}) - \mathrm{Var}(w_i^{pre})}_{\text{Change in inequality}} = \underbrace{\mathrm{Var}(\Delta w_i)}_{\text{Dispersion of changes}} + \underbrace{2\mathrm{Cov}(w_i^{pre}, \Delta w_i)}_{\text{Progressivity term}}.$$

Hence, inequality increases even if the shock is progressive if (and only if) $\mathrm{Var}(\Delta w_i) > 2|\mathrm{Cov}(w_i^{pre}, \Delta w_i)|$.

To gain intuition, write $\Delta w_i = \beta \cdot w_i^{pre} + \varepsilon_i$, where $\beta = \text{Cov}(\Delta w_i, w_i^{pre})/\text{Var}(w_i^{pre})$ is the progressivity slope and $\varepsilon_i \perp w_i^{pre}$. Then

$$\text{Var}(w_i^{post}) - \text{Var}(w_i^{pre}) = \text{Var}(w_i^{pre}) \cdot \beta(\beta + 2) + \text{Var}(\varepsilon_i).$$

The residual ε_i captures variation in wage changes *orthogonal* to initial position—dispersion in Δw_i among workers with the same w_i^{pre}. In our model, initial wage is a scalar summary, but as skills are multi-dimensional, two workers with identical w_i^{pre} can have very different skill bundles. Following automation, the effect on worker i depends on s_i, not just w_i^{pre}, creating $\text{Var}(\varepsilon_i) > 0$. By contrast, in a model with scalar skill $s_i \in \mathbb{R}$ and wages $w_i = g(s_i)$ for monotonic g, knowing w_i^{pre} perfectly identifies s_i. The wage change $\Delta w_i = g'(s_i) - g(s_i) = h(w_i^{pre})$ is a deterministic function of initial wage, implying $\varepsilon_i = 0$ for all i. In such a model with one-dimensional skills, shocks that are progressive necessarily reduce inequality.

B Appendix: Theory Meets Data

B.1 Estimation methodology

For simplicity we assume without loss of generality that in the initial steady state there is only one composite machine task with productivity normalized to $\log r$. This implies that the intercept term

$$
\begin{aligned}
\mu_{o,t} =& \frac{\log(P_{o,t})}{LS_o} + \sum_{\tau \in \mathcal{T}} \frac{\alpha_{o,\tau}}{\sum_{\tau \in \mathcal{T}_l} \alpha_{o,\tau}} \log\left(\alpha_{o,\tau}\right) + \left(\sum_{\tau \in \mathcal{T}_m} \frac{\alpha_{o,\tau}}{\sum_{\tau \in \mathcal{T}_l} \alpha_{o,\tau}} \left(z_\tau - \log r\right) \right) \\
=& \frac{\log(P_{o,t})}{LS_o} + \sum_{\tau \in \mathcal{T}_l} A_{o,\tau} \log\left(A_{o,\tau} \cdot LS_o\right) + \frac{1 - LS_o}{LS_o} \left(\log LS_o\right)
\end{aligned}
$$

depends only on occupation-specific prices, occupational labor shares $LS_o = \sum_{\tau \in \mathcal{T}_l} \alpha_{o,\tau}$ and elements of A. Our estimation approach treats A and LS_o as observable for all occupations, while the two vectors of prices in each A-regime, $\vec{P}$ and $\vec{P}'$, have to be estimated.

In what follows, let $\hat{o}_{i,t}$ denote the recorded occupation choice of worker i in period t and $-\hat{o}_{i,t}$ be the set of occupations not chosen in period t. We count pre-2000 and post-2000 occupations as distinct sets of occupations, with only one set of occupations available to the worker at any given time. Next, we define

$$
A^\star = \begin{pmatrix} A \\ A' \end{pmatrix}
$$

where A and A' correspond to the matrix defined in (6) before and after the year 2000. We denote the vector of prices in each regime, which can be shown to be independent of t as $\vec{P} = (P_1, \ldots, P_O)$ and $\vec{P}' = (P'_1, \ldots, P'_O)$, respectively.

For a given worker observed in occupations $(\hat{o}_{o,1}, \ldots, \hat{o}_{i,T})$ and with wage history $(w_{i,\hat{o}_1,1}, \ldots, w_{i,\hat{o}_T,T})$,

$$
\begin{pmatrix} s_1 \\ \vdots \\ s_{n_{\text{skill}}} \\ w_{i,\hat{o}_1,1} \\ \vdots \\ w_{i,\hat{o}_T,T} \end{pmatrix} = \begin{pmatrix} 0 \\ \vdots \\ 0 \\ \mu_{\hat{o}_1} \\ \vdots \\ \mu_{\hat{o}_T} \end{pmatrix} + \begin{pmatrix} I & 0 \\ A^\star_{(\hat{o}_1,\ldots,\hat{o}_T),\cdot} & I \end{pmatrix} \cdot \begin{pmatrix} s_1 \\ \vdots \\ s_{n_{\text{skill}}} \\ \varepsilon_{i,1} \\ \vdots \\ \varepsilon_{i,T} \end{pmatrix}, \quad \text{where} \begin{pmatrix} s_1 \\ \vdots \\ s_{n_{\text{skill}}} \\ \varepsilon_{i,1} \\ \vdots \\ \varepsilon_{i,T} \end{pmatrix} \sim \mathcal{N}\left(\begin{pmatrix} \bar{s}_1 \\ \vdots \\ \bar{s}_{n_{\text{skill}}} \\ \delta_1 \\ \vdots \\ \delta_T \end{pmatrix}, \begin{pmatrix} \Sigma_s & 0 \\ 0 & \varsigma^2 I \end{pmatrix} \right).
$$

Thus, $w_{i,\hat{o}_{i,\cdot},\cdot}$ and s_i are jointly normal, which yields easy to compute formulas for the distribution of $s_i|w_{i,\hat{o}_{i,\cdot},\cdot}$. The likelihood of observing $(w_{i,\hat{o}_{i,\cdot},\cdot},\hat{o}_{i,\cdot})$ is then given by

$$\mathcal{L}(w_{i,\hat{o}_{i,\cdot},\cdot},\hat{o}_{i,\cdot}|\nu,\varsigma,\bar{s},\Sigma_s,\vec{P},\vec{P}') =$$

$$\prod_i \int_s \Bigg[\left(\int_{w_{i,\cdot,-\omega.}} \prod_t P(\hat{o}_{i,t} = \omega_{i,t}|w_{i,\cdot,\cdot},\nu) \cdot f(w_{i,t,-\omega_t}|s,w_{i,\cdot,\omega.},\varsigma,\vec{P},\vec{P}') \right)$$

$$\cdot f(s|w_{i,\cdot,\omega.},\varsigma,\bar{s},\Sigma_s,\vec{P},\vec{P}') \Bigg] \cdot f(w_{i,\cdot,\omega.}|\varsigma,\bar{s},\Sigma_s,\vec{P},\vec{P}').$$

Maximizing this expression involves two key challenges. The first challenge is that we seek to maximize over a high-dimensional integral, which is intractable. To overcome this challenge, we use Monte Carlo integration to compute a numerical approximation of the likelihood instead of evaluating this expression analytically. That is, instead of maximizing the analytical likelihood, we instead maximize the mean of a simulated statistical object that converges to the likelihood value for large sample sizes, n_0. It can be shown that, as $n_0 \to \infty$, the argmax of this object converges to the true maximum likelihood estimate under mild regularity conditions. We find that, in practice, $n_0 = 40$ yields a sufficiently accurate approximation to deliver satisfactory results in a Monte Carlo exercise, which we report in Section 3.3.

Concretely, our implementation of this idea is as follows: For all individual workers i, we generate n_0 draws from

$$f(w_{i,\cdot,-\omega.}|w_{i,\cdot,\omega.},\varsigma,\bar{s},\Sigma_s,\vec{P},\vec{P}') = \int_s f(w_{i,\cdot,-\omega.}|s,w_{i,\cdot,\omega.},\varsigma,\vec{P},\vec{P}') f(s|w_{i,\cdot,\omega.},\varsigma,\bar{s},\Sigma_s,\vec{P},\vec{P}')$$

These draws can be generated by (i) drawing from the distribution $s_i|w_{i,\hat{o}_{i,\cdot},\cdot}$, (ii) computing $\varepsilon_{i,t}$ for every period (as a deterministic function of s_i and $w_{i,\hat{o}_{i,\cdot},\cdot}$), and (iii) computing the resulting vector of all occupational wages in every period. Using these wages, we then evaluate the mean of $P(\hat{o}_{i,t}|w_{i,\hat{o}_{i,\cdot},\cdot},\nu)$ to obtain an estimator for $\mathcal{L}_i(\theta)$:

$$\hat{\mathcal{L}}_i(w_{i,\cdot,\hat{o}_{i,\cdot}},\nu,\varsigma,\bar{s},\Sigma_s,\vec{P},\vec{P}') = \left(\frac{1}{n_0} \sum_j \prod_t P(\hat{o}_{i,t} = \omega_{i,t}|w_{j,t,\cdot},\nu) \right) \cdot f(w_{i,\cdot,\omega.}|\varsigma,\bar{s},\Sigma_s,\vec{P},\vec{P}') \tag{B.1}$$

The second challenge is that our maximization procedure must impose that prices satisfy the following restriction across regimes, which follows the GE structure of the model and the assumption that only A is allowed to change across the pre 2000 and post 2000 regimes:

$$H(\vec{P},\vec{P}') \equiv \Delta \log P_o + \frac{1}{\sigma}\Big(\Delta \log Y_o(\vec{P},\vec{P}') - \Delta \log Y(\vec{P},\vec{P}')\Big) = 0. \tag{B.2}$$

To account for the fact that equation (B.2) must hold for any set of parameters, we use the implicit function theorem to compute an augmented gradient for each entry in $\theta = (\nu, \varsigma, \bar{s}, \Sigma_s, \vec{P})$[B.1]:

$$\bar{D}_x \hat{\mathcal{L}} = D_x \hat{\mathcal{L}} + D_{\vec{P}'} \hat{\mathcal{L}} \cdot (-D_{\vec{P}'} H)^{-1} \cdot D_x H \quad \forall x \in \{\nu, \varsigma, \bar{s}, \Sigma_s, \vec{P}\}.$$

Holding constant all random variables in the estimator, we then proceed to maximize the object in (B.3) over the parameter space θ. This parameter space is large and requires efficient numerical optimization methods. We utilize stochastic gradient descent paired with auto-differentiation techniques that allow us to efficiently compute gradients of $\hat{\mathcal{L}}_i$.

Stochastic gradient descent To estimate the model parameters, including the joint distribution of skills, the following stochastic object has to be maximized:

$$\hat{\mathcal{L}}_i(w_{i,\cdot,\hat{o}_{i,\cdot}}, \nu, \varsigma, \bar{s}, \Sigma_s, \vec{P}, \vec{P}') = \left(\frac{1}{n_0} \sum_j \prod_t P(\hat{o}_{i,t} = \omega_{i,t} | w_{j,t,\cdot}, \nu) \right) \cdot f(w_{i,\cdot,\omega.} | \varsigma, \bar{s}, \Sigma_s, \vec{P}, \vec{P}') \tag{B.3}$$

To do this, we exploit the fact $s_i | w_{j,\hat{o}_{i,\cdot},\cdot}$ is normal and thus any can be written as

$$s_i = \mu_s^{cond} + L_s^{cond} \cdot u$$

for some easy to compute $(\mu_s^{cond}, L_s^{cond})$ and $u \sim \mathcal{N}(0, I)$. u is drawn once and then held constant throughout the maximization procedure, while μ_s^{cond} and L_s^{cond} depend on model parameters. We therefore proceed as follows:

(i) For each worker i, generate n_0 draws of u that remain fixed

(ii) Compute $s_i = \mu_s^{cond} + L_s^{cond} \cdot u$

(iii) Compute $\varepsilon_{i,t} = w_{i,\hat{o}_{i,t},t} - \mu_o - A_{o,\cdot} \cdot s_i$

(iv) Use these draws to obtain a sample $w_{j,\cdot,\cdot}$ of wages in every occupation-period cell.

(v) Compute $\hat{\mathcal{L}}_i(w_{i,\hat{o}_{i,\cdot},\cdot}, \hat{o}_{i,\cdot} | \nu, \varsigma, \bar{s}, \Sigma_s)$

We then employ stochastic gradient descent. That is, starting with some guess, we update our parameters as follows:

$$\theta_{t+1} = \theta_t - \eta \cdot \nabla \left(-\hat{\mathcal{L}}(\theta_t) \right)$$

[B.1] $\vec{P}'$ is computed numerically as a function of these variables by imposing that the system in equation (B.2) must hold.

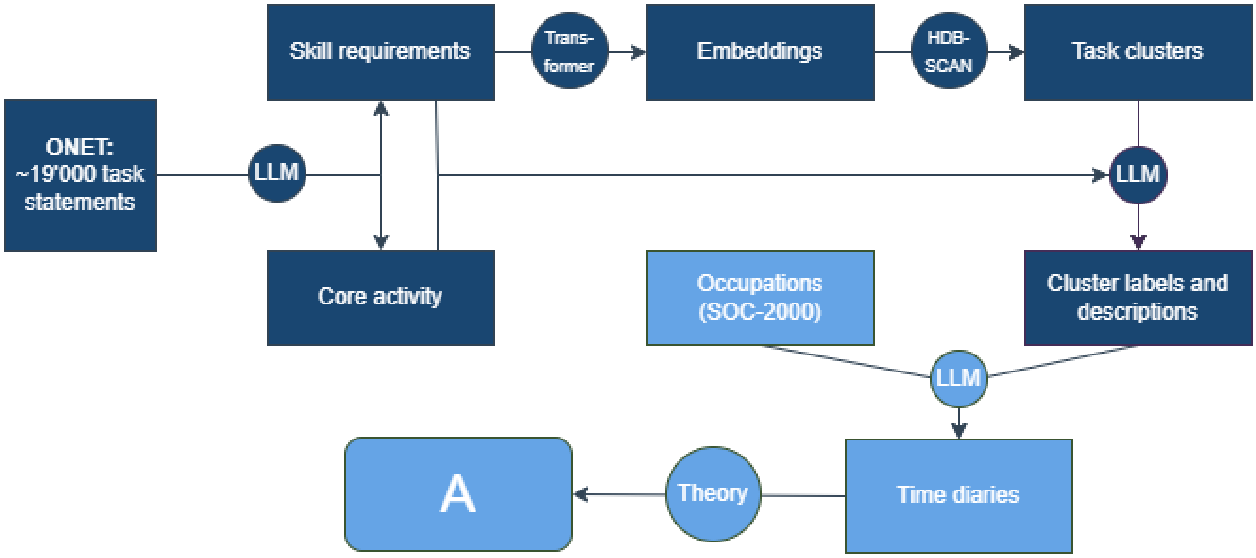


Figure B.1: Schematic overview of the measurement of tasks and the A matrix

Notes. Step 1 (colored in dark blue) involves the clustering of tasks, step 2 (colored in light blue) the measurement of occupational task weights.

for some sufficiently small $\eta > 0$. To further ease the computational load, we evaluate the likelihood at a subsample B only. To do this, we iterate over epochs. In each epoch, we randomly partition individuals into n groups:

$$\{1, 2, \ldots, I\} = B_1 \cup B_2 \cup \ldots \cup B_n, \quad B_i \cap B_j = \emptyset$$

and, on each iteration within an epoch, evaluate the likelihood based on batch $B_1, \ldots, B_n$ only. Using parallelization over individuals and auto-differentiation techniques, this reduces the computation time of the likelihood maximization procedure substantially and allows us to solve this problem even when the parameter space is very large.

B.2 Clustering of occupation-specific tasks

This section describes in more detail how we construct the task clusters on which our analysis is based. Our starting point is the list of detailed task statements in the O*NET database. Figure B.1 summarizes the NLP pipeline.

B.2.1 Extracting skill requirements

For each of the O*NET micro tasks, we extract a core activity and skill requirements using a LLM, specifically openAI's GPT-4o model. The system and user prompts are stated below.[B.2]

Table B.1 provides an illustration of the core activity and required skills this approach extracts for a set of skills; it also indicates the cluster the cluster will eventually be assigned to.

System Prompt.

```
<Role> You are an expert in labor economics, job analysis, and task
   classification. </Role>

<Overall goal>
You will be presented with a list of {len(tasks_chunk)}
   occupation-specific task statements. The ultimate goal is to group
   these and thousands of other tasks into clusters based on the type
   of activity and skills utilized, i.e. someone skilled at one task
   in a cluster could perform others in that cluster.

Your overall task is to prepare this clustering step by identifying,
   for each task statement:
(i) the fundamental  work activity; and (ii) the most essential skills
   and abilities (up to 5) required to perform this task effectively.

Key requirements for (i) the fundamental work activity:
- Definition: The fundamental work activity is a concise, abstract
   description that encapsulates the core activity  involved in the
   task statement (what is being done).
- Generalization: The activity label should be broad enough that if
   someone can perform one task under this label,
```

[B.2] The *temperature* parameter is set to 0.00001, which directs the model to provide its most confident response, minimizing variation across runs. Technically, the LLM predicts the next word in a sequence based on the preceding words and its prior training. Denoting by q_i the logit for candidate token i, the softmax function is used to scale the logits and map them into probabilities: $\frac{e^{q_i/T}}{\sum_{k=1}^{\text{size of vocabulary}} e^{q_k/T}}$. The parameter T is known as *temperature.* A higher temperature value "excites" previously low probability candidates, encouraging creativity, whereas a lower temperature value lowers the smaller outputs relative to the largest outputs. A lower value is thus preferable for contexts requiring high coherence and accuracy. Note, though, that even $T = 0$ does not result in deterministic output in practice, likely due to sources of randomness such as the state of the random-number generator. Moreover, the so-called "reasoning" generation of models does not support a temperature parameter. In practice, we have verified that the time allocation shares are highly comparable across different runs of the model.

```
 they'd be expected to handle any task requiring that same underlying
   competency.
- Terminology: Use concise and standardized, domain-agnostic terms
   that capture the core function, phrasing them in clear,
   natural-sounding language.
- Self-explanatory: The label must offer a succinct, self-contained
   summary that includes essential context for standalone
   understanding;
     do not merely reduce the statement to a vague abbreviation.
- Predominant activity: When multiple actions are present, select the
   one that best represents the overall purpose of the task.

Key requirements for (ii) the skills and abilities:
- Definition: skills refer to developed capacities that facilitate
   performance of activities that occur across jobs; abilities refer
   to relatively enduring attributes of an individual's capability for
   performing a particular range of different tasks. A "skill" is not
   simply a rewording of the task/activity description itself, but
   rather answer the question "What underlying capability makes
   someone good at this task?"
So for each skill you identify, ask: 'Would this skill enable
   performance across MULTIPLE different tasks and contexts?' If not,
   it's likely not a true underlying skill.
- Task: Identify the essential skills and abilities required to
   perform this task effectively and list them in descending order of
   importance.
- The number of skills can range from 1 to 5, depending on the
   complexity of the task; for straightforward ones, only include the
   core skills (at least 1); avoid padding with peripheral skills.
- The "most important skills" can include both capabilities and, where
   critical, knowledge domains, including:
i) Cognitive capabilities
Examples:  strategic planning, statistical analysis, diagnostic
   reasoning, technical writing
ii) Specialized technical capabilities
Examples: programming, surgical technique, database architecture
```

```
iii) Interpersonal capabilities, management, and leadership
Examples: negotiation, leadership, instruction, conflict resolution,
   team development, performance evaluation, delegation,
   organizational design, change management,
management of financial resources, management of personnel resources
iv) Physical/sensory capabilities
Examples: fine motor control, spatial awareness, physical endurance
v) Specialized expertise areas
Examples: mathematical modeling, designing scientific experiments,
   legal precedents, medical protocols
- Each skill in the list must follow this format: "Skill Name (Level)"
- Level must be one of "basic", "intermediate", "advanced", or
   "expert" using the following criteria:
basic: requires fundamental knowledge and minimal experience;
   intermediate: requires specialized knowledge and moderate
   experience; advanced: requires deep expertise and substantial
   experience;
expert: requires mastery-level knowledge, typically 8+ years of
   focused experience.
- Critical: When identifying skills, pay particular attention to
   specialized capabilities that typically command higher wages in the
   labor market, such as: Complex analytical or strategic thinking
   skills, Specialized technical expertise that requires extensive
   training, High-stakes decision-making capabilities, Skills
   involving the direction of others' work or significant resources,
   Expertise that is both scarce and in high demand. For high-wage
   occupations, ensure you separately list these skills rather than
   using generic descriptors.
</Overall goal>

<Detailed instructions>
Step 1) For each task statement, identify and summarize (i) the
   fundamental work activity; and (ii) the most important  skills.
Step 2) Return the output (activity; skills) for all
   {len(tasks_chunk)} task statements in the JSON format specified.
- List the skills in descending order of importance to the task (most
```

```
   crucial first).
- Never leave any task blank; if unsure, provide your best guess.
</Detailed instructions>

<Examples>
The following examples illustrate the level of abstraction desired
   (for reference only, do not copy these exact labels unless they
   truly match the task at hand).

Example work activities: "train and teach others at work", "operate
   vehicles", "operate industrial machinery," "provide advice or
   consultation," "coordinate the work of subordinates/peers,"
   "inspect or repair equipment," etc.

Example task: "Review statistical studies, technological advances, or
   regulatory standards and trends to stay abreast of issues in the
   field of quality control."
Activity: evaluate complex technical information
Skills: analytical thinking (expert), research (advanced), statistical
   analysis (advanced), reading comprehension (advanced)

Example task: "Wash glasses or other serving equipment at bars."
- Activity: cleaning
- Skills: manual dexterity (basic)

Example task: "Analyze financial statements to determine company
   valuation"
Activity: analyze and interpret financial data
Skills: market analysis (expert), numerical reasoning (advanced), data
   analysis (intermediate), financial modeling (advanced)

Example task: "Lead strategic planning for a multinational division
   with $500M annual revenue"
Activity: direct organizational strategy
Skills: leadership (expert), strategic planning (expert), financial
   analysis (advanced), business intelligence (advanced)
```

```
Example task: "Train new employees on safety procedures and equipment
   operation"
Activity: train and teach colleagues
Skills: verbal communication (intermediate), technical knowledge about
   equipment (intermediate), instructional planning (basic)

Example task: "Supervise and coordinate the work plan of customer
   service representatives and schedule shifts"
Activity: manage team operations
Skills: operational planning (intermediate), verbal communication
   (advanced), people development (advanced)

Example task: "Develop marketing strategy for new product launch"
Activity: create marketing strategies
Skills: strategic thinking (advanced), business knowledge (expert),
   creativity (intermediate), analytical reasoning (intermediate),
   written communication (advanced)

Example task: "Read operating schedules or instructions or receive
   verbal orders to determine amounts to be pumped."
- Activity: follow operational instructions
- Skills: reading comprehension (basic), verbal communication (basic)
</Examples>
```

User Prompt.

```
<List of task statements>
Here is the list of {len(tasks_chunk)} job task statements to analyze,
   along with their index numbers:

{task_list}
</List of task statements>
```

Task	Activity	Skills	Cluster
Smooth rough spots on walls and ceilings, using sandpaper	smooth surfaces	manual dexterity (basic), attention to detail (basic)	Performing Detailed Manual Tasks
Lubricate moving parts on gate-crossing mechanisms and swinging signals	lubricate moving parts	manual dexterity (basic), attention to detail (basic)	Performing Detailed Manual Tasks
Perform physically demanding tasks, such as digging trenches to lay conduit or moving or lifting heavy objects	perform physical labor	physical endurance (advanced), manual dexterity (intermediate)	Performing Physical Labor
Prepare reports of activities, evaluations, recommendations, or decisions	prepare reports	report writing (advanced), analytical reasoning (intermediate), attention to detail (intermediate)	Processing and Analyzing Records
Confer with officials of public health and law enforcement agencies to coordinate interdepartmental activities.	coordinate interdepartmental activities	collaboration (advanced), project management (advanced), communication skills (intermediate)	Coordinating Project Initiatives

Table B.1: Examples: detailed tasks, extracted characteristics, and cluster assignment

Notes. This table lists examples of detailed tasks (first column), that is the input, as well as the extracted core activity and skill requirements (LLM-generated), and the labeled cluster to which this task is assigned.

B.2.2 Embeddings and clustering

We use Alibaba's `gte-Qwen2-1.5B-instruct` model to create word embeddings of dimension 1,536 for the extracted skills for each task statement. To prepare the embeddings data for clustering, and noting that the HDBSCAN algorithm we are using performs best on data with low to medium dimensionality, we next perform a two-part dimensionality reduction step. We initially perform PCA, retaining the principal components that explain 95% of the variance in the embedding space. We then perform a subsequent UMAP step, which is useful to preserve both local and global data structures while shrinking the number of components to a level suited for the HDBSCAN algorithm. Finally, we use the HDBSCAN algorithms with the following hyperparameters `min_cluster_size` = 70, `min_samples` = 40, `cluster_selection_epsilon` = 0.05. The distance metric `option_hdbscan_metric` is Euclidean, aligned with the preceding UMAP step.

B.2.3 Labeling step & summary output

Finally, we use OpenAI's o3-mini-high model to create natural-language labels and a summary description for each of the task clusters. These cluster-level meta data are useful in two ways: in terms of interpretation, and as inputs to the LLM when constructing the occupation-level time shares across the task clusters. Practically, for each cluster we randomly select ten representative tasks and feed the core activity as well as the skill requirements for these tasks to the LLM, instructing it to generate a cluster label and a brief description, per the following prompts.

Table B.2 details all 38 task clusters, indicating the summary label and description.

System Prompt.

```
 <Role> You are a world-class expert in labor economics, task
   classification and  occupational analysis.  You use concise and
   standardized language that is consistent with established
   terminology in skills/occupational databases like O*NET or PIACC.
</Role>

<Overall goal>
The overarching goal is to create accurate and meaningful summary
   labels for clusters of job tasks.
Each cluster comprises many tasks, which grouped by the type of
   activity (what is being done) and the skills required (capacities
   that facilitate performance of activities); i.e., the general rule
   is that a person proficient in one task in a given cluster should
   also be able to perform others in that cluster.

Given this goal, you will be presented with a list of tasks --
   alongside the most essential skills required to perform each --
   that exemplify a particular cluster.

You primary task is to create an accurate and concise summary label
   for this cluster of tasks.
Your secondary task is to provide a concise description of this
   cluster, with reference to core skill requirements differentiating
   this cluster from others.

Requirements for the summary label:
```

```
- The label summarizes the common core activities (what is being
   done), while remaining specific enough to meaningfully
   differentiate this cluster from others.
- The label focuses on the essential underlying activity rather than
   the specific domain.
- The label is sufficiently specific to allow differentiating between
   occupations that have different skill requirements and wage levels.
- The label is concise (2-5 words),  uses natural sounding language
   aligned with established task/skill terminology,
    and where possible begins with a gerund (verb+ing form).

Requirements for the description:
- The concise description (1 sentence) summarizes the cluster, with
   reference to core skill requirements differentiating this cluster
   from others.
</Overall goal>

<Detailed instructions>
Step 1: Analyze the {len(tasks_chunk)} tasks by identifying the
   fundamental activities involved and core skills utilized across all
   them.

Step 2: Create a summary cluster label that satisfies the requirements
   outlined above.
Test your label to ensure that it meets each of the X requirements;
   revise and iterate until this is the case.

Step 3: Given the label, and considering the skills listed for the
   exemplary rasks, provide a concise description.
</Detailed instructions>

<Examples of cluster labels>
Here are examples of cluster labels to illustrate the desired level of
   abstraction. These serve for guidance only, you must create
   appropriate task-specific labels.
```

```
- Positive example: "Developing and Building Teams" (relevant across
   domains, but specific enough to distinguish from other
   interpersonal tasks)
- Positive example: "Analyzing quantitative data" (relevant across
   different occupations, distinct from qualitative analysis which
   would involve different skills)
- Positive example: "Performing gross motor or heavy manual physical
   labor" (connotes a broad range of tasks with similar skill
   requirements)
- Positive example: "Technical Operation and Maintenance Tasks" (not
   domain specific, connotes a skill requirement distinct from
   advanced technical analysis)
- Negative example: "Getting Information" (too unspecific)
- Negative example: "Performing Administrative Activities" (too broad,
   could involve routine tasks such as processing paperwork or
   advanced managerial tasks, i.e. tasks requiring very different
   skills)
- Negative example: "Communication" (too unspecific, could comprise
   anything from chatting with colleagues to arguing a complex case in
   court)
</Examples>
```

User Prompt.

```
 <List of tasks to analyze>
Here is the list of {len(task_descriptions)} tasks that are
   representative of the task cluster under consideration alongside
   the most important skills required to perform them:

{task_list}
</List of tasks to analyze>
```

B.3 LLM-generated time diaries

This section describes how, given the task clusters, we construct the occupational task weight matrix. In addition, we detail validation exercises.

Cluster label	Description
Performing Detailed Manual Tasks	This cluster involves executing precise, hands-on operations—ranging from cleaning and lubricating to marking and packaging—that rely on basic manual dexterity and careful attention to detail.
Performing Precision Finishing Tasks	This cluster encompasses tasks that involve fine manual adjustments and finishing operations—such as aligning, smoothing, and testing components—requiring intermediate manual dexterity and attention to detail.
Preparing and Planning Meals	This cluster involves tasks that span cooking, menu planning, and overseeing food safety and service, requiring strong culinary skills, dietary knowledge, and attention to detail.
Maintaining Records and Inventory	This cluster involves routine operational support tasks that require diligent record keeping, inventory management, and clear communication to sustain documentation, asset tracking, and service functions.
Coordinating Detail-Oriented Operations	This cluster involves routine tasks such as sorting, record-keeping, material distribution, and facility upkeep that require meticulous attention to detail and basic to intermediate organizational skills.
Delivering Public Presentations	This cluster involves speaking in formal and public settings—ranging from project briefings and lectures to courtroom testimonies—requiring advanced public speaking, communication, and subject matter expertise.
Documenting Technical Information	This cluster focuses on capturing and recording technical details and processes using advanced technical writing, documentation, and attention to detail.
Performing Clinical Procedures	This cluster involves executing patient-focused clinical tasks that combine advanced diagnostic reasoning, technical equipment operation, interpersonal communication, and therapeutic interventions to assess and treat medical conditions.
Providing Customer Service	This cluster involves direct customer interactions that require strong interpersonal, communication, time management, and organizational skills to assist, guide, and support various client needs in service-oriented settings.
Administering Regulatory Compliance	This cluster involves interpreting policies, reviewing and enforcing regulatory standards, and developing procedures, all requiring advanced regulatory knowledge, analytical reasoning, and communication skills.
Coordinating Emergency Response	This cluster involves executing and managing emergency procedures, crisis communication, threat monitoring, and strategic planning, requiring advanced emergency response and situational awareness skills.
Maintaining and Managing Records	This cluster involves systematically updating, retrieving, and organizing diverse records and data through strong attention to detail and organizational skills.
Reviewing and Editing Information	This cluster involves accurately reviewing, editing, and verifying various forms of information—from written materials to operational data—requiring advanced attention to detail and precision.
Ensuring Regulatory Compliance	This cluster involves meticulous inspection, record management, and analytical review to verify adherence to regulatory standards and operational protocols.
Performing Physical Labor	This cluster encompasses physically demanding tasks that require manual dexterity, physical endurance, and fundamental technical and safety skills across diverse settings including construction, cleaning, material handling, animal care, and exercise instruction.
Creating Technical Visual Representations	This cluster involves transforming data, technical specifications, and artistic ideas into precise visual media by integrating advanced drafting, design, and multimedia editing skills.
Designing and Implementing Systems	This cluster centers on planning, designing, and integrating technical systems across diverse fields, emphasizing advanced project management, engineering design, and technical expertise.
Processing and Analyzing Records	This cluster involves tasks focused on maintaining, recording, and evaluating data—including financial, production, and medical records—where strong numerical reasoning, analytical skills, and meticulous attention to detail are essential.
Operating, Calibrating, and Inspecting Equipment	This task cluster involves technical operations focused on handling electronic recording, imaging, and sound equipment, requiring precise calibration, systematic inspections, and adept problem-solving skills.
Inspecting and Evaluating Quality	This cluster involves detailed inspections and analyses that rely on advanced analytical reasoning and attention to detail to assess product, site, and process quality, ensuring standards and performance are met.
Performing Skilled Manual Operations	This cluster involves executing diverse manual tasks—ranging from assembly, finishing, and equipment maintenance to operation and cleaning—that require intermediate to advanced manual dexterity, attention to detail, and technical proficiency.
Negotiating and Coordinating Contracts	This cluster involves engaging stakeholders through advanced negotiation and communication skills to secure agreements and manage procurement activities while coordinating legal, regulatory, and project management requirements.
Repairing and Maintaining Equipment	This cluster encompasses preventative maintenance, technical repair, and equipment installation tasks that require advanced system knowledge, manual dexterity, and safety awareness.
Managing Safety Operations	This cluster involves overseeing operational activities with a strong emphasis on safety compliance, hazard assessment, and technical oversight across diverse industrial, emergency, and technical settings.
Monitoring and Inspecting Systems	This cluster involves actively operating, adjusting, and inspecting automated processes and equipment by employing advanced technical troubleshooting, precision measurement, and quality control skills to ensure optimal system performance.
Analyzing and Optimizing Systems	This cluster involves applying advanced technical analysis, simulation, and maintenance skills to assess performance, recommend design changes, and ensure operational integrity across diverse systems.
Analyzing Natural Phenomena	This cluster involves applying advanced scientific analysis, technical expertise, and data interpretation to evaluate, classify, and redesign natural and biological systems across diverse domains.
Instructing and Training	This cluster involves delivering instruction, training, and mentorship across diverse subject areas, relying on advanced instructional techniques, verbal communication, and subject matter expertise.
Mediating and Consulting Clients	This cluster involves interpersonal guidance tasks—including counseling, referrals, conflict investigation, and dispute resolution—that require advanced communication, empathy, and problem-solving skills to address diverse client issues effectively.
Developing and Delivering Instruction	This cluster encompasses tasks centered on planning, designing, and conveying educational programs and curricula, leveraging advanced instructional design, curriculum development, and communication skills across varied content areas.
Communicating and Educating	This cluster involves effectively conveying information, instructions, and feedback through verbal channels, integrating clear reporting, problem-solving, and instructional skills across diverse contexts.
Engaging in Continuous Learning	This cluster encompasses tasks that require ongoing research, information synthesis, and professional development to remain current with industry trends, technology advancements, and scientific progress.
Collaborating Across Functions	This cluster comprises tasks requiring effective teamwork, communication, and coordination across diverse professional areas to address problems, manage operations, and support technical and client-oriented activities.
Coordinating Project Initiatives	This cluster involves planning, overseeing, and collaborating on diverse project tasks, leveraging advanced project management, communication, and leadership skills.
Coordinating Administrative Tasks	This cluster encompasses planning, scheduling, and organizing a range of administrative operations, requiring strong organizational, communication, and project management skills.
Coordinating Strategic Initiatives	This cluster involves planning, organizing, and supervising diverse activities—ranging from educational events to disaster recovery and recruitment—requiring advanced leadership, strategic planning, and team management skills.
Producing Technical Documentation	This cluster involves drafting and compiling technical reports, proposals, and documentation through advanced technical writing, analytical reasoning, and data presentation skills, with elements of programming and research support.
Performing Strategic Analysis	This cluster involves advanced quantitative research, financial and cost analyses, and strategic planning to assess deviations, forecast outcomes, and drive management recommendations.

Table B.2: Task cluster labels and descriptions

B.3.1 Methodology

To generate the time diaries we use the latest version of GPT-o3-mini-high for this step. We loop over each occupation using the following prompts. These are designed to break the complex task into clear sequential steps, draw on high-quality inputs, and convert the qualitative assessment into a numerical output.

System Prompt.

```
You are an expert in occupational classification (the system being
   used is {occ_system} at the {occ_level}-digit level) and analyzing
   occupational time allocation.
You combine precision in classification with deep knowledge of how
   different occupational groups allocate their time across tasks.
You focus on accurate, structured data output, and your time share
   predictions MUST sum to exactly 1.0. You are precise and
   conscientious.
```

User Prompt.

```
<Objective and context>
We want to accurately estimate what percentage of their work time
   workers in a specific occupation group spend on various tasks.
The occupation group is {occ_title}, as classified following the
   occupational classification system {classification_description}.
The reference period to consider is the {option_timeperiod}.
</Objective and context>

<List of tasks>
The tasks to consider are as follows:

{task_list}
</List of tasks>

<Instructions>
Follow these steps to generate accurate time allocation estimates:

1. Analyze core functions, activities and responsibilities of
   {occ_title}
```

```
2. For each task listed above:
   - Review the task carefully
   - Assess the importance and frequency of this task for {occ_title}
     in the {option_timeperiod}, drawing on high-quality evidence,
     expert knowledge and statistical data.
3. Having done this for all tasks, convert assessments to time
    allocation shares:
     For each task:
   - Convert assessment to percentage of work time
   - Translate to decimal (e.g., 25%  0.2500)
   - Document: task_name: 0.XXXX
   - Add to running_sum
4. Verification (required):
   Calculate total_sum to 4 decimals
   If total_sum != 1.0000:
   - Calculate scaling = 1.0000/total_sum
   - Multiply EACH share by scaling
   - Recalculate sum
   STOP: Submit shares only if sum = 1.0000

Critical Requirements:
- Use 4 decimal precision throughout
- Show calculations
- Final shares MUST sum to 1.0000
- No rounding of intermediate values
- Calculate time shares for all {task_count} tasks.
</Instructions>
```

In rare instances, the LLM does not generate time shares that sum to 1, despite the above instructions. This is reminiscent of human responses in time diary surveys. We therefore programmatically normalize the LLM-predicted shares, just as we do using the conventional, human survey responses discussed below. Over the course of the project we moved from OpenAI 4o to o3-mini-high, which greatly reduced the need for this ex-post normalization of time shares.

B.3.2 Validation

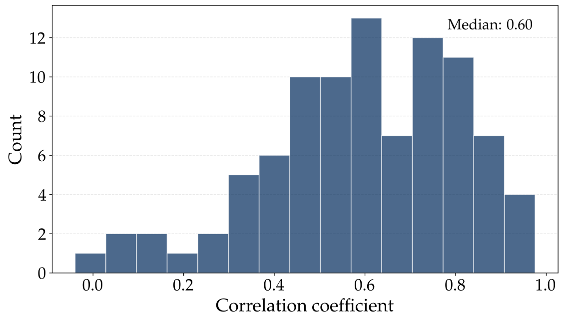
The use of LLMs in measurement invites some immediate questions: How could the LLM know this information? And are the resulting measurements reliable? Regarding the first question, LLM training data comprises virtually the entire internet, including vast amounts of unstructured data on what people across different occupations do at work, as well as summaries of time diary surveys reported in research papers. Since these data sources generally do not reference our exact tasks or occupations, and much input data is qualitative, the LLM's quantitative output results from interpolation. Given the black-box nature of this data construction, we perform four complementary exercises that collectively demonstrate the robustness of the LLM-based measurement of the occupational task weights.

First, we compare LLM-generated task weights at the occupation-cluster level to the average importance rating that O*NET assigns to detailed tasks within each cluster. While O*NET importance weights do not directly map onto our A matrix entries—unlike time shares—they are strongly correlated with our baseline measures. Second, we exploit a unique 2012 supplemental survey by the German Federal Institute for Vocational Training (Bundesinstitut fuer Berufsbildung, BIBB) in which workers across many occupations report their time allocation across 17 tasks. Though the occupations and tasks differ from our baseline analysis, our LLM-based method is flexible enough to generate time diaries for German BIBB classifications. This comparison reveals highly correlated time shares between the two approaches. Third, we use O*NET's Generalized Work Activities as a task classification, with importance ratings as weights. LLM-generated time shares for these activities again align strongly with importance ratings. Finally, we establish LLM internal consistency: occupational task weights constructed by averaging across constituent minor categories are highly correlated with those derived by directly querying the model for major groups. This section provides more details.

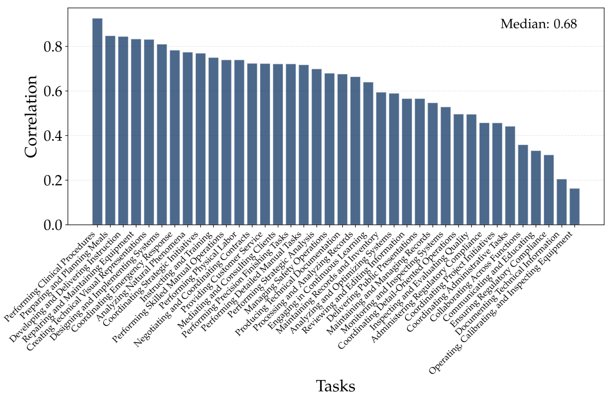
O*NET importance weights. O*NET provides for each 8-digit occupation not only a list of detailed task statements—from which we constructed our task clusters—but also assigns to each task a categorical rating on a Likert scale from 1-5 that indicates the importance of this task to its associated occupation.[B.3]

We begin aggregating up the importance ratings to the level of our aggregate task clusters by occupation. To do this, we first collapse occupations to the SOC-2019 minor group level. We then create weights for each occupation-cluster pair. These weights correspond to the shares of the detailed tasks associated to that occupation group that belong to a given cluster, weighted by importance ratings. Let k index occupations (at the minor-group level), τ index task clusters, and t index detailed tasks. Further, let $\mathcal{T}_\tau$ denote the set of detailed O*NET tasks associated

[B.3] In addition to "importance," O*NET also provides scales for "relevance" and "frequency." We found that incorporating these scales makes no difference to our results.



(a) Occupation-level correlations



(b) Task-level correlations

Figure B.2: Comparison of LLM-generated task weights to aggregated O*NET importance weights
Notes. This figure compares the A weights obtained via LLM for the post-2000 period to those constructed by averaging over O*NET importance weights The left panel plots the distribution of occupation-level correlation; the right panel shows the task-level correlations across occupations instead.

with cluster τ, $\mathcal{T}_k$ the set of detailed tasks associated with occupation k, and $\omega_{k,t}$ be the weight attached to task τ by occupation k. Then we construct the (relative) weight occupation k puts on cluster τ, $\omega_{k,\tau}$ as

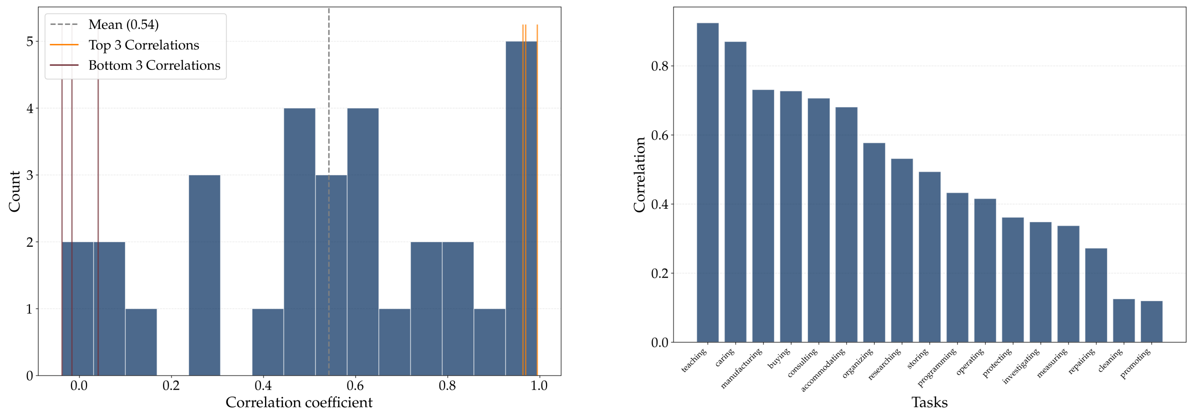
$$\omega_{k,\tau} = \frac{\sum_{t \in \mathcal{T}_\tau} \mathbf{1}_{t \in \mathcal{T}_k} \cdot \omega_{k,t}}{\sum_{t \in \mathcal{T}_k} \omega_{k,t}},$$

where $\mathbf{1}_{t \in \mathcal{T}_k}$ is an indicator function that equals 1 if task t belongs to occupation k and 0 otherwise. That is, the weight occupation k puts on cluster τ is greater if a large fraction of the detailed tasks associated to k are linked to τ or if those tasks have especially high importance weights for k. Next, the SOC-2019 occupations are cross-walked to the SOC-2000 classification used in our analysis using the official crosswalks available from https://www.onetcenter.org/taxonomy.html.

Figure B.2 shows that the occupational task weights obtained through this method exhibit a strong positive correlation to those generated via the LLM-based method. The median correlation is 0.6 (across occupations), respectively 0.68 (across tasks).

BIBB time diaries. We use a supplemental survey conducted for the 2012 Employment Survey carried out by the German Federal Institute for Vocational Training (Bundesinstitut fuer Berufsbildung, BIBB) and the German Federal Institute for Occupational Safety and Health (BAuA). This survey asks a subset of surveyed workers to report their allocation of time to a pre-specified list of tasks such as "teaching" and "cleaning" on a given day.[B.4]

[B.4]The full list of 17 tasks is as follows: 'investigating', 'organizing', 'researching', 'programming', 'teaching', 'consulting', 'buying', 'promoting', 'repairing', 'accommodating', 'caring', 'cleaning', 'protecting', 'measuring', 'operating', 'manufacturing', 'storing'



(a) Occupation-level correlations

(b) Task-level correlations

Figure B.3: Comparison of LLM-generated task weights & BIBB survey

Notes. The left panel plots the distribution of occupation-level correlations between the LLM-predicted task weights and those constructed from the BIBB. The right panel shows the task-level correlations across occupations instead.

We proceed in three steps. First, we construct occupation-task level time allocation shares from the BIBB. We consider the sample of individuals in West Germany aged 15-65 who have completed their training and who report a valid occupation ISCO-08 2-digit occupation. For each individual, we normalize the time shares to sum to one. Then we average time shares across occupations and retain those occupations comprising at least ten surveyed workers. Second, we re-run the same LLM-based process as in the main analysis, but now requiring responses for the same set of tasks considered in the BIBB and looping over ISCO-08 2-digit occupations. Third, we compare the BIBB-based and LLM-based responses.

Overall, the two different approaches yield highly comparable results. Figure B.3a shows correlations at the level of occupations; the mean correlation is 0.54, the standard deviation is 0.31. The lowest correlations are reported for "Customer Service Clerks," "General and Keyboard Clerks" and "Numerical and Material Recording Clerks." A major source of discrepancy is that for these occupations, survey respondents in the BIBB put substantial weight on the task "programming" (which in the original German language context could also be interpreted as "using a computer"). With further clarification on the interpretation of tasks, we expect that the LLM and BIBB would yield results that are more comparable still. Figure B.3b reports task-level correlations of weights across occupations. The tasks with the lowest overlap are "promoting" and "cleaning," while the alignment is greatest for "teaching" and "caring."

Comparison to O*NET importance weights for GWAs. Next, we compare LLM-generated time allocation shares with O*NET occupation-level importance weights for Generalized Work Activities (GWAs). We use the GWAs from ONET 5.0, as this database aligns with the SOC-2000 classification used in our main empirical analysis. We construct relative importance weights for

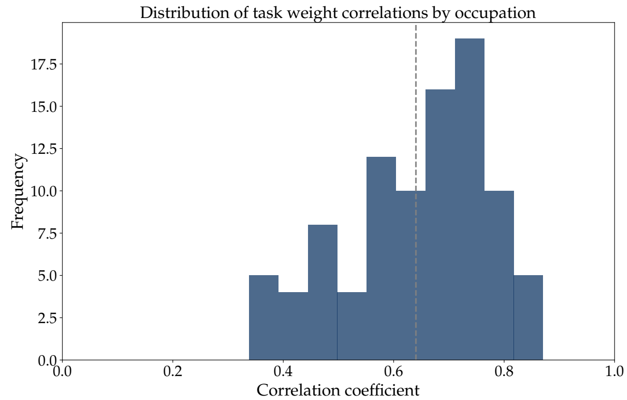


Figure B.4: LLM-time shares for GWAs correlate with O*NET importance weights

each GWA by occupation and aggregate to the minor-group level. We then generate LLM-based time allocation shares for identical GWAs across the same occupational categories and compare the resulting two A matrices.

Figure B.4 displays the distribution of occupation-level correlations between LLM-generated time shares and O*NET importance weights. The distribution is markedly right-skewed, with a central tendency around 0.6-0.7, indicating substantial alignment between our LLM-based approach and established occupational measurements. Figure B.5 presents task-specific correlations across occupations, grouped by correlation strength. Tasks involving cognitive and managerial functions show the strongest correspondence (correlations >0.75), while more specialized technical tasks exhibit moderate alignment. Even the lowest-correlating tasks maintain coefficients above 0.2, suggesting our approach captures meaningful variation across the entire task spectrum.

LLM consistency in aggregation across occupational hierarchies. Figure B.6 demonstrates that the LLM-generated task weights are consistent across different levels of occupational aggregation. We compare weights derived directly from major occupational groups with those constructed by averaging across their constituent minor occupational categories (using unweighted means). The very high correlation coefficient (0.89) confirms that our task approach maintains consistency regardless of aggregation level.

B.4 Occupational labor shares

To construct occupation-level labor shares, i.e., compensation over value-added, we take the following approach, where industries are indexed by j and occupations by o:

(i) Construct weights s_{oj} corresponding to the share of industry-j payments to labor going to

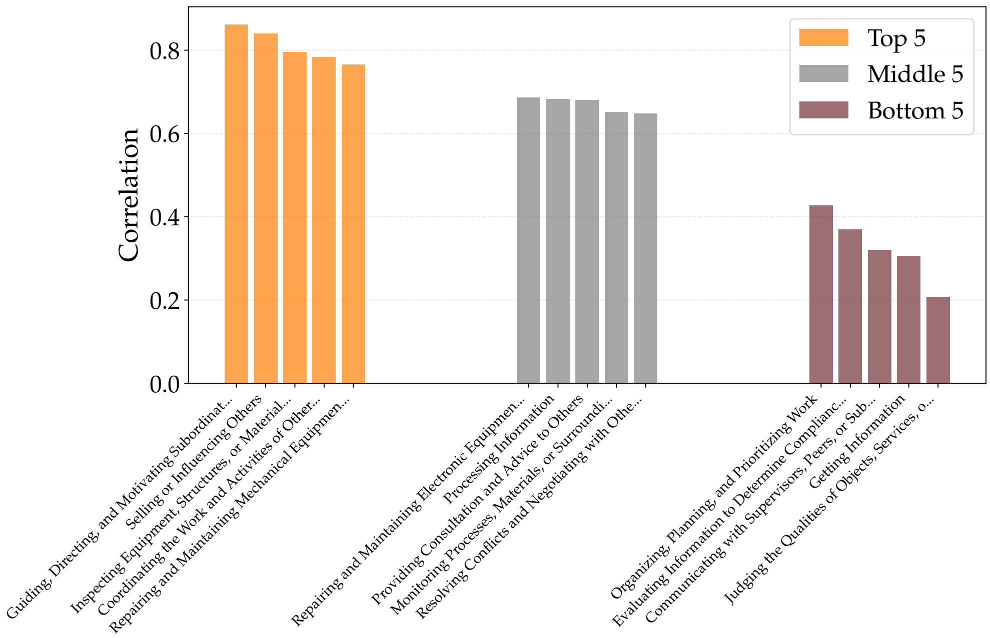


Figure B.5: Correlation across occupations by task

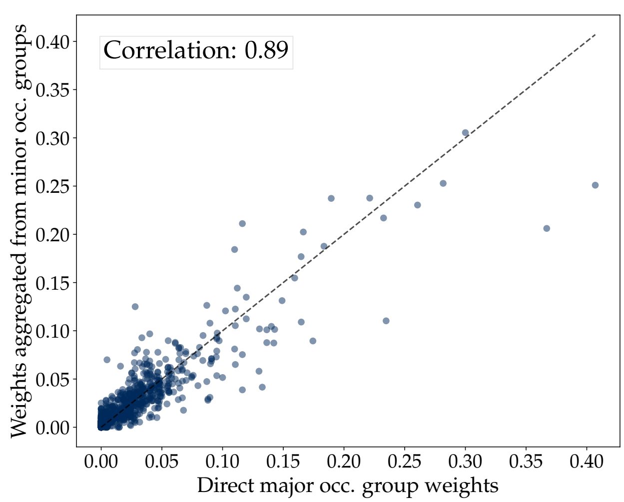


Figure B.6: Comparison of task weights at different occupational levels of aggregation occupation o:

$$s_{oj} = \frac{\text{(wage payments to } o \text{ in } j)}{\sum_o \text{(wage payments to } o \text{ in } j)} \tag{B.4}$$

(ii) Assume that value-added in j due to o is proportional to s_{oj}:

$$\text{VA}_{oj} = s_{oj} \cdot \text{VA}_j \tag{B.5}$$

(iii) Compute

$$\text{LS}_o = \frac{\sum_j \text{wage payments to } o \text{ in } j}{\sum_j \text{VA}_{oj}} \tag{B.6}$$

In practice, we use the 2002 wave of the BLS Occupational Employment and Wage Statistics (OEWS) to construct s_{oj}. This wave uses the same SOC-2000 occupational classification as our (harmonized) NLSY dataset and NAICS-2002 industry codes. Data on VA_j come from the BEA-BLS Integrated Industry-level Production Accounts (1987-2020). In addition, to construct the numerator of equation (B.6) we use the same apportionment method as in equation (B.5).[B.5] Industry-level data are averaged across sample years. We then link the OEWS data on s_{oj} with the BEA/BLS industry-level data by merging at the 2-digit NAICS level, retaining only those industries with a 1:1 mapping.[B.6]

The (unweighted) average labor share across occupations is 0.61, with a minimum of 0.49 (Farming, Fishing, and Forestry Occupations) and a maximum of 0.75 (Legal Occupations).

B.5 Additional figures and tables

[B.5] Using the wage bill information from the OEWS instead suffers from the problem that magnitudes of compensation differ from those in the BEA-BLS accounts; using the latter is, therefore, internally more consistent.

[B.6] The BEA/BLS data provide a crosswalk from the "production account classes" to NAICS-2007; NAICS-2007 and NAICS-2002 are identical at the 2-digit level.

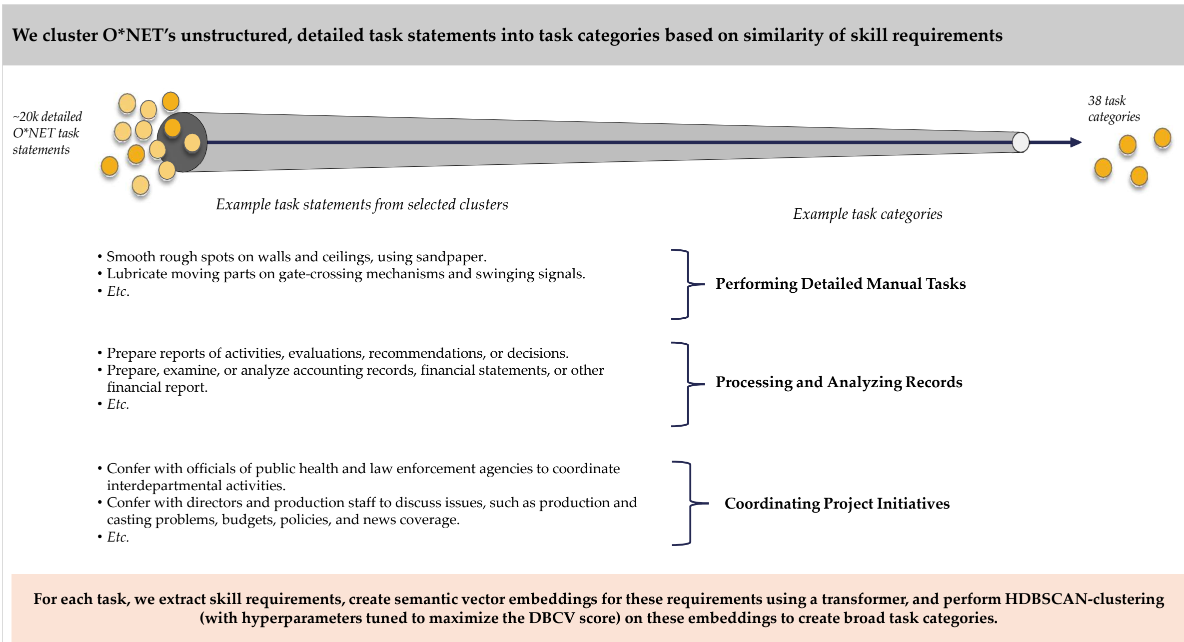
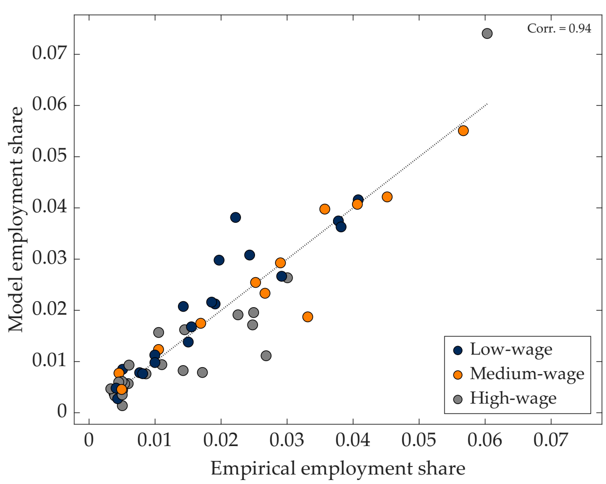
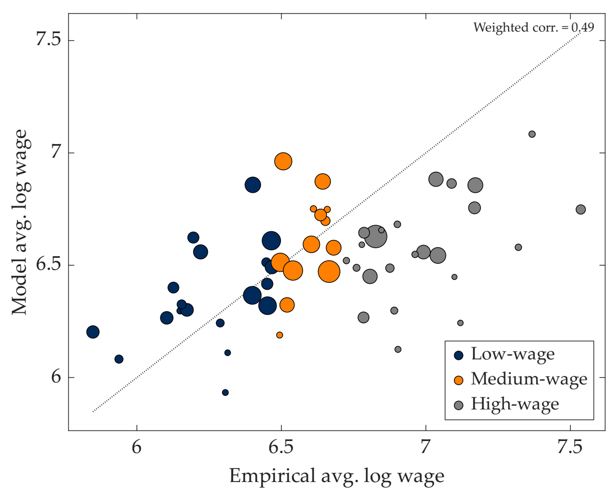


Figure B.7: Examples of mapping from detailed tasks to clusters



(a) Employment shares



(b) Wages

Figure B.8: Model vs. data: occupation-level variables

Notes. In the right-hand panel, the correlation is weighted by employment shares.

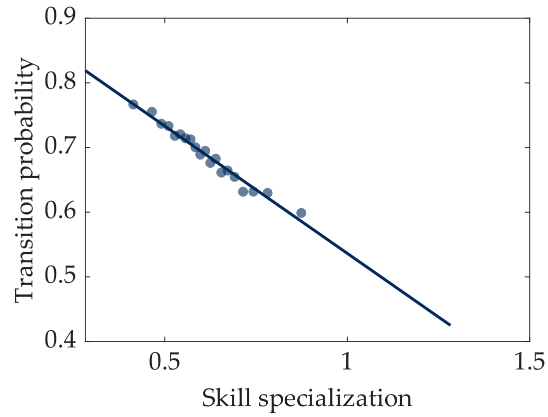
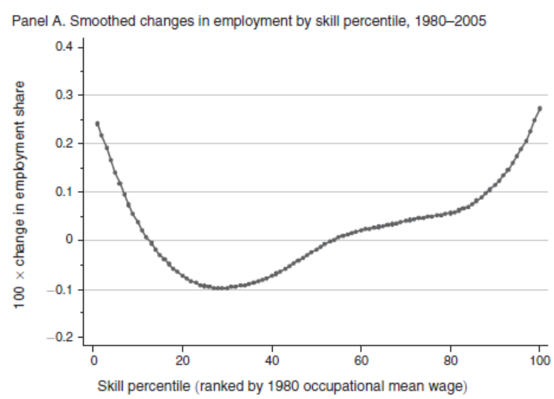
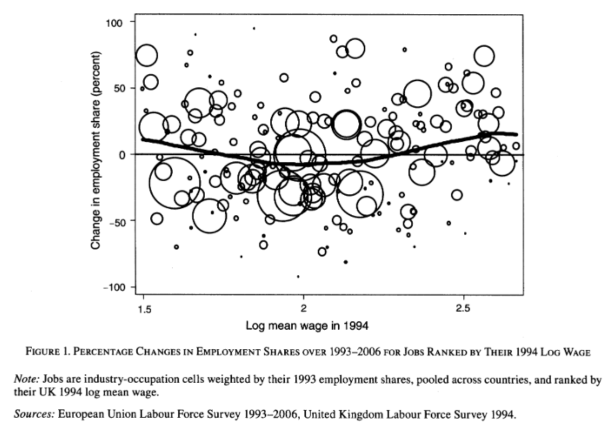


Figure B.9: Specialization generates occupational persistence

Notes. This figure shows a binscatter plot at the individual-level, relating the normalized frequency of occupation switches to the coefficient of variation of skills.



(a) Autor and Dorn (2013)



(b) Goos *et al.* (2009)

Figure B.10: Employment polarization according to the empirical literature

Notes. The left panel shows Figure 1, panel A from Autor and Dorn (2013), which is based on data from the U.S. Census and American Community Survey. The right panel shows Figure 1 from Goos *et al.* (2009), which is based on labor force surveys for the EU and the UK.

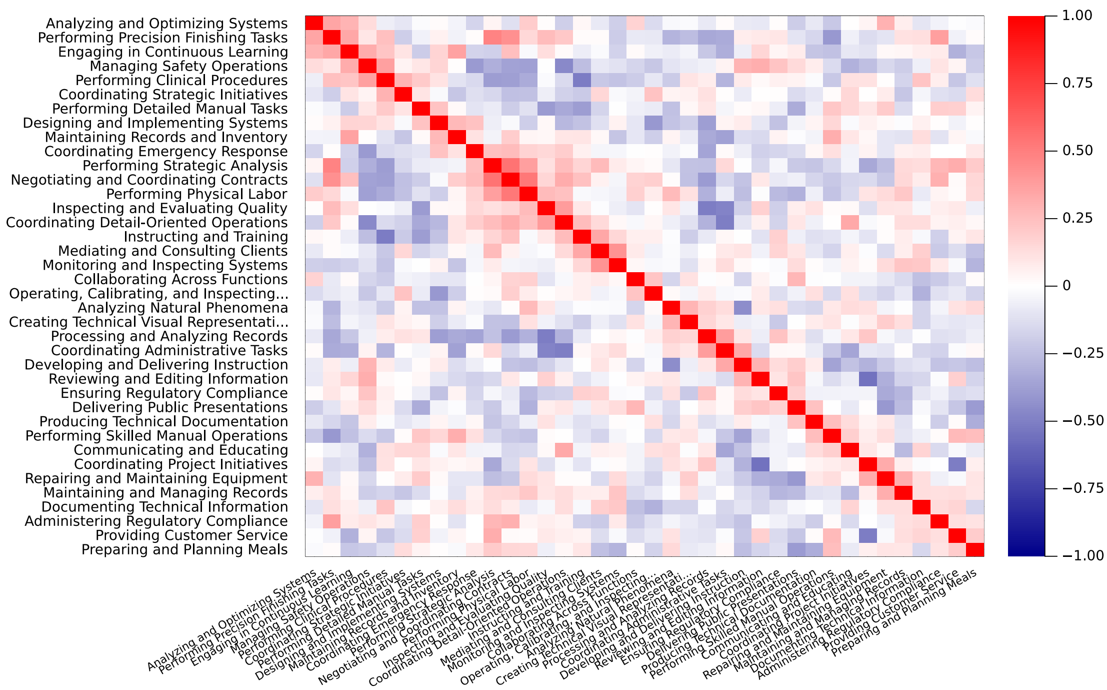


Figure B.11: Skill correlations

Notes. Estimated skill correlations are indicated by colors.

C Appendix: The Labor Market Results of LLMs

C.1 Historical examples of job transformation

This section illustrates job transformation using two historical case studies. First, consider weavers' work over the course of the 19th century. We draw on Bessen's (2011) analysis of records from the Lawrence Company, which operated several mills in Massachusetts. Table C.1 shows how weavers' duties shifted away from a diverse set of manual tasks like preparing, dressing, and letting off the reed, toward spending more time on a narrower set of tasks, like fixing mechanical issues or replacing inputs such as bobbins.

Period	**Preparatory tasks**		**Tasks while machine running**							**Tasks while power loom stopped**							
	Prepare warp	Dress warp	Let off warp	Pick shuttle	Beat reed	Take up cloth	Adjust warp tension	Replace empty bobbin	Monitoring	Fix smashes	Adjust temples	Back up loom	Replace empty shuttle	Fix broken weft	Fix broken warp end	Remove cloth, cleaning	Replace warp
Handloom	●	●	●	●	●	●	●		●		●		●	●	●	●	●
Early power loom (~1820)							●	●	●	●	●	●	●	●	●	●	●
1833							●	●	●	●		●	●	●	●	●	●
1883							○	●	●	●			●	●	●	●	○

Table C.1: Job transformation: the case of weavers

Notes. Legend: ● = Task performed; ○ = Reduced frequency; Empty = Task not performed.
Replication of Bessen (2011, Table 2).

Second, machinists in manufacturing experienced job transformation during the late 20th century. Bartel *et al.* (2007) study how investments in IT-enhanced machinery transformed skill requirements in valve manufacturing. These enhancements consisted of capital-embedded improvements in computer numerically controlled (CNC) tools, which fix the raw material and automatically machine valve components based on designs entered into the machine's operating software. Previously, operators used routine machining skills and physical strength to position valves correctly and had to choose and move appropriate cutting tools. With more sophisticated CNC machines, work shifted toward setup, monitoring, and adjustment – tasks demanding greater programming and problem-solving skills. Atalay *et al.*'s (2020) systematic analysis of job advertisements corroborates this account. They show that machinists experienced major task shifts in the 1980s with CNC adoption, with increased emphasis on computer programming, problem-solving, and technical engineering knowledge.

C.2 Automation exposure measures: Webb (2019)

Figure C.1 shows the average standardized exposure score of each task cluster for the three types of technology considered by Webb (2019): AI, Robots, and Software.

It can be observed that the task cluster identified as most exposed to AI in Webb (2019) is "Analyzing Natural Phenomena" which, according to the LLM's summary description, involves "applying advanced scientific analysis, technical expertise, and data interpretation to evaluate, classify, and redesign natural and biological systems across diverse domains." By contrast, "Processing and Analyzing Records," our primary example of a task category exposed to LLMs has a close to average exposure score.

This difference is indeed to be expected. The technology cluster labeled by Webb (2019) as "AI" comprises a broader set of tools, including neural networks and deep learning algorithms more broadly, compared to the study by Eloundou *et al.* (2023), which explores task-level exposure to LLMs more specifically. Thus, it is indeed to be expected that different tasks would be exposed to the two technologies, respectively. Our framework suggests that this difference is potentially important, even if the degree of automation of two distinct tasks, as measured by the decline in the labor share, for instance, may carry labor market consequences that differ in important ways depending on the bundles of tasks the exposed tasks form part of and the distribution of task-specific skills.

C.3 Additional results, figures and tables

C.3.1 Occupation-level outcomes

The main text presents results at the level of individual workers. This section instead reports *occupation*-level outcomes. Figure C.2 displays changes in employment by occupation, aggregated to the two-digit level for ease of interpretability. The occupations projected to decline the most in terms of employment tend to be white-collar professions such as Life, Physical, and Social Science, Architecture and Engineering, or Legal Occupations. Figure C.3 displays the change in average wages at the occupation level, plotted against exposure. The relationship displays a similar inverted-U shape as seen in Figure 13, but partly reflects the change in skill composition induced by occupational switching.

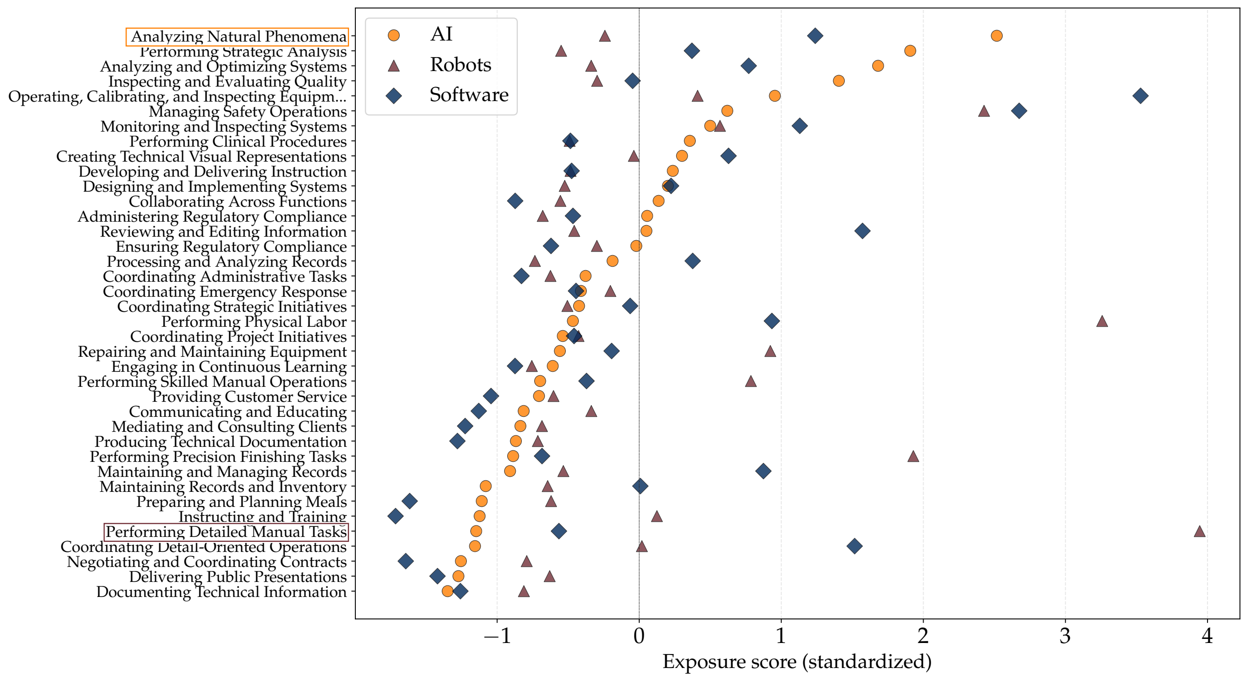


Figure C.1: Technology-specific exposure scores at the task level (Webb, 2019)

Notes. This chart shows for each of three technologies the standardized exposure score for our task clusters based on the results of Webb (2019). In Webb's approach, AI patents are identified by terms like "neural network," "deep learning," or "generative model" in titles or abstracts. Software patents contain terms such as "software" or "program" while excluding hardware-related terms like "chip" or "circuit." Robot patents are selected through the inclusion of the term "robot" in titles or abstracts.

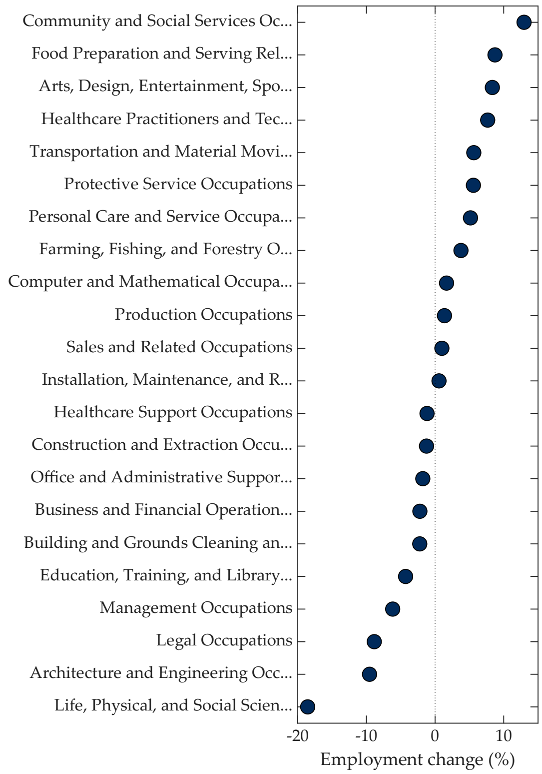


Figure C.2: Changes in occupation-level employment
Notes. This figure displays the percent change in employment at the level of two-digit occupations.

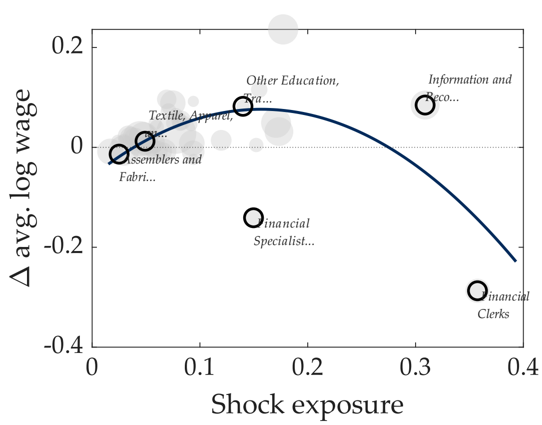


Figure C.3: Changes in occupation-level average wages
Notes. Each dot is an occupation. The vertical axis measures the difference in the average log wage after vs. before the shock. The horizontal axis measures shock exposure. Dot sizes correspond to pre-shock employment shares.