AI Automation and Labor Market Outcomes[§]

Yakup Kutsal Koca[*]

August 18, 2026

Please click here for the most recent version

Abstract

Existing measures of occupational AI exposure ignore worker reallocation, failing to capture true economic impact. In this study, I address this by developing a general equilibrium framework that explicitly models the occupational choice problem for workers. Using administrative German data, I recover the unobserved, heterogeneous comparative advantage vectors that drive these choices. I then simulate a productivity-enhancing AI shock derived from task-level automation scores. The results show that the ability to reallocate determines the distribution of welfare gains: *generalists* capture the largest relative wage gains by pivoting to high-growth sectors. Conversely, *specialists*, defined by concentrated comparative advantages, experience significantly smaller gains. This relative penalty affects both low-skill workers trapped in stagnating tasks and high-skill professionals forced into costly transitions, as the limited transferability of their skills restricts them from from accessing the highest returns in the labor market.

JEL Codes: J01, J20, J24, J62, E24, O33.

[§]The data access was provided via on-site use at the Research Data Centre (FDZ) of the German Federal Employment Agency (BA) at the Institute for Employment Research (IAB) and subsequently remote data access.

[*]University of Southern California. E-mail address: koca@usc.edu

1 Introduction

Automation technologies have had significant impact on the wage structure in the United States (Acemoglu and Restrepo (2022)). AI technologies that can potentially automate many tasks may have significant implications for the labor market (Trammell and Korinek (2023), McElheran et al. (2024)), especially considering the recent improvements in these technologies and the widening use cases (Bick, Blandin, and Deming (2024) and Handa et al. (2025)). There are several studies so far that measure AI exposure scores for occupations (among them, Brynjolfsson, Mitchell, and Rock (2018), Webb (2020), Felten, Raj, and Seamans (2018), Eloundou et al. (2023), and Handa et al. (2025)). However, the information content of these measures may be lacking in terms of understanding the welfare impacts, since these studies do not take into account the worker reallocation resulting from the AI technologies shifting prices and labor demand. Rather than simply determining who is displaced, these general equilibrium effects shape how workers sort into new roles in response to changing incentives. Consequently, the ultimate economic impact depends not just on initial exposure, but on the ability of workers to reallocate across occupations.

With this concern in mind, I study the impact of AI, specifically LLM technologies, on the wage distribution across workers and the employment distribution across occupations. I build a framework where workers can reallocate across occupations as a result of prices changing due to the AI technologies. Central to this analysis is the AI shock, which I measure using task-level automation scores following Eloundou et al. (2023) and Eisfeldt et al. (2023). I model this as a structural change that distorts relative occupational prices, which prompts workers to re-optimize their career choices based on their comparative advantages in the new environment.

Consistent with the recent automation literature (Acemoglu and Restrepo (2022), Acemoglu and Restrepo (2018), Acemoglu, Autor, et al. (2022), and Humlum (2021)), I model production at the task level. In this framework, tasks serve as the fundamental unit of production. AI technologies are capable of automating a subset of these tasks, in which case capital fully replaces labor. Workers reallocate their labor to the remaining non-automated tasks. This automation functions as a productivity shock: as AI takes over automatable components, the marginal product of labor in the remaining tasks increases, raising the effective productivity of workers in that occupation. On the supply side, I characterize how workers sort across occupations through a dynamic discrete choice framework. In every period, forward-looking workers decide whether to remain in their current occupation or switch to a new one, subject to switching costs. Wages are determined by both observable characteristics and an unobserved component: comparative advantage. Because comparative advantage is heterogeneous across both workers and occupations, two workers with identical observable traits may make different occupational choices.

Quantifying the welfare effects of this sorting requires recovering the unobserved comparative advantage vectors. Ideally, identification would rely on observing a worker's wage

history across all occupations to assess their potential productivity in every field. However, because individual employment histories are sparse, such comprehensive data is unavailable. To overcome this identification challenge, I model unobserved heterogeneity using finite latent worker types. Each type represents a group of workers who share a specific, unobserved comparative advantage vector. While an individual worker's history is limited, the pooled employment history of all workers belonging to a specific type spans the full set of occupations, allowing for identification. I estimate these types and the structural wage parameters using an Expectation-Maximization (EM) algorithm following Arcidiacono and Miller (2011). Furthermore, to identify the switching costs separately from the wage parameters, I leverage the finite dependence property, which allows me to express the dynamic relationship between transition probabilities and value functions without solving the full dynamic programming problem at every step of the estimation.

For estimation, I use an administrative German panel data which tracks the employment history of workers between years 1998-2021[1]. This panel dataset contains the employment status, occupational choice and wages of the 2% randomly selected sample of all individuals in Germany. The data includes the occupations and earnings history of workers along with some individual characteristics such as age and schooling. The large number of observed occupation-to-occupation transitions in this dataset allows me to identify the comparative advantage parameters[2]. I map the German occupational codes (KldB-2010) to their US O*NET equivalents for the automation analysis.

With the structural estimates of the comparative advantage vectors and switching costs, I compute counterfactual wages and allocations under the AI shock. AI technologies make workers more productive by releasing time from automated tasks to be reallocated toward non-automated tasks. This additional productivity causes wages to increase, while on the other hand reducing the price of the exposed occupations due to increased quantity. I solve for the new steady state where workers have re-sorted to maximize their lifetime value given the initial distortion in occupational prices caused by the AI shock.

I find that the efficiency of reallocation is the key determinant of welfare in the new equilibrium. *Generalist* workers, those without a sharp comparative advantage in specific fields, capture the largest relative wage gains. Their high task reallocation intensity reflects a strategic pivot to high-growth sectors, allowing them to fully capitalize on the productivity boom. In contrast, *specialist* workers, who possess a distinct comparative advantage in only one or a few occupations, experience significantly smaller relative gains. These workers incur a high opportunity cost when switching, as they must forfeit the specific edge that defines their current productivity. This relative penalty affects both low-skill workers, whose comparative advantage is confined to manual or service tasks, and high-skill professionals forced

[1]While the raw data covers earlier years, I use the data starting 1998. For a discussion on the data cleaning procedure, please see Section 4.

[2]Since most granular task descriptions and AI exposure metrics are tied to the US O*NET classification, my analysis involves mapping the German occupational codes to their O*NET equivalents.

to abandon specific domains, as the limited transferability of their skills restricts them from accessing the highest returns in the labor market.

I contribute to the automation literature, mainly to those concerning AI automation, including works such as Brynjolfsson, Mitchell, and Rock (2018), Eloundou et al. (2023), Felten, Raj, and Seamans (2018), Webb (2020), Humlum and Vestergaard (2024) and Handa et al. (2025), by studying the effects of the AI automation in a general equilibrium that incorporates worker reallocation. While the exposure scores are informative, there is no guaranteed one-to-one relationship between initial AI exposure and the equilibrium wages of workers in that occupation. For example, if highly exposed translators have a comparative advantage confined to the occupation "translators", then they may experience significantly smaller relative wage gains as they are unable to leverage outside options. On the other hand, if another highly exposed group, computer scientists, have a comparative advantage in an adjacent occupation, such as engineering, then they can reallocate to that occupation to fully capture the productivity gains generated by the AI shock.

This study also adds to the previous works that study the general equilibrium effects of automation shocks using a reduced form analysis, such as Acemoglu and Restrepo (2022). In Acemoglu and Restrepo (2022), the authors estimate a propagation matrix that measures how an automation shock to a set of tasks ripples through to other tasks based on historical data. My contribution to this strand of literature is that I *structurally estimate* this propagation matrix, which depends on the primitives of the model environment, worker characteristics, and comparative advantage vectors. A reduced-form approach is not feasible for the current AI shock because these technologies are in the early adoption phase, lacking sufficient historical data for estimation. A structural approach overcomes this limitation, providing the flexibility to quantify labor market responses to varying degrees of AI integration based on estimated structural parameters rather than past trends.

In parallel work, Smeets, Tian, and Traiberman (2025) study the AI shock using Danish data. Like this study, they account for worker reallocation using a dynamic discrete choice model to uncover comparative advantages. However, they find that lower-income workers are worse off in absolute terms, whereas I find that all groups experience wage gains, though to varying degrees. This divergence in distributional outcomes is likely driven by distinct modeling assumptions regarding the production structure. Specifically, Smeets, Tian, and Traiberman (2025) assume strong complementarity between sectors, following Atalay (2017). In a framework with strong complementarity, automation can lead to labor demand and wage contractions in specific sectors even under productivity growth. In contrast, I model AI as a strictly productivity-enhancing shock within a task-based framework, isolating the supply-side reallocation incentives in a setting where the marginal product of labor increases.

The rest of the paper is organized as follows. Section 2 presents the general equilibrium framework, detailing the task-based production hierarchy and the dynamic occupational choice problem. Section 3 describes the structural estimation of the wage switching cost

parameters, and focuses on the identification of latent worker types. Section 4 outlines the administrative German panel data, the mapping procedure between German and US occupational codes, and the construction of the AI automation scores. Finally, Section 5 presents the estimation results and analyzes the counterfactual steady-state equilibrium under the AI productivity shock.

2 Theory

2.1 Production

The economy is populated by human workers[3] normalized to unity. Time is discrete. Every period, each worker chooses among O different occupations to work. Each occupation consists of a series of tasks, where tasks can either be performed by human workers or the AI technology.

AI technology is perfect substitute for human labor, and it has zero rental cost. AI technology is infinitely more productive than human workers in the tasks it can perform. Therefore, human workers create a bottleneck in the production in the sense that the AI technology can be scaled infinitely whereas the human worker output is costly to scale. Human worker output is the sole determinant of the output of the occupations, whereas the tasks that are performed by the AI technology has no effect on output.

To solidify the idea, consider the occupation of *translator*. Suppose the AI technology performs the translation and the only task to be performed by human workers is to review the translation. Then, the translator output is entirely determined by how fast a human translator can review the translation of the AI technology[4].

The production technology for occupation o at time t is denoted by Y_{ot} and is assumed to have the following function form.

$$Y_{ot} = M_o \left[\sum_{\tau \in \tau_o \setminus \tau_A} Y_{\tau t}^{\frac{\theta-1}{\theta}} \right]^{\frac{\theta}{\theta-1}} \tag{1}$$

where τ_o denotes the set of tasks associated with occupation o and τ_A denotes the set of tasks that are automated. $M_o > 1$ is a productivity multiplier due to time reallocation of human workers from the automated tasks to the task they are performing. Denoting the

[3]Throughout the paper I use *workers* and *human workers* interchangeably, as well as *AI*, *AI technologies* and *automation technology*.

[4]Another example would be a researcher doing a literature review. 50 years ago this task would involve going to a library and skimming through journals to find relevant studies. Now, this task only involves the *skimming through the literature* and not the *going to the library* part for most cases. Hence, the researcher should be able to allocate the time from commuting to the library to searching through the internet, and the output of this task is solely determined by how productive the researcher is in searching through the web for relevant studies.

share of automated tasks in occupation o by m_o, the formula productivity multiplier is as follows.

$$M_o = \frac{1}{1 - m_o} \tag{2}$$

Production in task τ is solely determined by human workers' productivity.

$$Y_{\tau t} = \sum_{n \in L_{ot}} z_{not} \tag{3}$$

where L_{ot} is the set of workers in occupation o at time t and z_not denotes the productivity of worker n in occupation o at time t.

Finally, the consumption basket of the households is a CES combination of the consumption of the individual occupation outputs.

$$c_{nt} = \left(\sum_{o=1}^{O} \mu_o^{\frac{1}{\rho}} c_{ont}^{\frac{\rho-1}{\rho}} \right)^{\frac{\rho}{\rho-1}} \tag{4}$$

Here, μ_o denotes the occupation demand shifter for the occupation o output. Denoting the time t aggregate price level as P_t and the occupation o price level as P_{ot}, and along with the market clearing conditions, optimal consumption dictates

$$Y_{ot} = \mu_o Y_t \left(\frac{P_{ot}}{P_t} \right)^{-\rho} \tag{5}$$

Every occupation is a perfectly competitive market, and the wage of a worker is equal to the marginal productivity of the worker times the price of the occupation output.

$$w_{nt} = M_o \times p_{o_{nt}} \times z_{no_{nt}t} \tag{6}$$

where o_{nt} is the occupation choice of worker n at time t.

2.2 Labor Supply

Having set up the labor demand side, this section provides details about the labor supply. Each period, workers face a problem where they have to choose between staying in their current occupation or switching to another occupation. They are subject to switching costs upon switching occupations, and also switching cost shocks. There is no savings, that is, everyone in the economy is hand-to-mouth consumers. Workers maximize their expected lifetime utility given as

$$\mathbb{E}_0 \sum_{t=0}^{T} \beta^t u(c_{nt}) \tag{7}$$

The flow utility of worker n who is working in occupation o_{nt} at time t is as follows.

$$u(c_{nt}) = w_{nt} + S_n(\xi_{nt}, o_{nt} | o_{nt-1}, \omega_{nt}) \tag{8}$$

where $S_n(\cdot)$ denotes the cost associated with switching from occupation o_{nt-1} to o_{nt}. ω_{nt} denotes the worker characteristics that the switching costs depend on. For ease of expression, switching cost function can be separated into two terms as follows.

$$S_n(\xi_{nt}, o_{nt}|o_{nt-1}) = s_n(o_{nt}|o_{nt-1}, \omega_{nt}) + \xi_{o_{nt}o_{nt-1}nt} \tag{9}$$

so that the deterministic part $s_n(o_{nt}|o_{nt-1}, \omega_{nt})$ and the stochastic part $\xi_{o_{nt}o_{nt-1}nt}$ are separated. For convenience, replace o_{nt} with o' and o_{nt-1} with o. Under certain conditions, worker n's problem can be written as a Bellman equation as follows.

$$V_t(o', h_t, H_t) = \max_{o''} \{\mathbb{E}_t w_{no't} + s_n(o'|o, \omega_{nt}) + \xi_{o'ont} + \beta \mathbb{E}_t V_{t+1}(o'', h_{t+1}, H_{t+1})\} \tag{10}$$

where h_t is the set of individual state and H_t is the set of aggregate state variables. $v_t(\cdot)$ is the value function associated with choosing a specific occupation at time t whereas $V_{t+1}(\cdot)$ is the value function associated with optimal occupation choice from time $t+1$ and on. There is an expectation term on w, since I assume the wage shocks are unobserved before making the occupation choice. Switching cost shocks, on the other hand, are observed before the occupation choice. I also assume that the switching cost shocks follow Type I extreme value distribution ($\xi \sim F(0, \gamma)$) with scale parameter equal to γ and location parameter equal to 0. This yields the following recursive formulation (Rust (1987)).

$$\begin{aligned} v_t(o', h_t, H_t, \omega_{nt}) =& \mathbb{E}_t w_{no't} + s_n(o'|o, \omega_{nt}) + \xi_{o'ont} \\ &+ \gamma \int_{\xi} \log \sum_{o'} \exp\left(\frac{\beta}{\gamma} v_{t+1}(o', h_{t+1}, H_{t+1}, \xi)\right) dF(\xi) + \beta\gamma c^e \end{aligned} \tag{11}$$

where c^e is the Euler–Mascheroni constant. Time $t+1$ unconditional value function ($V_{t+1}(\cdot)$) can be manipulated to define a relationship between the value functions and the transition probabilities.

$$\begin{aligned} \mathbb{E}_t V_{t+1}(o', h_{t+1}, H_{t+1}, \omega_{nt+1}) =& \mathbb{E}_t \big[v_{t+1}(o', h_{t+1}, H_{t+1}, \omega_{nt+1}, \xi) \\ &- \gamma \log \pi\left(o'|o', h_{t+1}, H_{t+1}, \omega_{nt+1}\right) \big] + \gamma c^e \end{aligned} \tag{12}$$

This representation will be critical in estimation of the switching costs as it relates the value function to the conditional choice probabilities. Given two different workers who end up at the same occupation with the same individual states, the lifetime value differential can be reduced to utility differentials between the two workers. Utility differentials are linear functions of expected wages and switching costs. Conditional choice probabilities and the expected wages can be recovered from the data. This allows me to estimate switching costs via a regression. More details on the mathematical identity between the choice (transition) probabilities, expected wages and the switching costs can be found in Section 3.

2.3 State Variables

h_t represents the individual state variables, age, schooling category and the unobserved productivity parameters for the worker. In practice, estimating productivity parameters for

each worker is not feasible because (i) not all of the workers have a work history across all the occupations and (ii) even if they did, they need to have worked at least two periods for each occupation for identification of the productivity parameters. Instead, I assume that each worker belongs to one of the finite types $i \in \{1, \ldots, I\}$. This assumption resolves both problems mentioned before because all types are going to exhibit work histories across all the occupations. Due to the nature of the estimation procedure which will be explained in Section 3, all workers have a strictly positive probability of being any type. Hence, this rules out the concern where very little number of workers belonging to a type and causing identification problems. H_t, on the other hand, represents the only aggregate state variable which are the occupational output prices. Workers can perfectly forecast the individual state variables, age and the time-invariant type, since they are non-stochastic. Imposing a forecasting method for the aggregate state (occupational prices) is not necessary for the estimation purposes, since I am following Arcidiacono and Miller (2011) which allows me to work with empirical transition rates.

Wage of worker n is equal to worker n's productivity in occupation o multiplied by the price of the occupation worker n employed at.

$$w_{nt} = p_{o_{nt}} \times z_{no_{nt}t}$$

where z_{not} is the productivity of worker n in occupation o at time t. Productivity of a worker is, in logarithmic form, a linear function of their individual state variables.

$$\log z_{no_{nt}t} = \beta_1 \times \text{Age}_{nt} + \beta_2 \text{Age}_{nt}^2 + \beta_3 \text{Schooling} + AA_{i(n)o_{nt}} + \sigma_o \varepsilon_{nt} \tag{13}$$

where AA_{io} denotes the absolute advantage of worker type i in occupation o. However, absolute advantage parameters cannot be identified because there is a perfect collinearity between them and the occupational prices (p_{ot}). To see this, consider two economies where every worker is twice productive in the first economy compared to other, whereas the occupational prices are equal to the half of those in the second economy. Then, the average wages would be equal in both economies. For this reason, productivity parameters can be estimated only up to a difference from a benchmark type, which is essentially the comparative advantage. Hence, I estimate this equation setting type 1 as the benchmark type. ε_{nt} is the idiosyncratic wage shock which is unobserved before making the occupation choice. It is assumed to be independent of all state variables, individual or aggregate. σ_o regulates the standard deviation of the wage shocks, which depends on the occupation.

2.4 Equilibrium

This section lays out the equilibrium conditions for the post-AI equilibrium economy. For the pre-AI automation economy the exact conditions apply except $\tau_A = \emptyset$ or $M_o = 1 \, \forall o \in \{1, \ldots, O\}$.

Workers maximize their lifetime utility by choosing an occupation in a forward looking manner. The forward looking behavior is both due to the switching costs and also the

evolving occupational prices. Because the switches are costly, a worker might wait until it is the *right time* to switch occupations, that is, until they are hit with a favorable switching cost shock. I do not consider any functional form for forecasting the occupational prices since the estimation procedure does not require so.

$$\max_{c_{nt},o_{nt}} \mathbb{E}_0 \sum_{t=0}^{T} u(c_{nt}) \tag{14}$$

$$\text{s. to} \quad c_{nt} = w_{nt} + S_n(\xi_{nt}, o_{nt}|o_{nt-1}) \tag{15}$$

Each individual must be working in the occupation that offers the highest expected discounted lifetime utility. Mathematically

$$\mathbb{E}_t V_t^n(o) > \mathbb{E}_t V_t^n(o') \quad \forall o' \in \{1, \ldots, O\} \tag{16}$$

where $V_t^n(o)$ denotes the discounted lifetime utility given the occupation choice o for worker n at time t.

Firms combine task outputs ($Y_{\tau t}$) and produce occupational outputs (Y_{ot}). Each occupation is a perfectly competitive market. Individual behavior of the firms are irrelevant since I can ignore them and work on occupation level variables to establish the equilibrium.

There is no rental cost for the AI technology. With free entry condition, all firms make 0 profit. Therefore the marginal cost must be equal to the output price for all firms in all occupations.

$$p_{ot} = \frac{1}{M_o} \frac{w_{nt}}{z_{not}} \quad \forall\, n \in L_{ot} \tag{17}$$

In the equilibrium, there may be more than one type of labor working within an occupation, which may result with more than one wage per firm.

Market clearing condition for the labor market is

$$\sum_{o=1}^{O} L_{ot} = 1, \quad \forall t \tag{18}$$

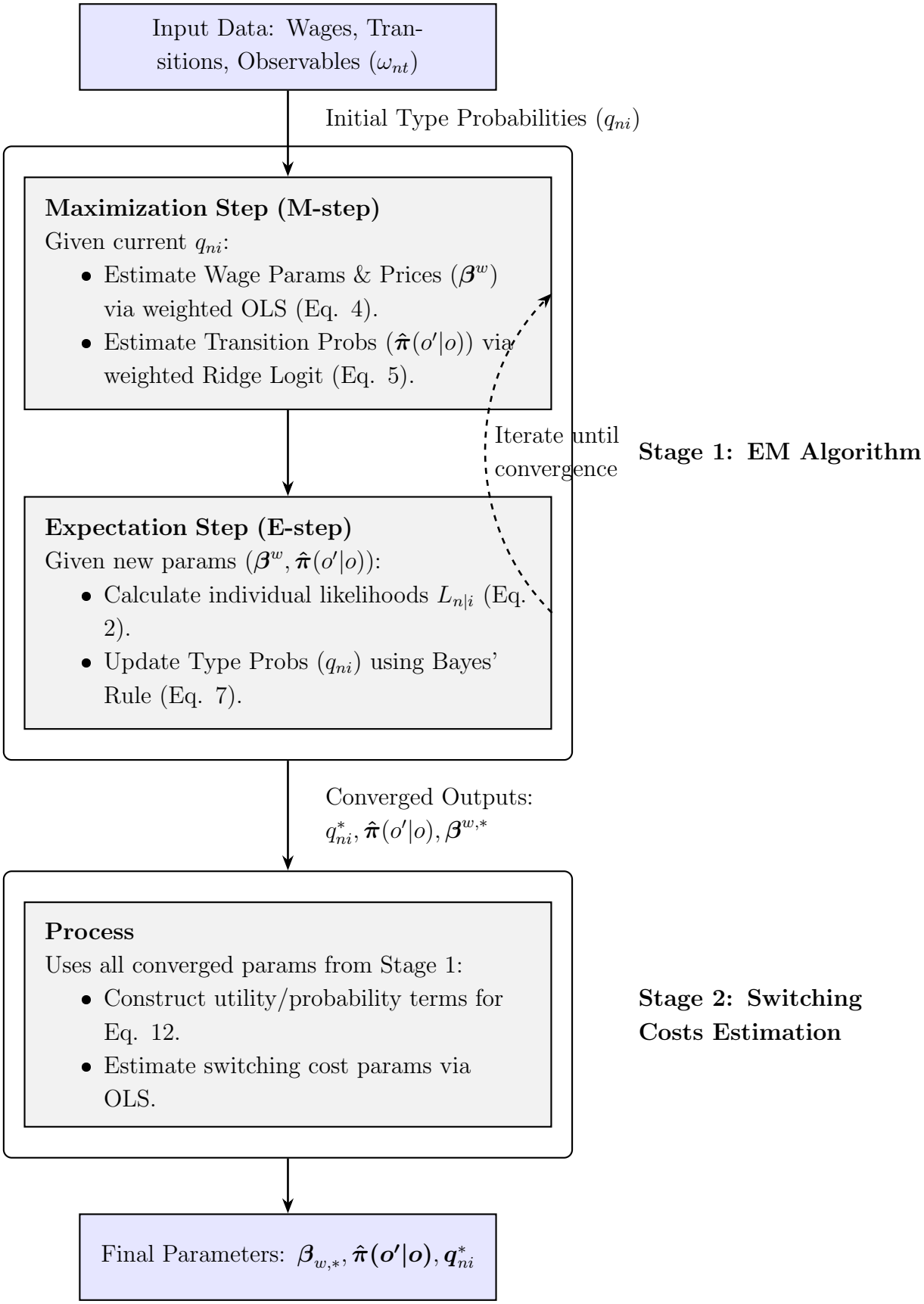
and for the goods market

$$C_{ot} = Y_{ot}, \quad \forall o, t \tag{19}$$

3 Estimation

Without the unobserved types, the estimation procedure would be running two sets of regressions to estimate the income regression and switching cost parameters that maximize the likelihood. Due to unknown worker type probabilities, one must also estimate the vector of type probabilities for each worker. Type probabilities are estimated such that workers'

Figure 1: Visualized Estimation Procedure



history is aligned with being that type. For example, suppose the wage regression for the type 1 workers indicate a positive comparative advantage for a particular occupation (call it occupation 1) and a negative comparative advantage for another occupation occupation 2). Consider a worker who earns more than average controlling for their observables in occupation 1 (see Figure 2). Suppose this worker also earn less than average in occupation 2, controlling for the observables. Then, this worker is likely to have a positive comparative advantage in occupation 1 and a negative comparative advantage in occupation 2. It follows that the worker must be attached a relatively high probability of being type 1 compared to the average type 1 probability of population.

There is no closed form solution for this type of models, and they are estimated via the Expectation-Maximization (EM) algorithm. The idea behind the EM algorithm is maximizing the total likelihood in two steps. First step involves maximizing the log-likelihood given a type probability vector for each worker. Next step involves updating the type probabilities via Bayesian update[5]. These two steps are repeated until the likelihood converges. I perform the EM algorithm for the first stage to calculate the wage regression parameters, occupational prices, type probabilities and transition matrices.

Switching costs can be recovered having obtained the estimates from the first stage. I provide additional details about the second stage of the regression where the switching costs are estimated in Section 3.2.

First, let us define the object that is to be maximized. The log-likelihood for a single observation conditional on the worker being type i as follows

$$L_{nt|i} = f(w_{no_{nt}t}|\omega_{nt}, \text{type} = i) \times \pi(o_{nt}|o_{nt-1}, \omega_{nt}, \text{type} = i) \tag{20}$$

where $f(\cdot)$ is the Gaussian distribution and $\pi(\cdot)$ denotes the transition probability from last period's occupation to this period's. ω_{nt} denotes the vector of observables for worker n at time t, that are age and schooling. Likelihood contribution due to a worker is then

$$L_{n|i} = \prod_{t=1}^{T} f(w_{no_{nt}t}|\omega_{nt}, \text{type} = i) \times \pi(o_{nt}|o_{nt-1}, \omega_{nt}, \text{type} = i) \tag{21}$$

Unconditional likelihood is the integral (in this case, the weighted sum) of the individual likelihood contributions with respect to the type probabilities. The total log-likelihood is therefore

$$L = \prod_{n=1}^{N} \prod_{i=1}^{I} q_{ni} \prod_{t=1}^{T} [f(w_{no_{nt}t}|\omega_{nt}, \text{type} = i) \times \pi(o_{nt}|o_{nt-1}, \omega_{nt}, \text{type} = i)] \tag{22}$$

[5]This is in idea similar to k-means clustering. However, k-means clustering assigns binary scores for clusters, whereas the EM algorithm assigns probabilities. Furthermore, EM is more flexible in the sense that I can define the norm as the likelihood defined by the data generating processes that I introduce, whereas with the k-means I would have to use the L2-norm which would not be suitable for defining a model-consistent distance.

where q_{ni} denotes the probability attached to worker n being type i. EM algorithm allows estimating q_{ni} and the conditional likelihood in two different stages instead of tackling a very high dimensional problem. However, $\pi(\cdot)$ depends on both the first and the second stage parameters, as transition probabilities depend on the comparative advantage vectors, as well as the switching cost parameters and shocks. Therefore, maximization of this likelihood involves finding a fixed point for the the contemporaneously estimated parameters for the transition probabilities and wage regression as well as the switching cost parameters estimated in sequence[6]. Doing this at every iteration of the EM algorithm is computationally infeasible. Furthermore, I would need to impose additional structure on the model by defining forecasting rules for the workers for the aggregate states[7].

To overcome this problem, I follow (Arcidiacono and Miller, 2011) for the estimation, which allows me to treat the transition probabilities as something to be empirically estimated from the data, instead of calculating as part of the fixed point or the recursive problem. Following sections describe in detail the estimation process, where it is also visualized in Figure 1.

3.1 First Stage

The EM algorithm starts off by initiating type probability vectors for all workers. The only hard rule for the initial type probabilities is that they must not be perfectly uniform. If the initial type probabilities are left very close to uniform then the maximization step yields the same parameters for every type and thus the EM algorithm cannot converge. There has to be some diversity between the type vectors so that the maximization step of the EM algorithm can generate different sets of parameters for each type, and start to converge from there.

I divide the occupations into 4 categories, based on a rough measure of how similar their names are. For each worker, I assign initial type probabilities depending on the share of their occupation history in each occupation category. Even when a worker spent all their career in one occupation group, the resulting initial type probabilities indicate a some nudge towards one type and are not very definitive.

3.1.1 Maximization Step

Maximization step involves estimating a weighted regression for the wage regression, and another weighted regression for generating empirical transition probabilities. Occupational prices are estimated as part of the wage regression. Occupational prices do not depend on worker types, and a the way to ensure that is to estimate the wage regression for all types

[6]Since this particular estimation is a finite dynamic programming problem, one needs to perform a backward recursion to solve for the transition probabilities

[7]In case of rational expectations one would have to find another fixed point for the forecasting rule and the actual realizations for the aggregate state variable.

in a single equation, where the occupational prices are not differentiated with respect to worker types.

I use 34 occupations and 4 types, which requires estimating 34×3 comparative advantage parameters. While technically feasible, some occupations are *similar* in the skills they require[8]. Instead of estimating 34 productivity variables for each type, I reduce the dimension in the occupation space by generating a lower-dimensional skills vector[9]. This idea is similar to generating a distance measure between the tasks. If two occupations are similar, then workers who have comparative advantage in occupation are likely to have comparative advantage in the other occupation as well.

To do so, I use quantified occupation characteristics from O*NET database, such as "importance" of mathematics, reading comprehension, negotiation, etc. Then I use the first 8 principal components of this very high-dimensional information[10]. Thanks to this dimension reduction I have to estimate 3×8 comparative advantage parameters, instead of 3×34 comparative advantage parameters.

$$\log w_{ino_{nt}t} = p_{ot} + \beta_1 \times \text{Age}_{nt} + \beta_2 \text{Age}_{nt}^2 + \beta_3 \text{Schooling} + \mathbb{1}\{i \neq 1\}\Gamma_o \boldsymbol{\beta}_{CA}^i + \sigma_o \varepsilon_{nt} \tag{23}$$

where Γ_o represents the skill shifters (first 8 principal components) and $\boldsymbol{\beta}_{CA}^i$ represents the skill vector corresponding to the first 8 principal components. This regression is estiamted via weighted OLS, where the estimation matrix is stacked for each type, and q_{ni} enter as observation weights.

For estimating the transition probabilities, Arcidiacono and Miller, 2011 suggests using the empirical distribution of the transitions. Specifically, a bin estimator as follows

$$\pi(h_2, o_2 | h_1, o_1) = \frac{\sum_{n=1}^{N} q_{ni} \mathbb{1}(h_{nt} = h_2, o_{nt} = o_2, h_{nt-1} = h_1, o_{nt-1} = o_1)}{\sum_{n=1}^{N} q_{ni} \mathbb{1}(hh_{nt-1} = h_1, o_{nt-1} = o_1} \tag{24}$$

Where h are individual state variables, age and schooling. In practice, this bin estimator is not a reliable measure because partitioning the data based on individual state variables and occupation result with very few observations for some (age, schooling and occupation) triplets. Traiberman (2019) faces the same problem and uses a linear probability model to approximate the bin estimator while Ransom (2022) uses a logit estimator. I find that both approaches lead to numerical instabilities in my case. With an LPM model, numerical instability arises because the predicted probabilities are not bound between 0 and 1. With the logit model, some estimations do not converge given very few observations for some transitions. Therefore, I rely on an L2-regularized (ridge) logit to keep the parameter estimates from taking unreasonable values when there are only a few observations for a transition[11].

[8] For example "Occupations in plastic-making and -processing, and wood-working and processing" and "Occupations in production and processing of raw materials, glass- and ceramic-making and processing".

[9] **smeetsFieldChoiceSkill** uses the same dimension reduction approach.

[10] First 8 principal components explain 82.5% of the variance in the entire information matrix. More information on the construction of the principal components can be found in Section C.2.

[11] It is also possible to assign some default parameters when there are very few observations for any

3.1.2 Expectation Step

In this step, type probabilities are updated by Bayesian updating. The intuition behind the updating procedure is as follows. Consider an EM estimation with two types. There are two set of parameters estimated for each type. If a worker's likelihood contribution can be better explained with a particular set of parameters belonging to a certain type, than the probability that the person is that type increases. Formally, the updating formula is as follows.

$$q_{ni}^{(m+1)} = \frac{L_{n|i} q^{(m)}(i|\omega_{n1}^{obs})}{\sum_{i'} L_{n|i'} q^{(m)}(i'|\omega_{n1}^{obs})} \tag{25}$$

where $q(i|\omega^{obs})$ relates the distribution of being type i to the observables at time $t = 1$ (Arcidiacono and Miller (2011)). This allows me to take into account that the initial individual state variables, age and schooling, might be reflective of the unobserved type. $q(i, \omega^{obs})$ is updated at every iteration of the EM algorithm since the type probabilities change at every iteration.

3.1.3 Identification of Worker Types

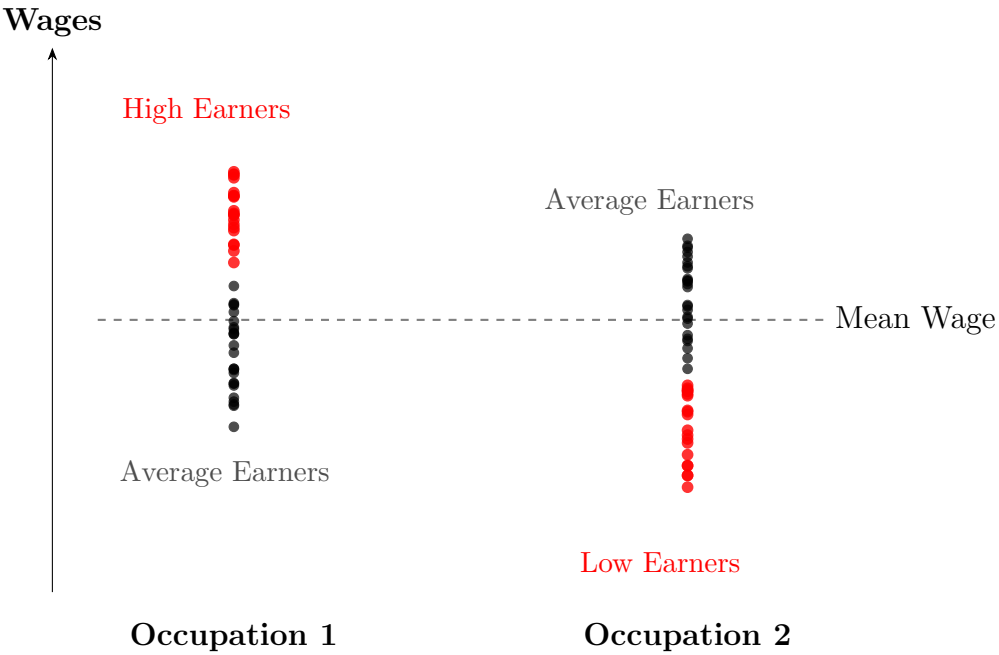
Identification of types relies on two observables, wages and transitions. Consider a group of workers who earn more than average in a certain occupation, after controlling for their observable characteristics such as age and schooling. This implies that their unobservable comparative advantage for that occupation is likely to be positive. These group of workers are likely to be the same type if they exhibit similar wage patterns in other occupations as well (see Figure 2).

Transitions also determine the worker types. Consider two types of workers, high and low skill. High skill occupations are mostly populated by high skill workers, whereas low skill jobs present a bit more diversity in terms of skill-mix. Transition histories for these two types of workers will look different in the sense that while the high skill workers can transition between low and high skill occupations, low skill workers can only transition between the low skill occupations (see Figure 3)

Identification of a price of an occupation also relies on the same conditions for the identification of the types working in that occupation. To see when the identification of an occupation price may not be possible, consider a type of workers are only employed in a single occupation. Because this type's comparative advantage cannot be identified due to not being employed in any other occupation, the price of the occupation cannot be identified in the wage equation as well. The estimates for occupation prices and the comparative advantage

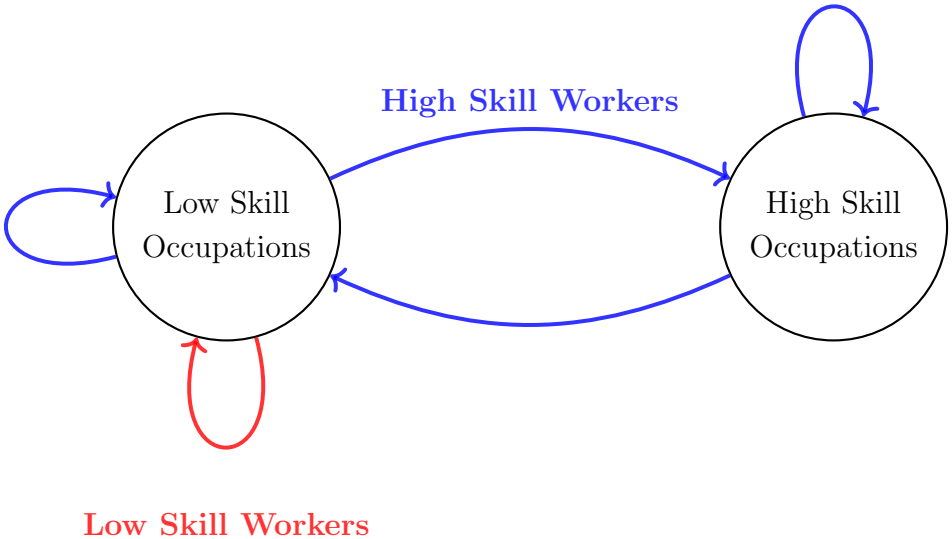
transition. This does not work in practice because the logit estimator may fail to converge with, for example, 10 transitions whereas it may converge with a single transition only. As such, the cases of failed convergence are not characterized by an observation threshold. Hence, assigning default large negative values for the failed cases creates a negative bias for the transitions with *enough* number of observations yet where the logit estimator does not converge.

Figure 2: Worker Types and Comparative Advantage (CA).



Note: A group of workers (red dots) who earn more than average in Occupation 1 also earn less than average in Occupation 2, suggesting they belong to a specific unobserved *type*.

Figure 3: Worker Types and Transitions



Note: High skill workers (blue) can transition between both high and low skill occupations. Low skill workers (red) are restricted to transitions within low skill occupations. The transition histories are likely to separate this two groups of workers into different types.

parameters are more robust when workers (especially of different types) transition between different occupations.

3.2 Second Stage

Estimation of the switching costs relies on exploiting two occupation transition paths that start at the same occupation (o) and ends at another occupation (both at o'). At the end of this transition paths, continuation values are the same for two workers with the same individual states. This allows writing down the transition probabilities in Eq. B15 in terms of flow utilities and switching costs (for the full derivation of the conditional choice probabilities see Section B.3).

$$\log\left(\frac{\pi_t(o'|o,h_t,H_t,\omega_{nt})}{\pi_t(o|o,h_t,H_t,\omega_{nt})}\right)+\beta\log\left(\frac{\pi_{t+1}(o'|o',h_{t+1},H_{t+1},\omega_{nt+1})}{\pi_{t+1}(o'|o,h_{t+1},H_{t+1},\omega_{nt+1})}\right)$$
$$=\frac{1}{\gamma}\mathbb{E}_t\left[w_{no't}-w_{not}\right]+\frac{1-\beta}{\gamma}s_n(o'|o,\omega_{nt})+\frac{1}{\gamma}\left[\xi_{o'ont}-\xi_{o'ont+1}\right] \tag{26}$$

Transition probabilities and expected wage differentials are recovered in the first stage. Following the first stage, I estimate this equation via OLS where the discount factor $\beta = 0.96$.

Estimation of the switching costs reliably requires a sufficient amount of transition between each occupation pair. This is not the case for each occupation pair, and to overcome that, I follow Ransom (2022) and Traiberman (2019), and define a measure of distance between two occupations, and assume that the switching cost is a linear function of distance along with a fixed cost for the origin equation. This transformation can be shown as follows.

$$s_n(o'|o,\omega_{nt})=\alpha_0^o+\alpha_1 d(o,o') \tag{27}$$

Ransom (2022) estimates switching costs between locations, which has a natural measure for *distance*. Traiberman (2019) uses the principal components to generate Mahalanobis measure, which scales down the principal components as well as taking into account the correlation between the principal components. I use the principal component vectors associated with each occupation and then measure the Euclidean distance between the associated vectors with every occupation. Euclidean distance does not scale the variance of the principal components. I find this more informative since principal components are ordered in terms of explaining the total variance hence the information content. Accordingly, a principal component of higher variance will have a larger *weight* with the Euclidean distance measure.

4 Data

4.1 Labor Panel

I use an administrative German data, Sample of Integrated Labour Market Biographies, which contains a 2 percent random sample from all individuals in Germany (Graf et al. (2023)). The data covers the period between 1975 and 2021, and has information on the ID, age, employment status, occupation, schooling, gender and income. Some information is anonymized further or provided in restricted detail. Wages are rounded to the nearest integer. For schooling, I define two categories, university education and all the other lower degrees since the data doesn't provide a fine grained information on the schooling variable.

I use the data from 1998 and onwards[12]. The occupation classification for the data is KldB 2010. While the data reported from after November 2011 uses KldB2010, the data reported before uses KldB 1998 classification, which are converted to KldB 2010 classification by the data provider.

I use O*NET database to get the task descriptions and the information on attributes such as skills and knowledge relevant for the tasks. O*NET database uses SOC 2019 classification, however, there are no direct crosswalks between SOC 2019 and KldB 2010 classifications. To this end, I use ISCO-08 occupation classification and first transcode SOC 2019 occupations to ISCO-08 classification and from there to KldB 2010 classification.

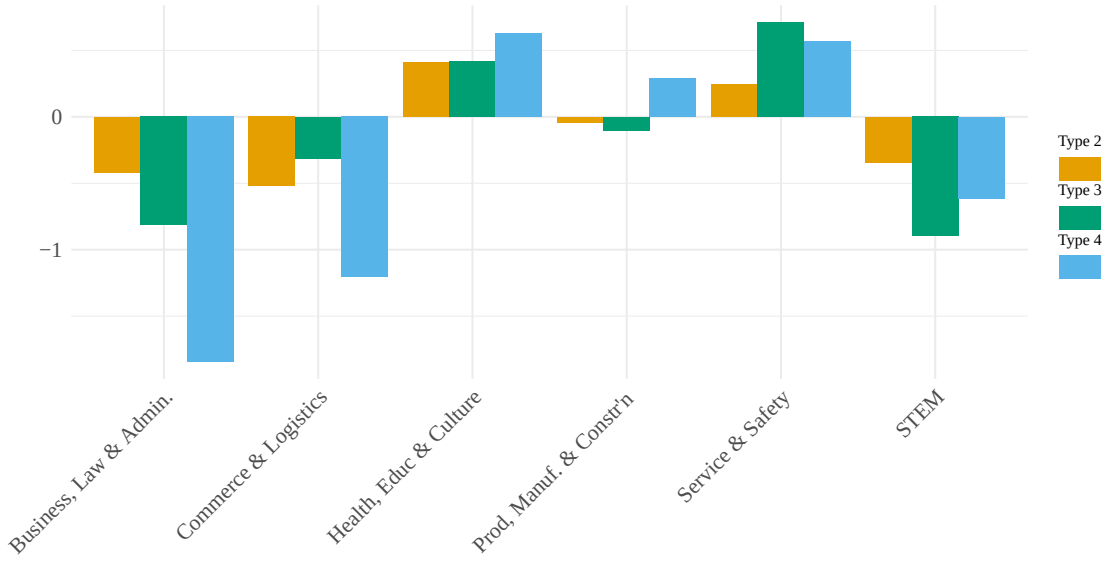
For generating the principal components, I follow Traiberman (2019) use "Knowledge", "Skills","Work Activities" and "Abilities" attributes from O*NET database. For each attribute, there are two measures, "Importance" and "Level". I follow Firpo, Fortin, and Lemieux (2011) and assign 2/3 and 1/3 geometric weights, respectively, to generate a single measure. This leaves me with 160 attributes for each occupation, for which I apply principal components dimension reduction to generate 8 new attributes. In Appendix section C, I detail the most important attributes for the first 8 principal components.

[12] After 1998 the employers started reporting earnings below some threshold. While the effect this change on the estimation is probably negligible, another reason I opted to start from 1998 is mainly due to memory limitations on the server that I am running the estimation. I estimate the wage regression for 4 types at the same time to estimate a single set of prices, and this requires stacking 4 wage regression matrix. Hence, the memory that the estimation requires roughly quadrupled. In addition to this, some data operations cannot be performed in-place, resulting in additional large matrices being created and increasing the RAM requirement. While there are some methods in R to bypass RAM limitations by using an SQL database or other methods that work over the hard disk instead of RAM, almost none of the options are usable due to the restrictions on the machine that I run the estimation on. For this reason, I further limit the sample size by randomly selecting some worker IDs after the data cleaning process. Even then, I have to rely on memory-efficient estimation processes such as incremental QR decomposition and using sparse matrices.

4.2 AI Exposure Scores

To compute the post-AI general equilibrium I need binary exposure scores for each task in every occupation. To achieve this, I follow Eloundou et al. (2023) and Eisfeldt et al. (2023) and generate the scores using an LLM. Specifically, I provide the LLM a task description and provide clear instructions to determine whether it can be automated by LLM technologies or not. I do this for every task description under all occupations in the O*NET database, and calculate the exposure scores by taking simple average across task automation scores[13,14,15]. Then, using relevant crosswalks, I aggregate the scores generated by the LLM to KldB-2010 2-digit level occupations. Exposure scores for the 2-digit KldB-2010 occupations are in Table A2.

Figure 4: Comparative Advantage Across Occupation Groups



5 Results

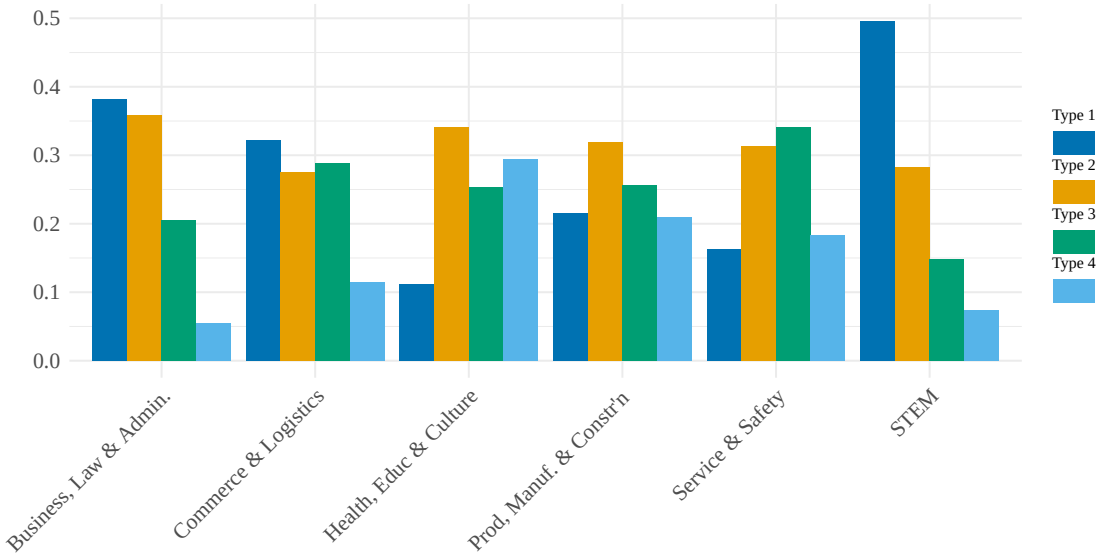
First I present a discussion of the estimation results; wage parameters, including the comparative advantage vectors, type distributions and switching costs. These are not equilibrium

[13]I use an initial "system" prompt to assign the LLM a clear role in identifying whether a task can be automated or not. The system prompt follows Eisfeldt et al. (2023) with some changes that takes into account the advancements in the LLMs since then.

[14]I experimented with OpenAI GPT-5, GPT-5 mini, GPT-5 nano, Google Gemini 2.5 Pro, 2.5 Flash, 2.5 Flash Lite and Claude Sonnet 4.0. Based on the input/output tokens generated for roughly 20,000 tasks, state of the art models (GPT-5, Gemini 2.5 Pro and Sonnet 4.0 are very slow and expensive to work with. Among the more speed and budget oriented model, I find that 2.5 Flash returns much more reasonable answers based on the reasoning output.

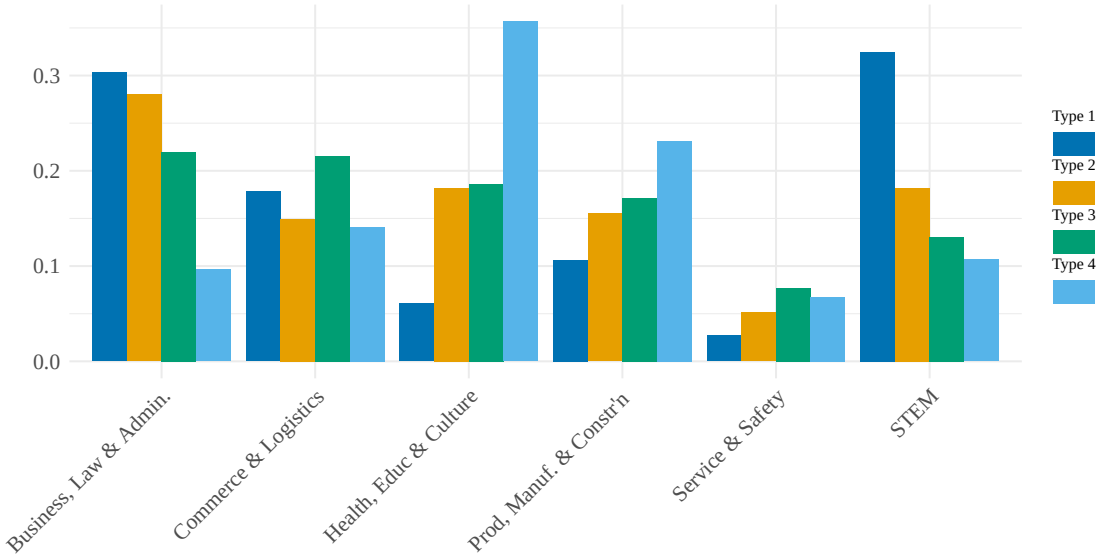
[15]Exact prompt given to Gemini 2.5 Flash model can be seen in Appendix Section D.

Figure 5: Share of Types Across Occupation Groups



Note: Total shares of all types for each occupation group is 1.

Figure 6: Distribution of Types Across Occupations



Note: Shares of all occupations for each type sum up to 1.

dependent objects, instead they are inherent attributes for the workers and the environment. Then I move onto the post-AI equilibrium where I discuss the wage changes and the employment shifts between occupations.

Figure 4 shows the comparative advantage of the three types in all occupation groups[16] in

[16]For the list of occupations and the corresponding occupation groups, please see Table A3.

comparison to the first type. Hence, type 1 workers has 0 comparative advantage across all occupations, but they might have positive comparative advantage against others in an occupation if others have negative comparative advantage in that occupation.

Type 1 workers has a comparative advantage in Business, Law & Administrative occupations along with STEM and Commerce & Logistics occupation groups. Type 2 and type 3 workers seem to be all-rounders whereas type 4 workers have a negative comparative advantage in all occupation groups other than Services & Safety, Health, Education & Culture and Production, Manufacturing & Construction (PMC) occupation groups.

When a worker type has a distinct comparative advantage in an occupation, they would populate the occupation, increase the competition for other worker types and drive them away. Similarly, type 4 workers are driven away from the occupation groups except Health, Education & Culture, Production, Manufacturing & Construction and Services & Safety, and they would face much less competition in these groups.

Figure 5[17] indicates that there is sorting of working types into occupations with respect to their comparative advantage. Half of the STEM population is type 4 workers, and they also make up on average more than 30 percent of the worker population in Business, Law & Administrative and Commerce & Logistics occupation groups. On the other hand, type 2 and type 3 workers do not exhibit a strong sorting pattern given their uniform comparative advantage.

5.1 Post-AI Equilibrium

In this section I present the post-AI equilibrum. To draw a reasonable comparison between the pre and post-AI equilibria, I also calculate the pre-AI equilibrium to get rid of any other important shock that might have occurred during the estimation period. I start from the average of the last three years' estimated occupational prices in both cases, and only adjust M_o parameter that accounts for the AI automation shock. I numerically solve for the post-AI equilibrium, starting from the original equilibrium prices. The exact algorithm for the numerical solution is described in Algorithm B2.

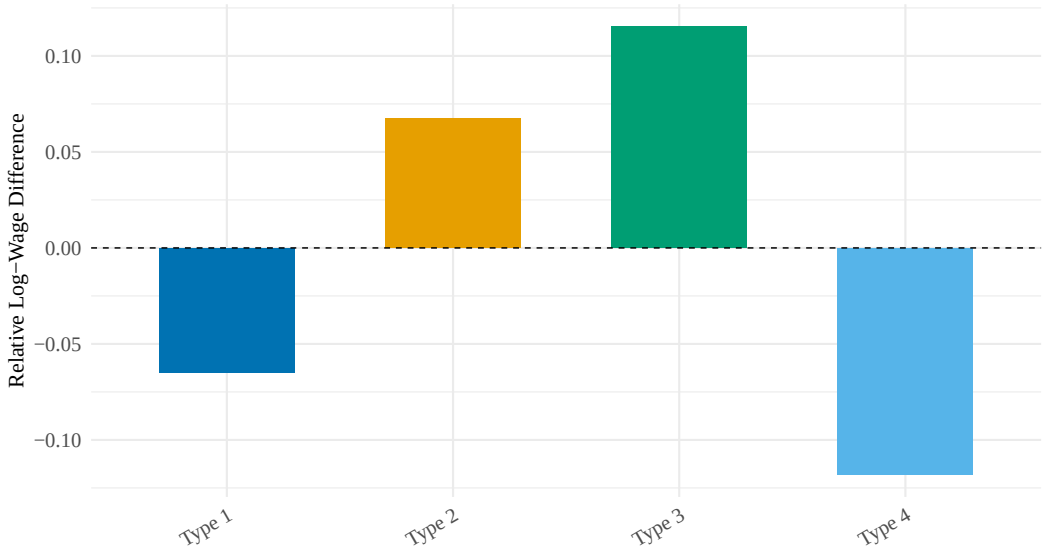
Since AI is introduced as a positive productivity shock, wages increase overall[18]. However,

[17] This figure does not account for the fact that worker types do not have equal masses. Figure 6 shows the same distribution of each worker type normalized for each type.

[18] In calculating the wage statistics by type, and how each types are distributed across occupations, I take weighted averages where the weights are posterior type probabilities for each worker. This results with some types working in some occupations where they don't have any comparative advantage for. For example, consider a group of workers who are type 4 with 80% posterior probability and type 1 with 10% posterior probability. Comparative advantage of these workers then strongly resembles the comparative advantage vector of a type 4 worker, meaning they are likely to be employed in occupations where type 4 has positive comparative advantage. However, taking weighted averages, 10% of this group of workers are type 1 workers and these type 1 workers are employed in the occupations where type 4 workers have a positive comparative advantage.

looking at the relative wages, there are winners and losers. Type 1 workers, who specializes in technical occupations, and type 4 workers, who specializes in services type occupations seem to do worse than the other types who are all-rounders (Figure 7). This indicates that the reallocation ability is significant in terms of reaping the benefits of a sizable fluctuation in prices (see Figure 4).

Figure 7: Relative Log-Wage Differences Across Types



Notes: y-axis indicate the difference between the wages in the post-AI and pre-AI equilibria, relative to the (unweighted) average wage difference across types. The picture with the average weighted by the type shares still look similar since type 1 and type 4 workers have a total of 45% share in the economy.

Figure 8 shows there is a strong positive correlation between the occupational price changes and the share of automated tasks in the post-AI equilibrium[19]. Since the AI shock boosts productivity of the workers, production increases in those occupations thus yielding lower prices. Occupational outputs are substitutes and this prevents prices from falling down too much. Hence, higher wages due to the productivity effect dominates and relatively few workers move from the highly exposed occupations.

Figure 9 illustrates the shift in worker allocation across occupation groups. The most signif-

This interpretation might be misleading in some cases where the two type probabilities are very positively correlated. To make sure this is not the case, I also construct wage and employment distribution by constructing binary types. If a worker have a large posterior probability of being a certain type, then they are assigned probability 1 for that type, and 0 for others. For example, for the 70& threshold, a worker has 100% probability of being type 1 if they have more than 70% posterior probability of being type 1. This ensures that workers represented in the following figures are mostly represented by a single type, especially with the threshold set as 70%. In both cases, however, results do not change in any direction or extent that may affect the interpretation in any way (Figure A1, A2, A4 and A5).

[19]While the average price in the post-AI equilibrium seems to be higher, price of the optimal consumption basket are still the same in both equilibria.

Figure 8: Changes in Occupation Prices vs. Non-Automated Task Shares

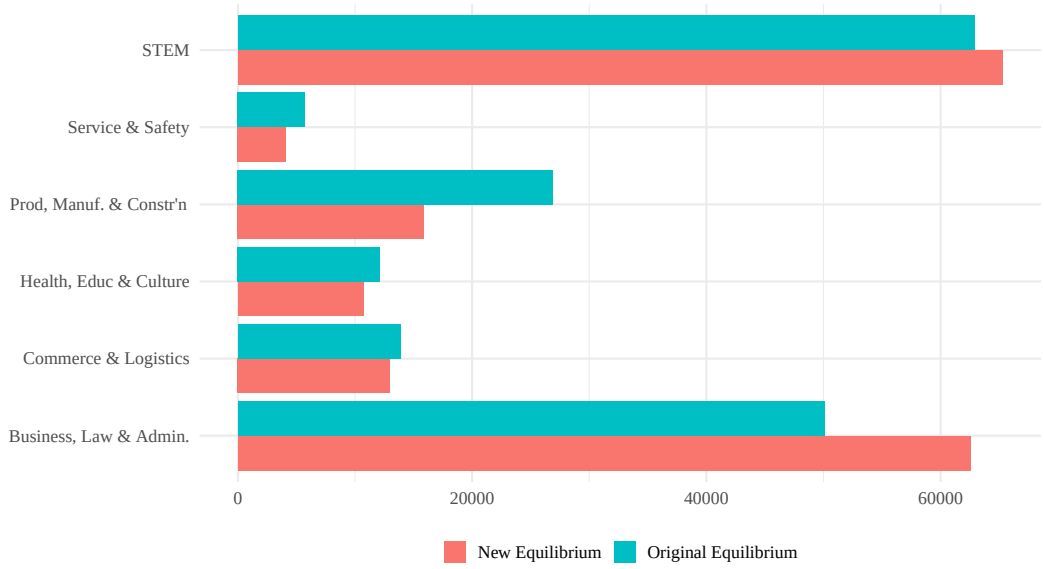
Notes: Left y-axis is for the occupational price differences between the post-AI equilibrium and the original equilibrium. Right y-axis is for the inverse automated task shares.

icant reallocation occurs from Manufacturing & Construction (low exposure) to Business, Law & Administration (high exposure), and this driven almost entirely by Type 1 workers.

This migration is explained by the heterogeneity in comparative advantage. While the AI shock boosts productivity in Business & Law, it also places downward pressure on prices. For most worker types (Types 2–4), the productivity gain is insufficient to offset the falling prices given their steep comparative disadvantage in these fields. Type 1 workers, however, possess a flat (non-negative) comparative advantage profile, making them the only group capable of absorbing this reallocation. Yet, despite this massive inflow, Type 1 workers do not realize significant welfare gains; their uniform (zero) comparative advantage means they merely drift across occupations without capturing a specific productivity premium.

Figure A6 reveals a distinct heterogeneity in the efficiency of reallocation. Type 1 workers exhibit the highest reallocation intensity, yet this mobility reflects *displacement* rather than strategic arbitrage. The occupations originally dominated by Type 1 workers become increasingly competitive due to inflows from *generalist* types, depressing occupational prices to the level where staying is no longer optimal. Consequently, Type 1 workers are forced to switch; they trade away their sharp comparative advantage to access more favorable price conditions in other sectors. In contrast, the *generalist* groups (Type 2 and Type 3) engage in *strategic* switching. Their movement is efficiently targeted toward occupations experiencing high productivity shocks, allowing them to maximize gains with less structural disruption. Finally, Type 4 workers, who specialize in low-skill tasks, exhibit the lowest Wasserstein

Figure 9: Worker Allocation Across the Occupation Groups



Notes: x-axis denote the number of workers.

distance. Their occupations experience neither a significant productivity boost nor an influx of labor competition. Constrained by comparative advantage profiles that do not favor alternative roles, these workers remain anchored to their original occupations.

6 Conclusion

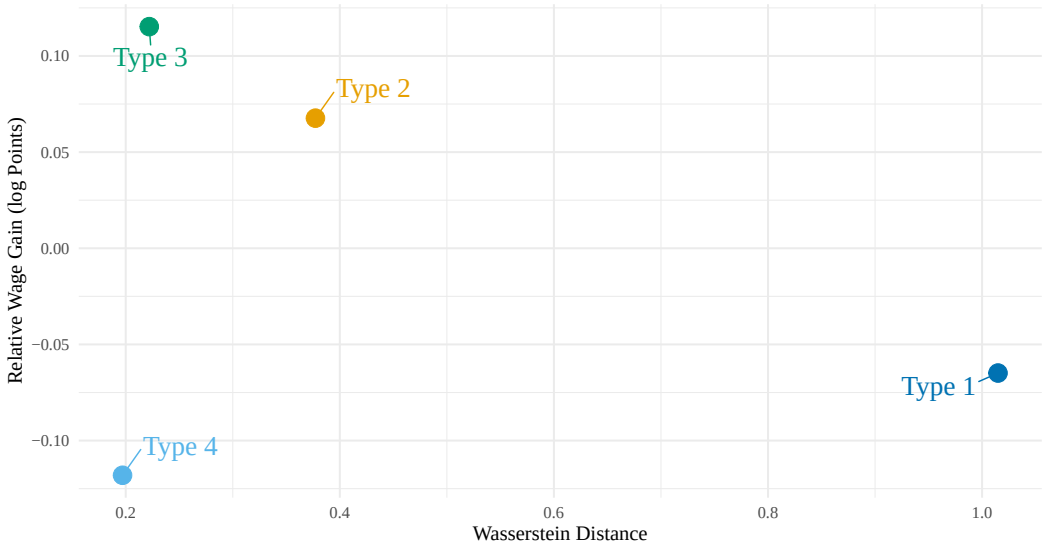
As AI technology becomes increasingly integrated into the workplace, the future of the labor market remains uncertain. To address this uncertainty, many predictions regarding occupational exposure to AI have been developed. However, these predictions offer an incomplete picture, as they fail to account for worker reallocation. Consequently, existing exposure scores are difficult to interpret in isolation. In this study, I build a framework where production can be performed by either human workers or AI technology. To analyze these reallocation dynamics, I estimate a model of occupational choice featuring heterogeneous worker types with varying comparative advantages across occupations.

I estimate 4 worker types, where two of them turn out to be specialists, i.e. they have distinct comparative advantages across the occupation groups, and the other two of them are generalists, i.e. they have uniform comparative advantage across the occupation groups. I find that the comparative advantage and the share of types across occupations exhibit a positive correlation, suggesting a sorting with respect to comparative advantage.

Using the automation scores I generate following Eloundou et al. (2023), I simulate an AI shock. Then I numerically solve for the new equilibrium starting from the prices distorted by the AI shock. I compare the original equilibrium and the post-AI equilibrium.

First, because the productivity gains from AI are asymmetrically distributed across occu-

Figure 10: Displacement vs. Relative Wage Change



Notes: x-axis denote the Wasserstein distance between the pre and post-AI type distributions across occupations.

pations, the optimal response for the labor market involves significant worker reallocation. However, the capacity to capture these gains is strictly governed by the transferability of human capital. Specialists, who derive their pre-shock wages from deep, task-specific comparative advantages, face a steep trade-off. For them, switching occupations entails forfeiting the specific edge that defines their productivity, effectively increasing the opportunity cost of moving. Consequently, they are often unable to chase the productivity boom. Generalists, in contrast, benefit from a "flexibility premium." Because their comparative advantage profile is relatively uniform, they incur minimal productivity losses when switching sectors. This allows them to engage in strategic arbitrage, pivoting to the highest-growth occupations to maximize their welfare gains.

While AI automation will inevitably generate divergent economic outcomes, accurate welfare assessment requires accounting for the margin of worker reallocation. This study demonstrates that a worker's current occupation is an insufficient statistic for predicting their post-shock trajectory. Instead, the ultimate distribution of gains is governed by unobserved comparative advantages and the capacity to move across occupations. Consequently, as AI adoption accelerates, policy frameworks must look beyond static occupational labels and prioritize understanding the labor mobility.

References

Acemoglu, Daron, David Autor, et al. (2022). "Artificial Intelligence and Jobs: Evidence from Online Vacancies". In: *Journal of Labor Economics* 40.S1.

Acemoglu, Daron and Pascual Restrepo (2018). "Modeling Automation". In: *AEA Papers and Proceedings* 108.

— (2022). "Tasks, Automation, and the Rise in U.S. Wage Inequality". In: *Econometrica* 90.5.

Arcidiacono, Peter and Robert A. Miller (2011). "Conditional Choice Probability Estimation of Dynamic Discrete Choice Models With Unobserved Heterogeneity". In: *Econometrica : journal of the Econometric Society*.

Atalay, Enghin (Oct. 2017). "How Important Are Sectoral Shocks?" In: *American Economic Journal: Macroeconomics* 9.4, pp. 254–280.

Bick, Alexander, Adam Blandin, and David J. Deming (2024). *The Rapid Adoption of Generative AI*. National Bureau of Economic Research Working Paper.

Brynjolfsson, Erik, Tom Mitchell, and Daniel Rock (2018). "What Can Machines Learn and What Does It Mean for Occupations and the Economy?" In: *AEA Papers and Proceedings* 108.

Burstein, Ariel, Eduardo Morales, and Jonathan Vogel (2019). "Changes in Between-Group Inequality: Computers, Occupations, and International Trade". In: *American Economic Journal: Macroeconomics* 11.2, pp. 348–400.

Eisfeldt, Andrea L. et al. (2023). *Generative AI and Firm Values*. SSRN Scholarly Paper.

Eloundou, Tyna et al. (2023). *GPTs Are GPTs: An Early Look at the Labor Market Impact Potential of Large Language Models*. Working Paper. arXiv.

Felten, Edward W., Manav Raj, and Robert Seamans (May 2018). "A Method to Link Advances in Artificial Intelligence to Occupational Abilities". In: *AEA Papers and Proceedings* 108, pp. 54–57.

Firpo, Sergio, Nicole M. Fortin, and Thomas Lemieux (2011). "Occupational Tasks and Changes in the Wage Structure". In: *SSRN Electronic Journal*.

Graf, Tobias et al. (2023). *Weakly anonymous Version of the Sample of Integrated Labour Market Biographies (SIAB) – Version 7521 v1Schwach anonymisierte Version der Stichprobe der Integrierten Arbeitsmarktbiografien (SIAB) – Version 7521 v1*.

Handa, Kunal et al. (2025). *Which Economic Tasks Are Performed with AI? Evidence from Millions of Claude Conversations*. Technical Report. arXiv.

Humlum, Anders (2021). *Robot Adoption and Labor Market Dynamics*. Working Paper.

Humlum, Anders and Emilie Vestergaard (2024). *The Adoption of ChatGPT*. University of Chicago, Becker Friedman Institute for Economics Working Paper No. 2024-50.

McElheran, Kristina et al. (2024). "AI Adoption in America: Who, What, and Where". In: *Journal of Economics & Management Strategy*.

Ransom, Tyler (2022). "Labor Market Frictions and Moving Costs of the Employed and Unemployed". In: *Journal of Human Resources* 57.S, S137–S166.

Rust, John (1987). "Optimal Replacement of GMC Bus Engines: An Empirical Model of Harold Zurcher". In: *Econometrica* 55.5, p. 999. ISSN: 00129682.

Smeets, Valerie, Lin Tian, and Sharon Traiberman (2025). *Field Choice, Skill Specificity, and Labor Market Disruptions*. Working Paper.

Traiberman, Sharon (2019). "Occupations and Import Competition: Evidence from Denmark". In: *American Economic Review*.

Trammell, Philip and Anton Korinek (2023). *Economic Growth under Transformative AI*. National Bureau of Economic Research Working Paper. National Bureau of Economic Research.

Webb, Michael (2020). *The Impact of Artificial Intelligence on the Labor Market*. Working Paper.

A Additional Tables & Figures

Table A1: Wage Equation Estimates by Worker Type

	All Types	Type 1	Type 2	Type 3	Type 4
Age	0.0419				
	(0.0001)				
Age^2	-0.0004				
	(0.0000)				
Schooling $\leq$ Secondary	-0.1822				
	(0.0002)				
1^{st} PC		0	-0.0499	-0.2102	-0.1739
			(0.0003)	(0.0003)	(0.0004)
2^{nd} PC		0	-0.1046	-0.0713	0.0269
			(0.0003)	(0.0003)	(0.0004)
3^{rd} PC		0	0.1575	0.5647	0.3885
			(0.0004)	(0.0004)	(0.0005)
4^{th} PC		0	0.1794	0.3952	-0.0704
			(0.0008)	(0.0008)	(0.0010)
5^{th} PC		0	-0.3357	-0.6689	-1.2483
			(0.0008)	(0.0008)	(0.0012)
6^{th} PC		0	0.8912	0.6521	1.5685
			(0.0009)	(0.0009)	(0.0011)
7^{th} PC		0	-0.9100	-0.0996	-1.8305
			(0.0012)	(0.0014)	(0.0017)
8^{th} PC		0	-0.1419	-1.3706	-0.3703
			(0.0011)	(0.0012)	(0.0015)

Notes: Occupation-year fixed effects are provided in Table **??**. Standard errors are reported in parentheses. The "All Types" column shows coefficients common to all worker types. Principal components (PC) are type-specific transformations of occupation characteristics. Type 1 coefficients for principal components are normalized to zero. All coefficients are statistically significant at the 1% level.

Table A2: Exposure Scores by Occupation

Occupation	Score	Occupation	Score
Advertising, marketing, and media design	0.63	Tourism, hotels, and restaurants	0.34
Financial services, accounting, and tax	0.61	Non-medical healthcare and body care	0.34
Computer science and ICT	0.59	Metal production and construction	0.33
Business management and organization	0.59	Raw materials, glass, and ceramic processing	0.31
Humanities, social sciences, and economics	0.56	Medical and health care	0.31
Purchasing, sales, and trading	0.54	Technical building services	0.31
Law and public administration	0.49	Teaching and training	0.30
Construction planning and surveying	0.49	Cleaning services	0.29
Technical R&D, construction, and production	0.44	Mechatronics and electrical engineering	0.27
Safety, security, and surveillance	0.42	Plastic, wood, and wood processing	0.26
Math, biology, chemistry, and physics	0.42	Gardening and floristry	0.25
Traffic and logistics	0.41	Machine-building and automotive technology	0.23
Retail sales	0.40	Education, social work, and theology	0.21
Agriculture, forestry, and farming	0.36	Food production and processing	0.21
Paper, printing, and technical media design	0.35	Vehicle and transport equipment operation	0.21

Note: Scores indicate the share of tasks within the occupation that are predicted to be automatable.

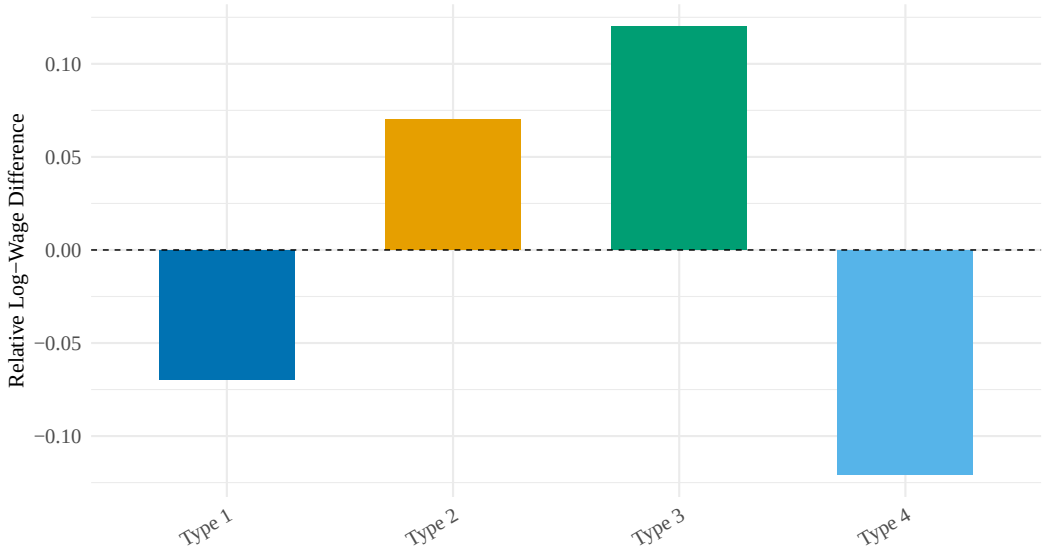
Table A3: Occupation Groups

Occupation	Group
Business management and organization	Business, Law & Admin.
Financial services, accounting, and tax	Business, Law & Admin.
Law and public administration	Business, Law & Admin.
Traffic and logistics	Commerce & Logistics
Vehicle and transport equipment operation	Commerce & Logistics
Purchasing, sales, and trading	Commerce & Logistics
Retail sales	Commerce & Logistics
Medical and health care	Health, Educ & Culture
Non-medical healthcare and body care	Health, Educ & Culture
Education, social work, and theology	Health, Educ & Culture
Teaching and training	Health, Educ & Culture
Humanities, social sciences, and economics	Health, Educ & Culture
Advertising, marketing, and media design	Health, Educ & Culture
Product design, craftwork, and fine arts	Health, Educ & Culture
Raw materials, glass, and ceramic processing	Prod, Manuf. & Construction
Plastic, wood, and wood processing	Prod, Manuf. & Construction
Paper, printing, and technical media design	Prod, Manuf. & Construction
Metal production and construction	Prod, Manuf. & Construction
Machine-building and automotive technology	Prod, Manuf. & Construction
Textile and leather production	Prod, Manuf. & Construction
Food production and processing	Prod, Manuf. & Construction
Building construction	Prod, Manuf. & Construction
Interior construction	Prod, Manuf. & Construction
Technical building services	Prod, Manuf. & Construction
Agriculture, forestry, and farming	Service & Safety
Gardening and floristry	Service & Safety
Safety, security, and surveillance	Service & Safety
Cleaning services	Service & Safety
Tourism, hotels, and restaurants	Service & Safety
Mechatronics and electrical engineering	STEM
Technical R&D, construction, and production	STEM
Construction planning and surveying	STEM
Math, biology, chemistry, and physics	STEM
Computer science and ICT	STEM

Table A4: Switching Cost / Average Wage by Occupation

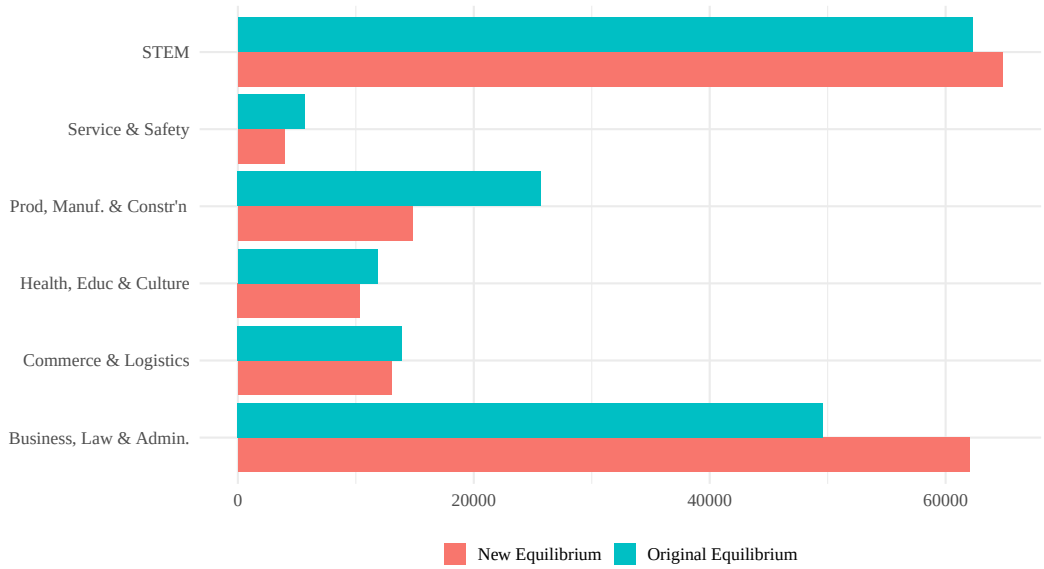
Occupation	$\bar{S}_n(\cdot\|o)/\bar{w}$	**Occupation**	$\bar{S}_n(\cdot\|o)/\bar{w}$
Agriculture, forestry, and farming	0.163	Traffic and logistics	0.524
Gardening and floristry	0.208	Vehicle and transport equipment operation	0.422
Raw materials, glass, and ceramic processing	0.178	Safety, security, and surveillance	0.352
Plastic, wood, and wood processing	0.380	Cleaning services	0.388
Paper, printing, and technical media design	0.264	Purchasing, sales, and trading	0.504
Metal production and construction	0.631	Retail sales	0.471
Machine-building and automotive technology	0.697	Tourism, hotels, and restaurants	0.334
Mechatronics and electrical engineering	0.506	Business management and organization	1.012
Technical R&D, construction, and production	0.559	Financial services, accounting, and tax	0.910
Textile and leather production	0.163	Vehicle and transport equipment operation	0.537
Food production and processing	0.408	Medical and health care	0.612
Construction planning and surveying	0.355	Non-medical healthcare and body care	0.297
Building construction	0.367	Education, social work, and theology	0.348
Interior construction	0.268	Teaching and training	0.262
Technical building services	0.436	Humanities, social sciences, and economics	0.262
Math, biology, chemistry, and physics	0.400	Advertising, marketing, and media design	0.480
Computer science and ICT	0.902	Product design, craftwork, and fine arts	0.169

Figure A1: Relative Log-Wages Across Types - Probability Threshold = 0.5



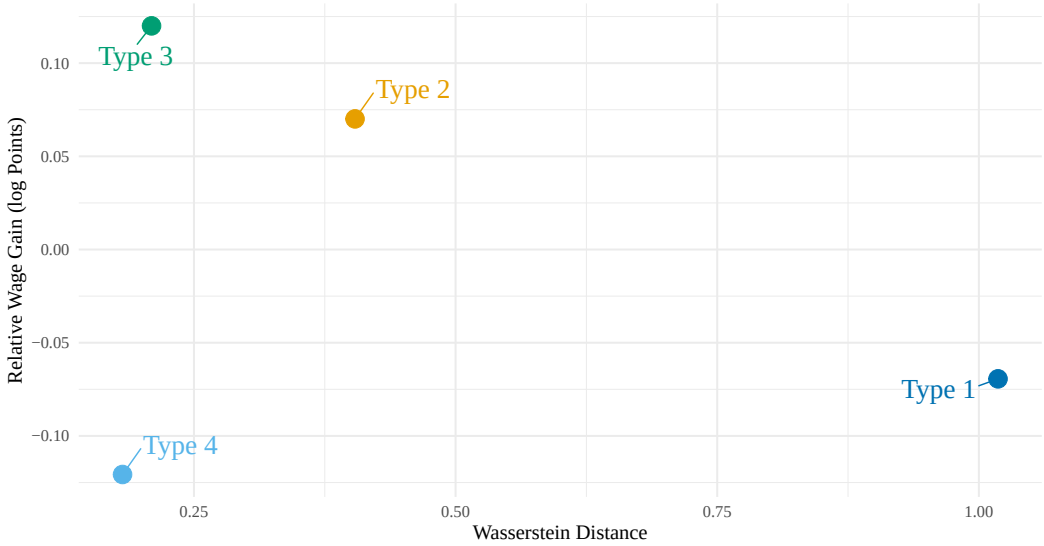
Notes: Posterior type probabilities meeting or exceeding the threshold are assigned a probability of 1, otherwise 0. y-axis indicate the difference between the wages in the post-AI and pre-AI equilibria, relative to the (unweighted) average wage difference across types. The picture with the average weighted by the type shares still look similar since type 1 and type 4 workers have a total of 45% share in the economy.

Figure A2: Worker Allocation Across the Occupation Groups - Probability Threshold = 0.5



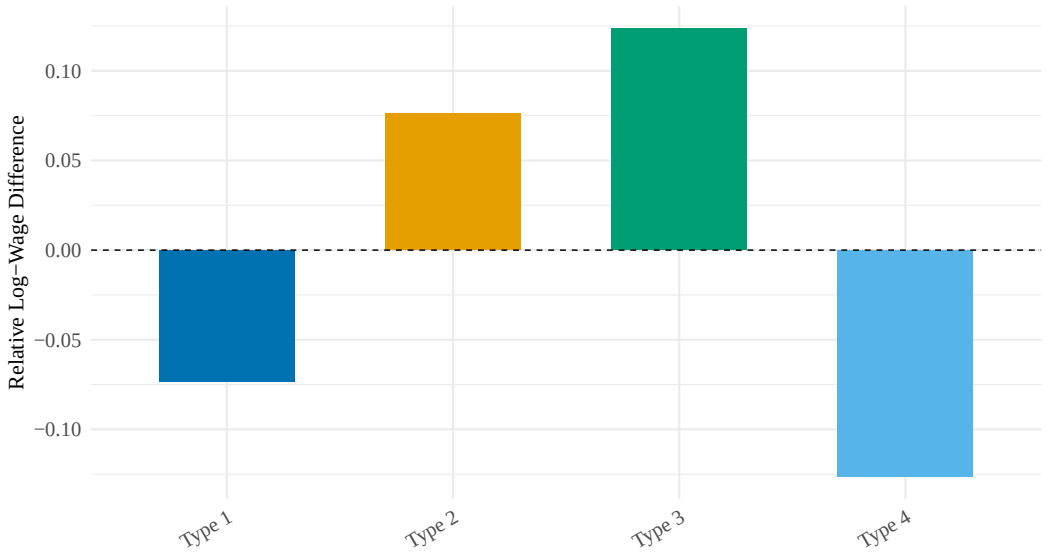
Notes: Posterior type probabilities meeting or exceeding the threshold are assigned a probability of 1, otherwise 0. x-axis denote the number of workers.

Figure A3: Displacement vs. Relative Wage Change - Probability Threshold = 0.5



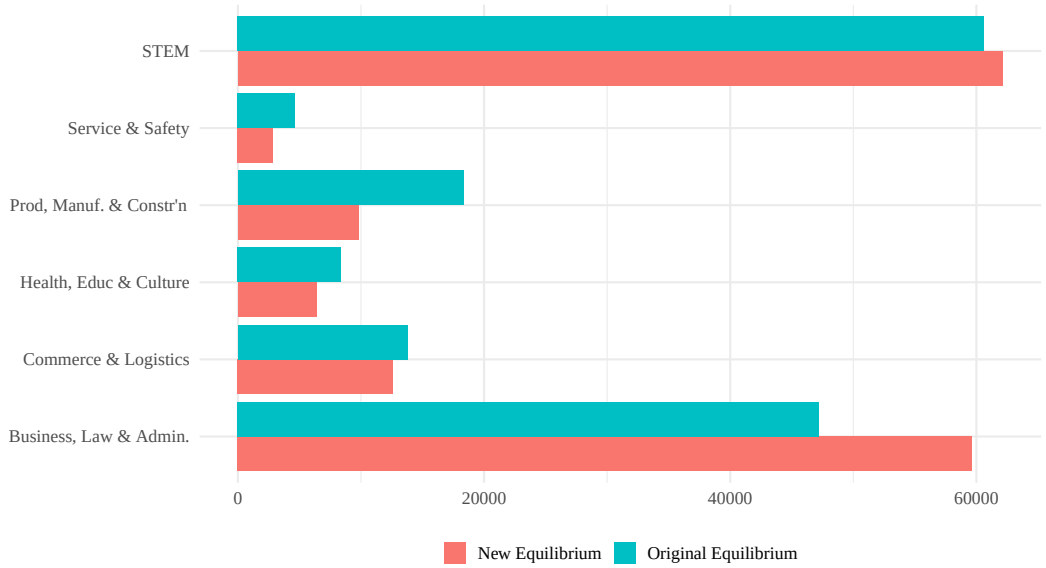
Notes: x-axis denote the Wasserstein distance between the pre and post-AI type distributions across occupations.

Figure A4: Relative Log-Wages Across Types - Probability Threshold = 0.7



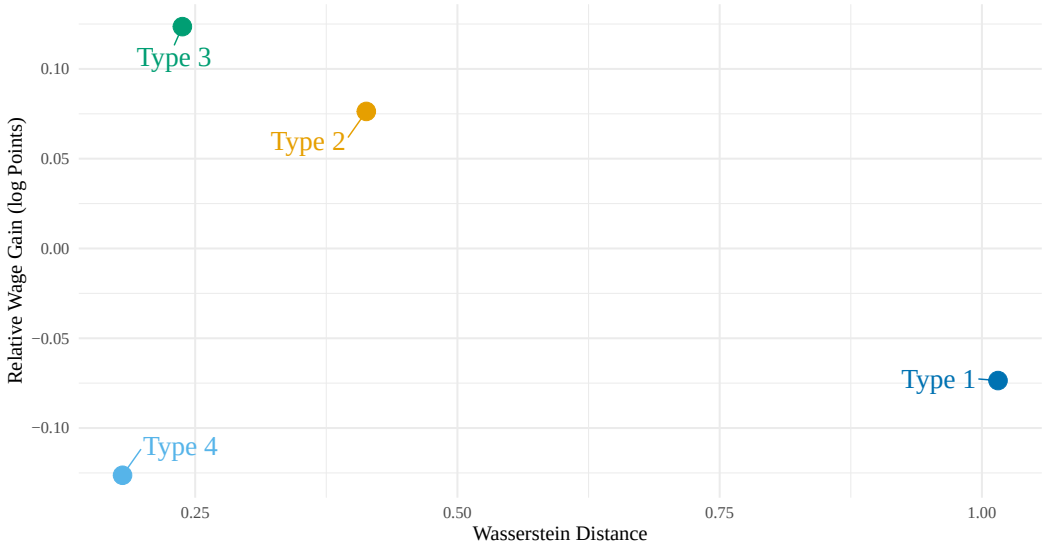
Notes: Posterior type probabilities meeting or exceeding the threshold are assigned a probability of 1, otherwise 0. y-axis indicate the difference between the wages in the post-AI and pre-AI equilibria, relative to the (unweighted) average wage difference across types. The picture with the average weighted by the type shares still look similar since type 1 and type 4 workers have a total of 45% share in the economy.

Figure A5: Worker Allocation Across the Occupation Groups - Probability Threshold = 0.7



Notes: Posterior type probabilities meeting or exceeding the threshold are assigned a probability of 1, otherwise 0. x-axis denote the number of workers.

Figure A6: Displacement vs. Relative Wage Change - Probability Threshold = 0.7



Notes: x-axis denote the Wasserstein distance between the pre and post-AI type distributions across occupations.

B Estimation Appendix

B.1 EM Algorithm

The section follows Section 3 to provide more details on the estimation procedure.

The EM algorithm is started by initiating a type distribution. To do that, I partition the occupations in 4 groups (the same as A3), and initiate probabilities based on the share of histories spent on each occupation groups. For example, a worker who spent all of their career in a single occupation group would be assigned an initial probability of 0.7 to the type initially associated with that group. During the estimation, types can be disassociated with the occupation groups or may become associated with some other occupations. The idea behind the initiation is to create enough diversity between types so that the wage and transition parameters generated by the maximization step can be diverse enough. Then, during the expectation stage, the probabilities calculated are not very uniform and the algorithm would slowly converge from that point.

Following the initiation of the type distributions, the algorithm for the EM estimation is provided in Algorithm B1.

I use $\lambda = 1e-2$ for the L2-regularization. After the log-likelihood converges, I re-estimate the parameters for the transition probabilities, this time with no penalty ($\lambda = 0$).

$$\boldsymbol{\beta}^{i,\pi,*} := \arg\min_{\boldsymbol{\beta}^{i,\pi}} ||q_{ni}\left(\mathbb{1}\boldsymbol{d}(o'|o) - X\boldsymbol{\beta}^{i,\pi}\right)||^2 \tag{B6}$$

Algorithm B1 Expectation-Maximization (EM) Algorithm for the First Stage

1: Initialize type probabilities $q_{ni}^{(0)}$ for all n, i.

2: **while** the total log-likelihood has not converged **do**

Maximization Step

3: Estimate transition parameters $\boldsymbol{\beta}^{i,\pi,*}$ for each type i via weighted L2-regularized (ridge) regression:

$$\boldsymbol{\beta}^{i,\pi,*} \coloneqq \arg\min_{\boldsymbol{\beta}^{i,\pi}} ||q_{ni}^{(m)} \left(\mathbb{1}\boldsymbol{d}(o'|o) - X\boldsymbol{\beta}^{i,\pi} \right) ||^2 + \lambda ||\boldsymbol{\beta}^{i,\pi}||^2/2 \tag{B1}$$

4: Estimate wage parameters $\boldsymbol{\beta}^{i,w,*}$ for each type i via weighted OLS:

$$\boldsymbol{\beta}^{i,w,*} \coloneqq \arg\min_{\boldsymbol{\beta}^{i,w}} ||q_{ni}^{(m)} \left(\boldsymbol{w} - X\boldsymbol{\beta}^{i,w} \right) ||^2 \tag{B2}$$

5: Estimate type probability regression parameters $\boldsymbol{\beta}^{i,q,*}$ via weighted multinomial logit:

$$\boldsymbol{\beta}^{i,q,*} \coloneqq \arg\min_{\boldsymbol{\beta}^{i,q}} ||q_{ni}^{(m)} \left(\mathbf{1} - X_1\boldsymbol{\beta}^{i,q} \right) ||^2 \tag{B3}$$

Expectation Step

6: Calculate individual likelihoods $L_{n|i}$ using the new parameters $(\boldsymbol{\beta}^{i,\pi,*}, \boldsymbol{\beta}^{i,w,*})$.

7: Generate predicted type probabilities:

$$\boldsymbol{q}^{(m+1)} \left(i|\omega_{n1}^{obs} \right) = X_1\boldsymbol{\beta}^{i,q,*} \tag{B4}$$

8: Update worker-specific type probabilities using Bayes' rule:

$$q_{ni}^{(m+1)} = \frac{L_{n|i}\boldsymbol{q}^{(m+1)}(i|\omega_{n1}^{obs})}{\sum_{i'} L_{n|i'}\boldsymbol{q}^{(m+1)}(i'|\omega_{n1}^{obs})} \tag{B5}$$

9: **end while**

Then estimate the predicted transition probabilities for the second stage.

$$\hat{\boldsymbol{\pi}}(o'|o) \coloneqq \frac{\exp(X\boldsymbol{\beta}^{i,\pi,*}_{o'})}{1+\sum_o \exp(X\boldsymbol{\beta}^{i,\pi,*}_{o}))} \tag{B7}$$

Expected wage differentials can also be recovered from the wage regression.

$$\mathbb{E}_t\hat{w}_{no't} - \mathbb{E}_t\hat{w}_{no't} = X_{nt}\boldsymbol{\beta}^{i,w}_{no'} - X_{nt}\boldsymbol{\beta}^{i,w}_{no} \tag{B8}$$

Given Eq. B7 and Eq. B8 and the distance measure between the occupations, I have all the variables for the second stage regression. Hence, I calculate the parameters for the scale parameter (γ) for the switching cost shocks and the distance-cost multiplier (α_1) as well as the fixed cost for switching from each occupation ($\{\alpha^0_o\}^O_{o=1}$).

B.2 Log-Likelihood

Figure B1 provides the log-likelihood history for the estimation. I stop the EM algorithm when the improvement in the log-likelihood get smaller than $1e-6$.

The total likelihood converges monotonically, with the exception of the step immediately following the update of the initial type probabilities. This transient non-monotonicity occurs because the logistic regressions for transition probabilities struggle to converge under the initial type distributions. This issue likely stems from two factors: numerical instabilities associated with near-zero probabilities, or the initial distributions placing disproportionate weight on outliers — observations with characteristics significantly different from the average transitioning worker. However, since the likelihood resumes monotonic growth immediately after this phase, the anomaly is best explained by the initial probabilities being far removed from the likelihood-maximizing values. It is this initialization gap, combined with the resulting convergence difficulties in the logistic regression, that drives the irregularity in the first iteration.

B.3 Derivation of the Regression Formula for the Second Stage

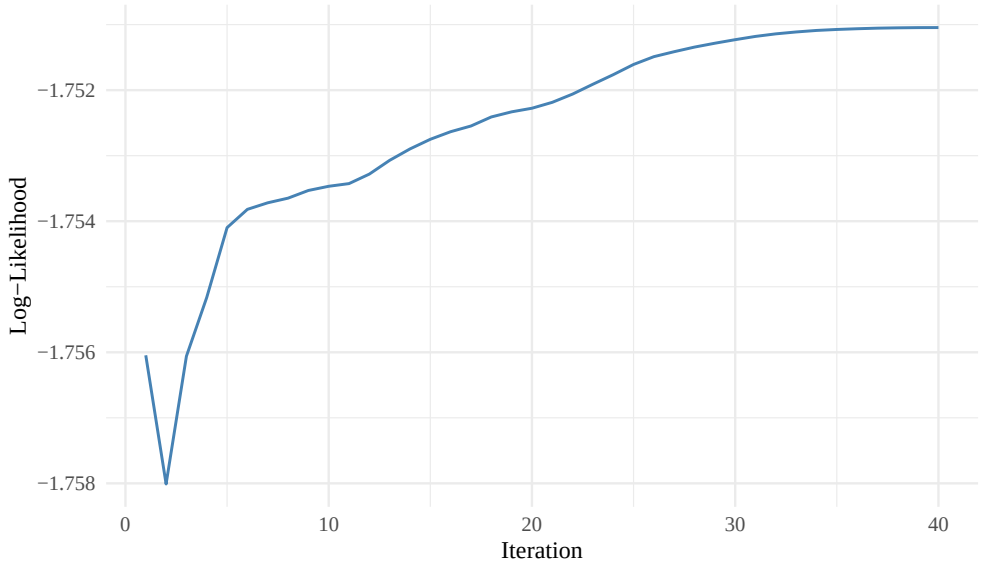
I start with the derivation of Eq. 12. To do that, first write down the relationship between the time $t+1$ unconditional value function and the time $t+1$ value function conditional on an occupation choice.

$$\mathbb{E}_t V_{t+1}(o', h_{t+1}, H_{t+1}, \omega_{nt+1}) = \gamma \int_\xi \log \sum_{o''} \exp\left(\frac{1}{\gamma} v_{t+1}(o'', h_{t+1}, H_{t+1}, \omega_{nt+1}, \xi)\right) dF(\xi) + \gamma c^e \tag{B9}$$

Due to the logit property induced by the Type 1 switching cost shocks, the probability of staying at o' at time $t+1$ is as follows.

$$\pi_{t+1}\left(o'|o', h_{t+1}, H_{t+1}, \omega_{nt+1}\right) = \frac{\exp\left(v_{t+1}(o', h_{t+1}, H_{t+1}, \omega_{nt+1}, \xi)/\gamma\right)}{\sum_{o''} \exp\left(v_{t+1}(o'', h_{t+1}, H_{t+1}, \omega_{nt+1}, \xi)/\gamma\right)} \tag{B10}$$

Figure B1: Log-Likelihood for the First Stage



Note: At 40^{th} iteration, the improvement in the log-likelihood is less than $1e-6$.

Taking the log of this probability, I get

$$\log \pi_{t+1}\left(o'|o', h_{t+1}, H_{t+1}, \omega_{nt+1}\right) = \frac{v_{t+1}(o', h_{t+1}, H_{t+1}, \omega_{nt+1}, \xi)}{\gamma} - \log \sum_{o''} \exp\left(\frac{1}{\gamma} v_{t+1}(o'', h_{t+1}, H_{t+1}, \omega_{nt+1}, \xi)\right) \tag{B11}$$

Substituting this into Eq. B9 yields the following expression.

$$\mathbb{E}_t V_{t+1}(o', h_{t+1}, H_{t+1}, \omega_{nt+1}) = \int_{\xi} \Big(v_{t+1}(o', h_{t+1}, H_{t+1}, \omega_{nt+1}, \xi) - \gamma \log \pi_{t+1}\left(o'|o', h_{t+1}, H_{t+1}, \omega_{nt+1}\right) \Big) dF(\xi) + \gamma c^e \tag{B12}$$

Since the terms inside the integral are conditional over future shock ξ, the integral is equivalent to the time t expectation.

$$\mathbb{E}_t V_{t+1}(o', h_{t+1}, H_{t+1}, \omega_{nt+1}) = \mathbb{E}_t \big[v_{t+1}(o', h_{t+1}, H_{t+1}, \omega_{nt+1}, \xi) - \gamma \log \pi_{t+1}\left(o'|o', h_{t+1}, H_{t+1}, \omega_{nt+1}\right) \big] + \gamma c^e \tag{B13}$$

Combining above equation with Eq. 11, I get to the following equality.

$$\begin{aligned} v_t(o', h_t, H_t, \omega_{nt}) = {} & w_{no't} + s_n(o'|o, \omega_{nt}) + \xi_{o'ont} \\ & + \beta v_{t+1}(o'', h_{t+1}, H_{t+1}, \omega_{nt+1}) - \beta\gamma \log \pi_{t+1}\left(o''|o', h_{t+1}, H_{t+1}, \omega_{nt+1}\right) \\ & + \beta\gamma c^e \end{aligned} \tag{B14}$$

Now suppose this worker stays at o' at time $t+1$. The value function can be written in terms of flow utility for two periods, transition probability and the continuation value.

$$\begin{aligned} v_t(o', h_t, H_t, \omega_{nt}) = \mathbb{E}_t w_{no't} + s_n(o'|o, \omega_{nt}) + \xi_{o'ont} \\ + \beta \mathbb{E}_t \left[w_{no't+1} - \gamma \log \pi_{t+1}(o'|o', h_{t+1}, H_{t+1}, \omega_{nt+1}) \right] \\ + \beta^2 \mathbb{E}_t V_{t+2}(o', h_{t+2}, H_{t+2}, \omega_{nt+2}) + \beta\gamma c^e \end{aligned} \tag{B15}$$

Following is the same identity for a worker with the same individual states who stays in occupation o at time t and moves onto occupation o' at time $t+2$.

$$\begin{aligned} v_t(o, h_t, H_t, \omega_{nt}) = \mathbb{E}_t w_{not} \\ + \beta \mathbb{E}_t \left[w_{no't+1} + s_n(o'|o, \omega_{nt+1}) + \xi_{o'ont+1} - \gamma \log \pi_{t+1}(o'|o, h_{t+1}, H_{t+1}, \omega_{nt+1}) \right] \\ + \beta^2 \mathbb{E}_t V_{t+2}(o', h_{t+2}, H_{t+2}, \omega_{nt+2}) + \beta\gamma c^e \end{aligned} \tag{B16}$$

Because Type 1 switching cost shocks induces logit probabilities, the value function differential between the two workers starting from the same occupation (before the occupation choices made in time t) can be related to the time t transition probabilities.

$$\frac{v_t(o', h_t, H_t, \omega_{nt}) - v_t(o, h_t, H_t, \omega_{nt})}{\gamma} = \log \left(\frac{\pi_t(o'|o, h_t, H_t, \omega_{nt})}{\pi_t(o|o, h_t, H_t, \omega_{nt})} \right) \tag{B17}$$

Subtracting Eq. B15 from Eq. B16, I get

$$\begin{aligned} \gamma \log \left(\frac{\pi_t(o'|o, h_t, H_t, \omega_{nt})}{\pi_t(o|o, h_t, H_t, \omega_{nt})} \right) = \mathbb{E}_t \left[w_{no't} - w_{not} \right] + \left[s_n(o'|o, \omega_{nt}) - \beta s_n(o'|o, \omega_{nt+1}) \right] \\ - \gamma\beta \log \left(\frac{\pi_{t+1}(o'|o', h_{t+1}, H_{t+1}, \omega_{nt+1})}{\pi_{t+1}(o'|o, h_{t+1}, H_{t+1}, \omega_{nt+1})} \right) + \xi_{o'ont} - \xi_{o'ont+1} \end{aligned} \tag{B18}$$

Under the assumption that $s_n(\cdot)$ does not depend on any time variant worker characteristics such as age, this equation can be further simplified.

$$\begin{aligned} \gamma \log \left(\frac{\pi_t(o'|o, h_t, H_t, \omega_{nt})}{\pi_t(o|o, h_t, H_t, \omega_{nt})} \right) = \mathbb{E}_t \left[w_{no't} - w_{not} \right] + (1-\beta) s_n(o'|o, \omega_{nt}) \\ - \gamma\beta \log \left(\frac{\pi_{t+1}(o'|o', h_{t+1}, H_{t+1}, \omega_{nt+1})}{\pi_{t+1}(o'|o, h_{t+1}, H_{t+1}, \omega_{nt+1})} \right) + \xi_{o'ont} - \xi_{o'ont+1} \end{aligned} \tag{B19}$$

Combine the probability terms to the left hand side and divide both sides by γ to get

$$\begin{aligned} \log \left(\frac{\pi_t(o'|o, h_t, H_t, \omega_{nt})}{\pi_t(o|o, h_t, H_t, \omega_{nt})} \right) + \beta \log \left(\frac{\pi_{t+1}(o'|o', h_{t+1}, H_{t+1}, \omega_{nt+1})}{\pi_{t+1}(o'|o, h_{t+1}, H_{t+1}, \omega_{nt+1})} \right) \\ = \frac{1}{\gamma} \mathbb{E}_t \left[w_{no't} - w_{not} \right] + \frac{1-\beta}{\gamma} s_n(o'|o, \omega_{nt}) + \frac{1}{\gamma} \left[\xi_{o'ont} - \xi_{o'ont+1} \right] \end{aligned} \tag{B20}$$

From the first stage, transition probabilities and expected wage differentials are calculated. I estimate this equation via OLS where the discount factor $\beta = 0.96$.

B.4 Construction of the Post-AI Equilibrium

First, I calculate the preference shifters (μ_o) by using the expenditure shares. Starting from Equation 5, I can single out the consumption preference shifters as follows.

$$\mu_o = \frac{Y_{ot}}{Y_t}\left(\frac{P_{ot}}{P_t}\right)^{\rho} \tag{B21}$$

Multiplying and dividing the right hand side by $(P_t/P_{ot})^{\rho-1}$, the consumption preference shifters can be expressed in terms of expenditure shares, individual occupation prices and the aggregate price level.

$$\mu_o = \frac{Y_{ot}}{Y_t}\frac{P_{ot}}{P_t}P_{ot}^{1-\rho}P_t^{\rho-1} \tag{B22}$$

There is no, to my knowledge, a direct source for occupational prices. Hence, I recover occupation prices and quantities from the stage estimations and from there, I can calculate every term on the RHS except the aggregate price level. However, μ_o are independent of P_t and. Hence, I calculate $\mu_o \times P_t^{1-\rho}$ and normalize the sum of μ_o to 1. For the elasticity of occupational outputs substitution (ρ), I set it to 1.78 following Burstein, Morales, and Vogel (2019)[20]. I calculate two steady states, one with no AI automation and one with AI automation. While the data is assumed to represent the steady state with no AI automation, I am calculating the no-switching cost shocks steady state (switching costs are still in place while there are no additional switching cost shocks). Therefore, to make a plausible comparison between the two steady states, I also calculate the steady state for the equilibrium with no AI technologies as well.

The iterative algorithm to find the steady state is captured in Algorithm B2. The success of this algorithm especially relies on preventing a large number of workers from switching simultaneously. When this happens, prices change significantly, and in the next iteration workers who just switched find it more profitable to switch to their previous occupation. By restricting the switches, specifically to a randomly selected 1% of the potential switchers, the fluctuations in the occupational prices get to a manageable level where the counterfactual wages across occupations gradually equalize until no worker finds it profitable to switch. Furthermore, updating the prices gradually also help, although it is by itself unable to prevent prices to jump back and forth.

Based on some trial runs, I observe that it would take either an extremely long time or impossible to get to an equilibrium where not even a single worker would want to switch. A few workers switching can fluctuate the prices just enough to make some others finding it more profitable to move back where the switching costs are low enough, which puts the algorithm into an endless loop. Therefore, I find it reasonable to stop the algorithm when there are only 100 workers who find occupation switching profitable. At this point, *true* and *effective* price vectors are virtually identical and running the algorithm further would not give me any additional precision.

[20] Authors of the study estimate this parameter with 30 occupations, using the US data.

Algorithm B2 Equilibrium Solver for Occupational Prices and Worker Allocation

Initialization:

Calculate aggregate price level $P_{agg}^{(0)}$ and aggregate output $Y^{(0)}$ from initial worker allocation.

Calculate initial occupation outputs $Y_o^{(0)}$ for all o.

Calculate initial *true* occupational prices:

$$P_{o,true}^{(0)} = \mu_o^{\frac{1}{\rho}} \left(\frac{Y_o^{(0)}}{Y^{(0)}} \right)^{-\frac{1}{\rho}}$$

Set effective prices $P_o^{(0)} \leftarrow P_{o,true}^{(0)}$.

Set iteration $k \leftarrow 0$, $num\_switchers \leftarrow \infty$.

Iteration:

while $num\_switchers > 100$ **and** $\max_o \left| \log(P_o^{(k)}) - \log(P_{o,true}^{(k)}) \right| \geq 0.005$ **do**

$k \leftarrow k + 1$

For each worker n in occupation $o_n^{(k-1)}$, find the optimal new occupation o'_n:

$$o_n^* = \underset{o'}{\operatorname{argmax}} \left\{ P_{o'}^{(k-1)} z_{ni(n)o'} M_{o'} - \frac{1}{1-\beta} s_n(o' | o_n^{(k-1)}, i) \right\}$$

Identify set of potential switchers $S \leftarrow \{n \mid o_n^* \neq o_n^{(k-1)}\}$.

$num\_switchers \leftarrow |S|$.

Select a random subset $S' \subset S$ of size $\lfloor \eta^{switch} \times num\_switchers \rfloor$.

Update worker allocation: $o_n^{(k)} \leftarrow o_n^*$ for $n \in S'$, and $o_n^{(k)} \leftarrow o_n^{(k-1)}$ for $n \notin S'$.

Calculate new individual outputs $Y_o^{(k)}$ and aggregate output $Y^{(k)}$ from new allocation $o_n^{(k)}$.

Update *true* occupational prices:

$$P_{o,true}^{(k)} = \mu_o^{\frac{1}{\rho}} \left(\frac{Y_o^{(k)}}{Y^{(k)}} \right)^{-\frac{1}{\rho}}$$

Update effective occupational prices (with damping):

$$P_o^{(k)} \leftarrow \lambda^P P_o^{(k-1)} + (1 - \lambda^P) P_{o,true}^{(k)}$$

end while

C Data Appendix

C.1 Data Preparation

The raw data contains more than 77 million observations after episode splitting and more than 55 million observations before episode splitting. Data providers perform episode splitting whenever two episodes overlap. In such cases, they generate two extra episodes for the overlapping period. For example, consider an episode with start and end dates 01/01/2001 and 12/31/2001. Consider another episode with start and end dates 09/01/2001 and 09/01/2002. Data providers create two additional episodes with start and end dates 09/01/2001 and 12/31/2001, where the information is transferred from the two overlapping episodes, therefore, ending up with 4 episodes instead of the original 2. These episode splitting artificially boost the transition probabilities from and to the same occupation, leading biased estimates for the first stage. To eliminate these generated episodes, I only keep the observations if they contain the mid-year between their start and end dates.

I first keep the observations that belong to two sources, Employee History (BeH) and Benefit Recipient History (LeH). Benefit Recipient History keeps track of people who receive unemployment benefit or unemployment assistance. Employee History data has the information on people with an active employment and it is the one that the most estimation parameters rely on, other than the transition probabilities. Since I am not taking into account the unemployment in the post-AI equilibrium, unemployment histories do not alter the post-AI worker allocation.

There are some other data sources that I drop from the data. Unemployment Benefit II Recipient History (LHG) starts from 2005, being much later than the 1998 threshold, I drop this data source. Similarly, I also drop Participants-In-Measures History Files (MTH/XMTH) due to starting from 2000 for MTH and 2005 for XMTH. Jobseeker Histories (ASU/XASU). The last data source, Jobseeker Histories (ASU/XASU) starts from 1997 for ASU and from 2005 for XASU. With this data source, however, there are many episode splittings (split episodes cover half of the observations) and I drop this data source from the estimation altogether.

There are also some missing episodes for some workers. This prevents transition probabilities from being correctly estimated. Whenever there are such cases, I keep the observations with the longest no-gap history. If there are at least 2 set of observations with equal length, then I keep the most recent one as estimation of the more recent periods would be more important in terms of making predictions about the future.

I also drop one occupation based on the criterion that it spans less than 1/1000th of all the observations. I merge two occupations with a low count of observations and related titles "Occupations in product design, artisan craftwork, fine arts and the making of musical instruments" and "Occupations in the performing arts and entertainment".

C.2 Principal Components

Each O*NET SOC occupation is associated with some metrics, classified under "Knowledge", "Skills", "Abilities" and "Work Activities". There are 160 attributes in total. Some attributes are very similar, and same attributes such as "Mathematics" is a part of both "Skills" and "Abilities" metrics. Hence, I use the principal components to reduce the very high dimension of attributes. Table C1 shows the most significant loadings for the first 8 principal components. Figure C1 shows the where the occupations stand in the first 2 principal components space.

Table C1: Most Positive and Negative Loadings for the Principal Components

Principal Component 1

Most Negative		Most Positive	
Task	Loading	Task	Loading
Manual Dexterity	-0.102	Written Expression	0.114
Extent Flexibility	-0.101	Written Comprehension	0.114
Handling and Moving Objects	-0.100	Writing	0.114
Static Strength	-0.099	Reading Comprehension	0.114
Dynamic Strength	-0.099	Active Learning	0.114

Principal Component 2

Most Negative		Most Positive	
Task	Loading	Task	Loading
Working with the Public	-0.068	Quality Control Analysis	0.148
Customer and Personal Service	-0.053	Mechanical	0.148
Fine Arts	-0.052	Inspecting Equipment, Material	0.149
Service Orientation	-0.052	Operation Monitoring	0.150
Establishing Interpersonal Relat.	-0.042	Physics	0.153

Principal Component 3

Most Negative		Most Positive	
Task	Loading	Task	Loading
Programming	-0.142	Working with the Public	0.154
Interacting With Computers	-0.130	Resolving Conflicts and Negotiating	0.155
Computers and Electronics	-0.126	Psychology	0.169
Near Vision	-0.110	Assisting and Caring for Others	0.180
Engineering and Technology	-0.102	Therapy and Counseling	0.196

Principal Component 4

Most Negative		Most Positive	
Task	Loading	Task	Loading
Sales and Marketing	-0.240	Biology	0.139
Management of Material Resources	-0.196	Assisting and Caring for Others	0.151
Management of Financial Resources	-0.195	Identifying Objects, Actions, and Events	0.158
Production and Processing	-0.178	Medicine and Dentistry	0.170
Selling or Influencing Others	-0.178	Documenting/Recording Information	0.242

Principal Component 5

Most Negative	Most Positive

Continued on next page

Table C1 – continued from previous page

Most Negative		Most Positive	
Task	Loading	Task	Loading

Task	Loading	Task	Loading
Fine Arts	-0.329	Economics and Accounting	0.113
History and Archeology	-0.276	Processing Information	0.113
Thinking Creatively	-0.228	Performing Administrative Activities	0.125
Philosophy and Theology	-0.209	Number Facility	0.132
Sociology and Anthropology	-0.185	Determine Compliance	0.173

Principal Component 6

Most Negative		Most Positive	
Task	Loading	Task	Loading
Geography	-0.321	Monitor Processes, Materials	0.113
Transportation	-0.302	Finger Dexterity	0.116
Telecommunications	-0.238	Arm-Hand Steadiness	0.118
Law and Government	-0.205	Instructing	0.120
Spatial Orientation	-0.198	Training and Teaching Others	0.144

Principal Component 7

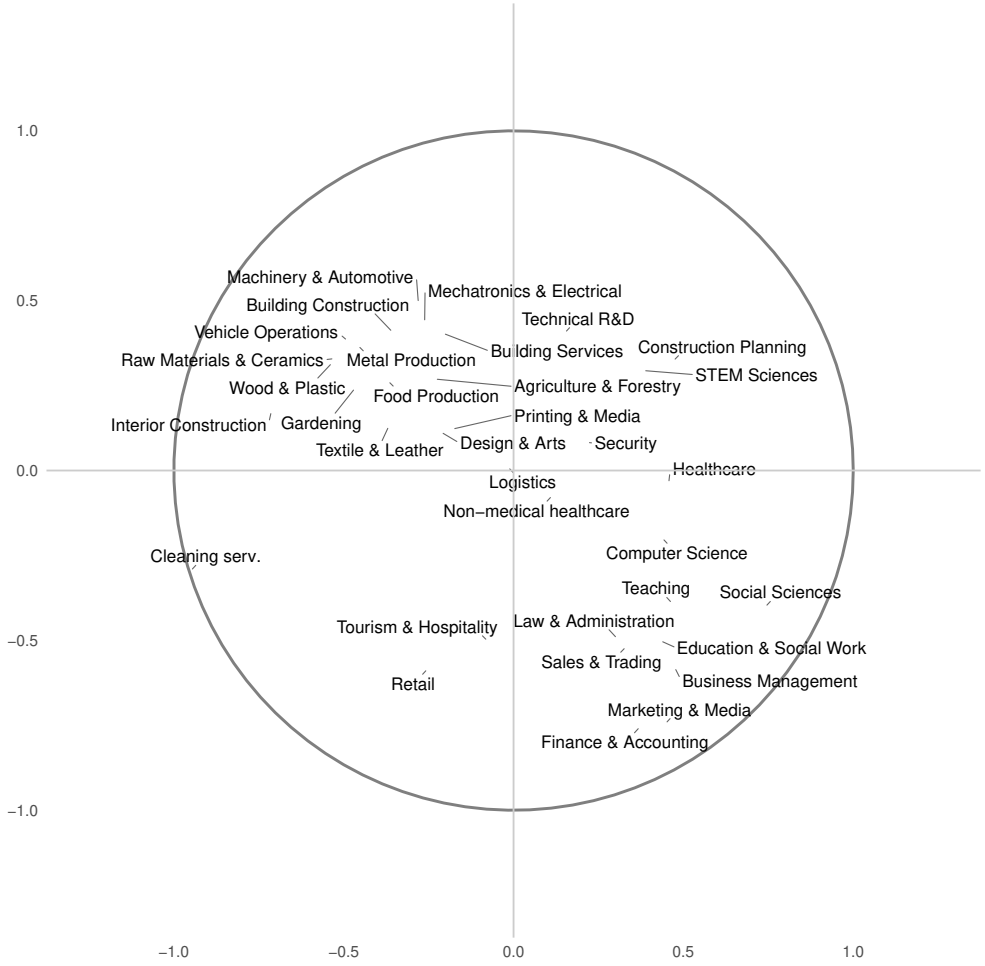
Most Negative		Most Positive	
Task	Loading	Task	Loading
Food Production	-0.218	Time Sharing	0.185
Biology	-0.162	Finger Dexterity	0.190
Estimating the Characteristics of Info.	-0.131	Telecommunications	0.234
Dynamic Flexibility	-0.114	Clerical	0.238
Geography	-0.100	Customer and Personal Service	0.259

Principal Component 8

Most Negative		Most Positive	
Task	Loading	Task	Loading
Chemistry	-0.243	Peripheral Vision	0.117
Biology	-0.241	Speech Clarity	0.118
Economics and Accounting	-0.240	Installation	0.141
Customer and Personal Service	-0.219	Sound Localization	0.148
Sales and Marketing	-0.190	Selective Attention	0.152

Note: I use the first 8 principal components in the estimation of the comparative advantage parameters, which explain more than 80 percent of the total variance in the entire space.

Figure C1: Occupations in the first Two Principal Components Dimension



Note: x-axis is for the first principal component and y-axis is for the second principal component. The principal components are scaled with the variance of the first principal component, ensuring that all occupations would be in the interior of the unit circle.

C.3 Crosswalks

There are no direct crosswalks between O*NET SOC-2010 classification and KldB-2010 classification. To match the occupations, I first generate a crosswalk from SOC 2010 to ISCO-08, then to KldB-2010 classification.

I generate two crosswalks for two cases, (i) automation scores and (ii) principal components. The idea behind both are the same and as follows.

First, the crosswalk between SOC-2010 and ISCO-08 matches 6-digit SOC-2010 occupations to 4 digit ISCO-08 occupations. I use the automation scores for the tasks associated with 6-digit SOC-2010, if the O*NET task statements are available for that 6-digit occupation. If not, I look at the children occupation in the SOC-2010 classification, and get to the parent

task statements by combining them. If the 6-digit's children do not exist in the O*NET task statements, I combine the task statements of the siblings, and use them as if they are the task statements that belong to the 6-digit occupation.

With regards to the principal components, the procedure is exactly the same, with the only difference being that instead of combining the task statements, I take simple averages of the principal components. The simple average could be of the children occupations if they have attributes listed in the O*NET database, or of the siblings if not.

There are 34 (excluding army occupations and the 2 eliminated occupations due to the low observation count) occupations in the 2-digit KldB-2010 classification, and there are around 1000 6-digit occupations in the SOC-2010 classification. Using the information on the children or siblings in the occupation hierarchy is an exemption, and there are on average more than 20 6-digit SOC-2010 occupation for each KldB-2010 occupation, which should make any bias due to missing information on 6-digit SOC-2010 occupation negligible.

D LLM Prompt

Following is the initial prompt given to Gemini 2.5 Flash model.

Consider the most powerful Google Gemini large language model (LLM). This model can complete many tasks that can be formulated as having text/audio/video input and text/audio/video output. This model have access to up-to-date facts from internet or any information or database that is relevant for the task.

You are a helpful assistant who wants to label the given tasks according to the rubric below. Equivalent quality means someone reviewing the work would not be able to tell whether a human completed it on their own or with assistance from the LLM. If you aren't sure how to judge the amount of time a task takes, consider whether the tools described exposed the majority of subtasks associated with the task.

Exposure rubric:

E1 - Direct exposure: Label tasks E1 if direct access to the LLM through an interface alone can reduce the time it takes to complete the task with equivalent quality by at least half. This includes tasks that can be reduced to: - Writing and transforming text and code according to complex instructions, - Providing edits to existing text or code following specifications, - Writing code that can help perform a task that used to be done by hand, - Translating text between languages, - Summarizing documents, - Providing feedback on documents, - Answering questions about a document, - Generating questions a user might want to ask about a document, - Writing questions for an interview or assessment, - Prepare training materials based on general knowledge, - Inform anyone of any information via written or spoken media.

E0 - No exposure: Label tasks E0 if none of the above clearly decrease the time it takes for an

experienced worker to complete the task with high quality by at least half. Some examples: - If a task requires a high degree of human interaction (for example, in-person demonstrations) then it should be classified as E0. - If a task requires precise measurements then it should be classified as E0. - If a task requires any use of a hand or walking then it should be classified as E0. - LLM cannot make any decisions that might impact human livelihood (e.g. hiring, grading, etc.). If any part of the task involves collecting inputs to make a final decision (as opposed to analyzing data to inform a decision or make a recommendation) then it should be classified as E0. The LLM can make recommendations. The LLM cannot do anything that legally requires a human to perform the task. - If there is existing technology not powered by an LLM that is commonly used and can complete the task then you should mark the task E0 if using an LLM or LLM-powered tool will not further reduce the time to complete the task. When in doubt, you should default to E0.

Read the following tasks that belong to an occupation. Then do the following: 1. Reason very briefly step by step to decide which of the labels (E0 or E1) from the exposure rubric you were given applies to the task's exposure to LLM. Report. 2. Give an explanation for which label applies and report the label that you think fits best. Report only the label that you determined for the task. 3: 3. Given the amount of speculation required in step 1, describe your certainty about the estimate – either high, moderate, or low.

Direct or coordinate an organization's financial or budget activities to fund operations, maximize investments, or increase efficiency.

LLM Response: *The task 'Direct or coordinate an organization's financial or budget activities to fund operations, maximize investments, or increase efficiency' involves high-level strategic decision-making, leadership, and ultimate responsibility for financial outcomes. These decisions directly impact the organization's stability, employee livelihoods, and overall business operations. The rubric explicitly states: 'LLM cannot make any decisions that might impact human livelihood (e.g. hiring, grading, etc.). If any part of the task involves collecting inputs to make a final decision (as opposed to analyzing data to inform a decision or make a recommendation) then it should be classified as E0.' While an LLM can analyze financial data, generate reports, draft proposals, or make recommendations (which are E1 activities), it cannot 'direct or coordinate' by making the final, responsible decisions or taking on the leadership role required for such a task. Furthermore, this task requires a high degree of human interaction, negotiation, and strategic judgment that an LLM cannot replicate with equivalent quality. Therefore, the core of the task is not exposed to LLM capabilities in a way that would reduce completion time by half with equivalent quality. E0 High*

...

...

...

I use the "chat" feature in the API, which allows me to respond to every output by the next

task description. This allows me to save fees and time by not having to input the initial prompt every time.