AI Automation and Labor Market Outcomes§

Yakup Kutsal Koca*

November 12, 2025

Click here for the most recent version

Abstract

I study the effects of AI automation on the wage distribution across workers and the employment distribution across occupations. I build a general equilibrium framework where workers can switch in response to changes in occupational prices. This allows workers who are directly exposed to mitigate the adverse effects of the AI automation, whereas the opposite effect takes place for workers who are not directly targeted by the AI technologies but face higher competition in the labor market since the exposed workers switching to their occupations.

JEL Codes:

§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 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. This general equilibrium effects can mitigate the income loss for those who are replaced. On the other hand, the GE effects are likely to be adverse for those who are not directly exposed to AI, but face higher competition in the labor market due to worker reallocation.

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 the occupations. I do this in a setting where workers can reallocate across occupations as a result of prices changing due to the AI technologies. There are two main ingredients of the general equilibrium. First ingredient is the automation scores for every occupation that I calculate following Eloundou et al. (2023) and Eisfeldt et al. (2023). AI technologies change the occupation outputs and therefore prices, and I quantify the changes by assuming a relationship between the automation scores and the occupation outputs. Second ingredient is the wage parameters, including the comparative advantage, for each worker that lets me calculate the counterfactual wages after the AI shock.

I set up a framework where in the production side the aggregate output is made up of occupation outputs, and each occupation consists of a series of tasks, as common in the automation literature (Acemoglu and Restrepo (2022), Acemoglu and Restrepo (2018) and Acemoglu, Autor, et al. (2022) and Humlum (2021)). Some of the tasks can be automated, in which case they are entirely performed by the AI technologies. Workers perform the non-automated task, and their earnings are a function of occupation prices, some observables such as age and schooling, and their unobserved comparative advantage vector across occupations.

On the labor supply side, workers face a dynamic discrete choice problem every period where they can either stay in their current occupation or switch to a different occupation subject to switching costs. The *dynamic* part stems from switching costs, and the workers need to calculate the expected lifetime value of each choice and weigh it against the switching costs. To calculate the occupation choices in an equilibrium, I need to calculate expected lifetime value of workers under any occupation choice. This requires knowing in full the wage parameters, including the productivity vectors.

Estimation of the unobserved productivity (or comparative advantage) vectors relies is through the wage regression, where the identification comes from the occupation to occupation transitions. The comparative advantage of a worker in occupation o would be roughly their wage in an occupation o at time t compared to the average wage in that occupation at time t, averaged over the years. However, this identification method fails because not even a single worker has an employment history across all occupations. To overcome this, I assume that there are finite worker types, where each type represents a latent group of workers who share a similar, unobserved comparative advantage vector. These types are not pre-defined but are estimated directly from the data. This way, the employment history of the set of workers who belong to the same type spans the set of all occupations, and the comparative advantage vector can be identified in this manner.

However, estimation of the type probabilities[1] and the wage regression[2] simultaneously is not a computationally feasible task. Therefore, I utilize an Expectation-Maximization algorithm which updates the type probabilities and wage regression sequentially, and eventually converging to the maximum. Specifically, I follow Arcidiacono and Miller (2011) which lets me estimate the wage regression parameters with unknown types first, and the switching costs in the next stage.

In order to estimate the switching costs, I utilize the relationship between the transition probabilities, which are recovered during the first stage of the estimation, value functions and the switching costs. Taking the difference of the value function of two workers who start and end up in the same occupations[3], the same relationship can be expressed in terms of the transition probabilities and the wage differential between the two workers, both of which are estimated during the first stage of the estimation, and switching costs, which can be estimated using the first two estimated variables.

[1]In the actual estimation process, each worker has a strictly positive probability of being any type due to the way the probabilities are constructed to maximize the likelihood.

[2]Conditional choice probabilities are also estimated as part of the likelihood. See Section B for additional details.

[3]This is due to the finite dependence property which allows differencing out the lifetime value functions (Arcidiacono and Miller (2011))

For estimation, I use an administrative German panel data which tracks the employment history of workers between years 1998-2021[4]. This is a panel data that 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. Employing this administrative data with many occupation to occupation transitions allow me to identify the comparative advantage parameters[5].

Having estimates of the wage equation parameters, I can compute counterfactual wages under any worker allocation. In other words, I can compute the equilibrium given any shocks to occupation prices. In essence, AI automation shock is a price shock as it distorts the occupation prices which triggers a worker reallocation. The way this shock works is it makes the workers more productive by having them allocate the time spent on the automated tasks to the non-automated tasks. This in turn increase the wages and makes some occupations more preferable to others, while on the other hand reducing the price of those occupations due to increased quantity. This structural framework is what allows me to trace the full effect of the AI shock, not only on the wages of *replaced* workers, but also on *non-exposed* workers who now face new labor market competition. This allows for a complete assessment of the distributional and welfare impacts.

In Section 2, I lay out the production hierarchy and the production technologies. Also in this section, I provide the equilibrium conditions; market clearing conditions for both the goods and the labor market, optimal consumption rules for household and profit maximization problem for the firms. Following that, I describe the labor supply side and the workers' dynamic choice problem. In Section 3, I provide details about the estimation of the wage parameters and the type probabilities. Next, I describe the second stage where I estimate the switching costs utilizing the first stage estimates. In Section 5, I provide relevant facts about the data including occupational transition rates and details for the construction of the panel that I use for the estimation. Following this, I describe the matching procedure from the German occupation classification (KldB-2010) to the US O*NET SOC classification, where I get some occupation metrics from. Furthermore, this section also covers the process for generating the AI automation scores. In Section 6, I describe how I numerically construct the post-AI equilibrium. This relies on shifting the workers between occupations such that their wages are maximized. Then I discuss the estimation results including the

[4]While the raw data covers earlier years, I use the data starting 1998. For a discussion on the data cleaning procedure, please see Section 5

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

wage distribution, employment distribution along with some important parameters such as comparative advantage vectors.

I contribute to the automation literature, mainly to those concerning the AI automation, including the aforementioned works such as Brynjolfsson, Mitchell, and Rock (2018), Eloundou et al. (2023), Felten, Raj, and Seamans (2018) and Webb (2020), Mckinsey & Company (2023), Humlum and Vestergaard (2024) 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 the AI automation exposure scores and the occupational prices or wages of the workers in that occupation. For example, if highly exposed translators have a positive comparative advantage only in the occupation "translators", then they might be adversely affected by the AI automation. On the other hand, if another highly exposed group, computer scientists, have a comparative advantage in an unaffected occupation, such as engineering, then they will move to that occupation and mitigate the adverse effects of the AI automation.

This study also adds to the previous work that study the general equilibrium effects of the automation shocks using a reduced form analysis, such as Acemoglu and Restrepo (2022). In Acemoglu and Restrepo (2022), authors estimate a propagation matrix of the automation shock that measures the propagation of an automation shock to a set of tasks to the other tasks. The contribution of this study on this strand of literature is that I am *structurally estimating* this propagation matrix, which depends on the primitives of the model environment and worker characteristics and comparative advantage vectors. A reduced form approach is not feasible for the case of AI automation shock since AI technologies are in adoption phase, however, a structural approach also gives me flexibility in terms of measuring the labor market responses to any degree of AI automation shock.

I show that ...

2 Theory

2.1 Production

The economy is populated by human workers[6] normalized to unity. Time is discrete. Every period, each worker chooses among O different occupations to work. Each occupation

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