Job Mobility Within and Across Occupations*

Attila Gyetvai†

Duke University

Job Market Paper, October 2020. Latest version here.

Abstract

This paper assesses the impact of occupational mobility on life cycle wage dynamics. I develop a model of job mobility which attributes differential returns to occupations to occupationally heterogeneous labor market frictions, compensating differentials, and non-pecuniary job switching costs. I estimate the structural model on linked Hungarian administrative data and use it to arbitrate between these mechanisms. High-skill occupations offer higher wages and more stable employment but lower non-wage amenities than low-skill ones. Coupled with less frequent offers and higher costs of switching from low-skill to high-skill jobs, workers who start in high-skill occupations have much steeper wage profiles. Together, these sources fully account for the life cycle inequality in wages, with each source contributing equally to fitting the empirical patterns.

*I thank Peter Arcidiacono as well as Arnaud Maurel, Seth Sanders, and Pat Bayer for their guidance and never-ending support. This paper benefited from useful discussions with Joe Altonji, Limor Golan, Thomas Le Barbanchon, Rasmus Lentz, Attila Lindner, Robert Moffitt, Jean-Marc Robin, Isaac Sorkin, Todd Stinebrickner, and members of the labor group at Duke, as well as comments by seminar and conference participants at SEA, HSE, EALE-SOLE-AASLE, ESWC, and the St. Louis Fed. Thanks to CERS-HAS for granting access to the data. Errors are mine.

†attila.gyetvai@duke.edu, attilagyetvai.com.

Wages vary widely among workers. One prominent explanation, the basic theory of human capital, attributes this empirical observation to differential returns to skills: some workers have higher stocks of human capital that command higher pay. Other theories, including labor market frictions and compensating differentials, offer competing explanations. Labor market frictions imply that some workers find better luck in rising through the ranks while some unlucky workers lose their jobs more frequently than others. Compensating differentials capture the idea that some workers require monetary compensation to endure undesirable job traits.

This paper argues that human capital contributes to wage inequality not only through differential returns to skills, but also through labor market frictions and compensating differentials that vary across skills. Returns to skills arise, in part, from labor market frictions: not only workers receive different payment for their skills, they also receive job offers and separation shocks at varying rates across skill levels. Another source of returns to skills is compensating differentials: workers at certain skill levels may receive a larger part of their compensation in non-wage job amenities, such as better working conditions or lower effort requirements.

I formalize this argument by proposing and estimating a model of job mobility within and across occupations. Occupations constitute an observable measure of skills; *[smooth this out]*. In the model, wage offers differ across occupations, capturing differences in pure returns to skills. Workers receive job offers and separation shocks at various rates across occupations, capturing skill-specific labor market frictions. Occupations offer different non-wage amenities, resulting in skill-specific compensating differentials. Finally, workers in different occupations face different non-pecuniary costs to switch into a certain occupation. Together, these mechanisms imply that jobs in various occupations differ not only in their flow values but also in the options they afford their workers.

I estimate the structural model using linked Hungarian administrative data on half of the country's population: notably, these data contain administrative information on workers' occupational transitions at a high frequency. Using these rich data, I uncover substantial occupational heterogeneity in the model mechanisms. I find that high-skill occupations offer higher wages and lower job separation rates than low-skill ones, and workers in high-skill occupations receive these better-paying offers more frequently than those in low-skill occupations. At the same time, low-skill occupations offer higher non-wage amenities. However, not everyone can take advantage of the wage gains from entering high-skill occupations: low-skilled workers face larger non-pecuniary costs to switch to these jobs, essentially locking them in lower wage trajectories. Taken together, these effects imply that high-skill jobs are more valuable than low-skill jobs: their lower levels of flow utility are offset by higher option values.

Using the structural estimates, I assess the contribution of occupational heterogeneity in each model mechanism to wage inequality. Upon entering the labor market, high-skilled workers who start in low-wage jobs expect rapidly rising wages: this steep wage growth stems from better wage offers in high-skill occupations. At the same time, low-skilled workers in high-wage jobs expect a sharp drop in their wages, due to more frequent job separations. Together, these forces imply that initial occupations shape workers' expected wage paths to a larger extent than their initial wages. Measures of model fit indicate that occupational heterogeneity in each of the captured mechanisms accounts for most of the empirical dispersion of wages. Allowing all mechanisms to interact with occupations fully accounts for the rising inequality of wages over the life cycle.

Literature: This paper contributes to multiple literatures. I build on a line of work on occupational mobility (McCall, 1990; Kambourov and Manovskii, 2009a,b; Groes, Kircher, and Manovskii, 2015). Papers in this literature typically posit that transitions result from workers' choices (Miller, 1984; Siow, 1984; Antonovics and Golan, 2012). In contrast, I argue that market frictions play an equally important role in understanding occupational mobility. More broadly, I contribute to papers that model career decisions in a dynamic discrete choice setting (Keane and Wolpin, 1997; Neal, 1999; Sullivan, 2010; Sullivan and

To, 2014). In comparison, the present model highlights the role of occupation-specific labor market frictions; my strategy extends the framework in accompanying work (Arcidiacono, Gyetvai, Jardim, and Maurel, 2020).

Another stream of papers focuses on estimating compensating differentials (Rosen, 1986; Sorkin, 2018; Arcidiacono, Hotz, Maurel, and Romano, 2020). I bring occupation-specific estimates to the table: I calculate worker's preferences towards occupations, parsed from their opportunities to enter or leave them. I also add to the vast literature on the job search behavior of heterogeneous workers (Postel-Vinay and Robin, 2002; Cahuc, Postel-Vinay, and Robin, 2006; Jolivet, Postel-Vinay, and Robin, 2006; Taber and Vejlin, 2020). Compared to these models, I uncover occupational heterogeneity in the search frictions.

Furthermore, my paper relates to two tangential literatures. An influential stream of papers attribute the dispersion of wages to worker- and firm-specific heterogeneity (Abowd, Kramarz, and Margolis, 1999; Card, Heining, and Kline, 2013; Card, Cardoso, Heining, and Kline, 2018; Bonhomme, Lamadon, and Manresa, 2019). I take a more structural approach and argue that occupation-specific heterogeneity explains the life cycle profile of wage inequality. Another tangentially related literature models the labor market through the lens of multiple job ladders or island economies (Pilossoph, 2014; Wiczer, 2015; Jarosch, 2015; Busch, 2017; Carrillo-Tudela and Visschers, 2020). Focusing on a different set of questions, these papers analyze the dissipation of income or productivity shocks through the reallocation of resources across islands; however, they typically impose the same amount of frictions on searching across islands. I add to this literature by modeling heterogeneous frictions across occupations and demonstrating that occupations provide a key source of heterogeneity in understanding life cycle wage dispersion.

Roadmap: The rest of the paper unfolds as follows. Section 1 presents the data and descriptive evidence that high-skill occupations offer faster wage growth than low-skill ones. Section 2 introduces the model. Section 3 discusses the identification of the structural parameters. Section 4 contains and interprets the estimation results. Section 5 connects these results to life cycle wage inequality. Finally, Section 6 concludes.

1 Descriptive Evidence

1.1 Data

I use matched employer-employee data from Hungarian administrative records covering half of the population, i.e., 4.6 million individuals, linked across 900,000 firms. These data have several strengths over other frequently used data sets: most importantly, they include detailed occupations from administrative sources, and they follow workers in continuous time.[1]

On the individuals' side, a de facto 50 percent random sample[2] of the Hungarian population are observed; every Hungarian citizen born on Jan 1, 1927 and every second day thereafter are included, yielding a sample of 4.6 million people. Labor market measures are recorded in up to two work arrangements at a time; importantly, one of these measures is occupation.[3] The occupational classification maps major occupational groups to skill content, proxied by the educational requirement of occupations within the group: Appendix Table A.1 presents this mapping for the occupations I include.

For the purposes of this paper, I restrict these rich data along several dimensions. First, I discard observations from 2011, the last year of the sample. The Hungarian occupation classification system was restructured in 2010, and the new classification is used from 2011. Unfortunately, the old and new occupation codes are not harmonized, and doing so would discard a substantial share of occupations. Second, I include males of age 22–50 who held at least one job in the sample, in an attempt to include as homogeneous analysis units in the sample as possible. By age 22, college students have presumably finished their education; the

[1]The data are constructed from continuous-time spells, and allow me to observe the approximate date of employer and occupation changes.

[2]I adopt the terminology of DellaVigna, Lindner, Reizer, and Schmieder (2017).

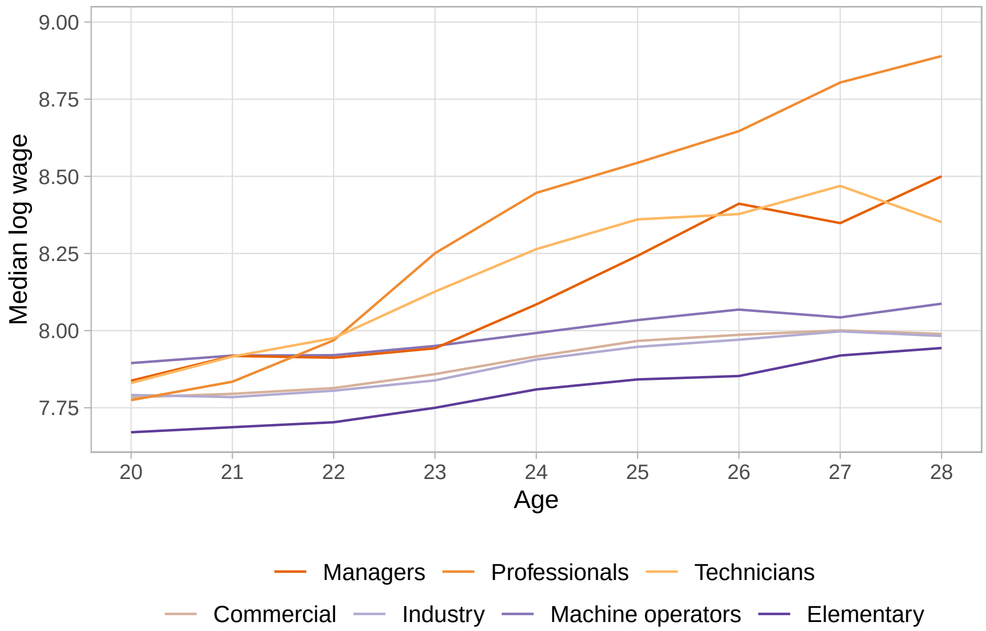
[3]4-digit occupations, based on the Hungarian Standard Classification of Occupations (HSCO). HSCO is similar in spirit to the Standard Occupational Classification (SOC) system by the BLS.

data do not contain information on educational attainment for most individuals, thus I resort to excluding younger workers. Furthermore, various pension rules are in effect for males from age 53; I drop older workers to ensure my results are not contaminated by their anticipatory behavior. Third, I include workers at privately owned firms. Hungary is a post-socialist welfare state, and thus has a large public sector: 30 percent of worker-month observations in the data are working in some form of a public sector job. Since these jobs are contracted differently than private sector jobs, including them would compromise the homogeneity of the sample. Fourth, I drop individuals who work in agricultural occupations and armed forces as transitions into and out of these occupations occur infrequently. Furthermore, I add office clerks to commercial occupations as they are comparable in skill content, and because they make up a small part of the sample on their own. The seven remaining occupations are Managers, Professionals, Technicians, Commercial occupations, Industry occupations, Machine operators, and Elementary occupations: see Appendix Table A.1. Fifth, finally, I only consider data on the primary work arrangement for all workers, even if they reportedly work in multiple parallel jobs.

These rich data accommodate the enormous requirements of estimating the model. I consider a discrete distribution of jobs, comprised of a handful of occupations and wage bins. Besides the eight occupations, I discretize wages into ten bins: I define the first wage bin as between 75 and 100 percent of the effective minimum wage, and bins 2 to 10 as nine-quantiles of the conditional wage distribution, truncated at the minimum wage. Therefore, job-to-job transitions fall into one of 6,400 cells:[4] Appendix Figure A.1 presents the entire job-to-job transition matrix. Most transitions occur within the same job. Conditional on an occupational switch, workers tend to move into occupations that are similar in terms of skill content to their current one. However, a sizable fraction of transitions occur from low-skill to high-skill occupations and vice versa.

[4] $8 \cdot 10 \times 8 \cdot 10 = 6,400$ job-to-job transition cells.

Figure 1.1: Diverging occupational wage paths early in the life cycle



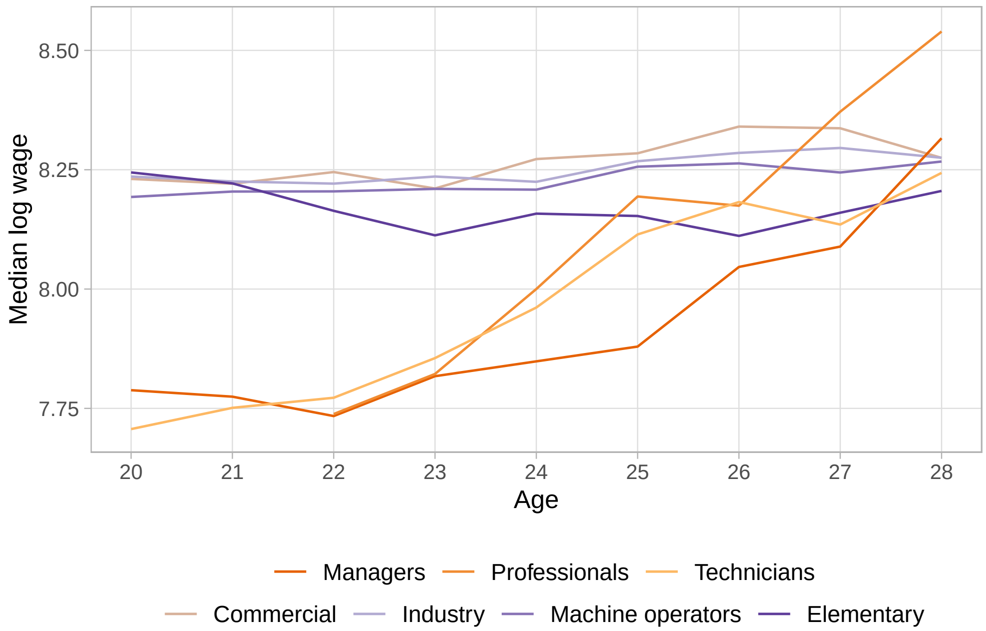
Notes: Median log wages of workers entering the labor market by age 22. Lines denote aggregate wages for workers whose first job is in respective occupation.

1.2 Young workers' wage growth across occupations

I now show that initial wages mask substantial occupational differences in subsequent wage growth. Figure 1.1 illustrates that occupations offer diverging wage trajectories to their workers. Young workers earn similar wages upon entering the labor market: for example, the median wages of office clerks and machine operators coincide at age 22. However, those who start their careers in high-skill occupations experience much faster wage growth than those in low-skill occupations. Ten years into their careers, office clerks earn 80 percent higher wages than machine operators.

Occupations offer diverging wage paths for workers at all wage levels. Figure 1.2 contrasts the observed wage paths of workers who start their careers in a low-wage job in a high-skill occupation to those of workers who start in high-wage low-skill jobs. At age 22, workers whose first job is in a low-skill occupation but pays above the median earn 58 percent higher wages than workers who start in a high-skill occupation but earning below median. However,

Figure 1.2: Initially low-paid high-skilled workers catch up



Notes: Median log wages of workers entering the labor market by age 22. Lines denote aggregate wages for workers whose first job is in respective occupation: high-skilled workers (managers, professionals, technicians) with low initial wages (below median wage at age 22), low-skilled workers (commercial, industry, machine operators, elementary) with high initial wages (above median wage at age 22).

these initial differences disappear in 3 years: high-skilled workers make up for their lower wages and quickly surpass their low-skill peers, who experience hardly no wage growth. By age 28, initially low-paid high-skilled workers earn 13 percent more on average than their low-skill peers—especially professionals whose wage premium reaches 33 percent.

Higher-skill occupations offering upward wage mobility begets the question: how can workers in low-skill occupations enter these more lucrative jobs? Upon switching into a higher-skill occupation, will they immediately earn more? Or, alternatively, would it be worthwhile for them to accept a low-paying offer in a high-skill occupation if it opens up better opportunities in the future? I attend to these questions in the remainder of this paper.

2 Model

This paper proposes a job posting model with numerous occupational employment states as well as a non-employment state. Contrary to wage posting models, firms post non-negotiable *job offers* where a job is an occupation-wage pair. From the individuals' perspective, job offers arrive at rates that are specific to their current occupation and the occupation of the job offer. Upon receiving an offer, they decide whether to accept it: their decision is governed by the cost of switching to the new job. This switching cost is stochastic, capturing the fact that individuals do not know ex ante whether they would accept a particular offer. I assume that the switching costs are logistically distributed. As a consequence, I am able to express the model in terms of the conditional choice probabilities (CCPs) — this object captures the probability that a worker will accept a particular job offer, conditional on their current job.

2.1 Notation

I ease notation by adhering to a set of guiding principles. In general, lower indices refer to origin states and upper indices refer to destination states. For occupations, I reserve the $(a, b, \ldots)$ indices and denote a generic occupation by o. For wages, I reserve the $(i, j, \ldots)$ indices and denote a generic wage rate by w. I denote jobs as an occupation–wage pair (a, i). Non-employment is denoted by N.

Putting these pieces together, the object p_{ai}^{bj}, as an example, denotes the probability of accepting a job offer in occupation b with wage j when the current job is in occupation a and pays i. Furthermore, I omit unobserved heterogeneity from the discussion of the model and its identification for the sake of clarity; I introduce types in Section 4.

2.2 Value functions

At any point in time, an individual is either employed in a job (a, i) or not currently employed. They receive job offers and decide whether they accept them; their decision is governed by an ex ante unknown cost of switching to the occupation of the offer. Employed individuals also receive job separation shocks. The rates of both offer arrivals and job separations differ across occupations.

Time is continuous. For an individual who is currently employed in job (a, i), two events may occur at any given instance. First, they may receive an offer from occupation o; these offers arrive at the rate λ_a^o which is specific to the pair of origin and destination occupations. The ex ante probability of receiving offers is integrated over all occupations that the offer may come from. Second, they may separate from their current job at the separation rate δ_a; job separations are exogenous and lead to immediate non-employment in the model. If none of these events occur, time gets discounted at the rate ρ.

The attained value is comprised of the flow utility and the option value of the current job. The flow utility $u_a(i)$ varies with the current occupation and wages.[5] The option value is the ex ante expected continuation value from the current job, i.e., the expected maximum value among attainable options, integrated over the set of possible job offers. In contrast to an exogenous job separation shock, upon which the worker is forced out of employment, accepting an offer is a choice. When workers receive an offer, they draw a switching cost $\tilde{c}_a^o$ from a known distribution with mean c_a^o and compare the value of the offered job, net of the cost of switching, to the value of staying in the current job.[6] Putting all the pieces together, the value of employment is written as

$$\left(\sum_o \lambda_a^o + \delta_a + \rho\right) V_a(i) = u_a(i) + \delta_a V_N + \mathbb{E}_{o,w,\tilde{c}} \left[\lambda_a^o \max\left\{V_o(w) - \tilde{c}_a^o, V_a(i)\right\}\right]. \qquad (2.1)$$

The value of non-employment is analogous to the above. An individual not in employment incurs the flow utility v and may receive a job offer from any occupation at rates specific to non-employment. If an offer arrives, they draw a switching cost which determines whether

[5] I discuss the structure I impose on flow utilities in Section 3.

[6] These stochastic switching costs can be equivalently expressed as deterministic costs amended by preference shocks drawn with job offers; see Appendix B.

they would accept the offer. Formally, the value function is

$$\left(\sum_o \lambda_N^o + \rho\right) V_N = v + \mathbb{E}_{o,w,\tilde{c}}\left[\lambda_N^o \max\left\{V_o(w) - \tilde{c}_N^o, V_N\right\}\right]. \tag{2.2}$$

2.3 CCP structure

I assume that switching costs are drawn from the logistic distribution with mean c_a^b. Combined with the assumption that wages are discrete with pmf $f^o(w)$, I can express the value functions in terms of conditional choice probabilities, i.e., the probability of accepting an offer conditional on receiving it:[7]

$$(\delta_a + \rho) V_a(i) = u_a(i) + \delta_a V_N - \sum_{o,w} \lambda_a^o \log(1 - p_{ai}^{ow}) f^o(w) \qquad \text{and} \tag{2.3}$$

$$\rho V_N = v - \sum_{o,w} \lambda_N^o \log(1 - p_N^{ow}) f^o(w). \tag{2.4}$$

The logistic assumption on the switching costs imply a particular structure on choice probabilities, and my identification strategy relies on this structure. Using this distributional assumption, the probability of accepting a job offer (b, j) conditional on receiving it when the current job is (a, i) can be written in terms of the value functions and mean switching costs as

$$p_{ai}^{bj} = \frac{\exp\left(V_b(j) - V_a(i) - c_a^b\right)}{1 + \exp\left(V_b(j) - V_a(i) - c_a^b\right)}. \tag{2.5}$$

Consequently, the probability of accepting (a, i) conditional on arrival when in (a, i) is

$$p_{ai}^{ai} = \frac{\exp\left(V_a(i) - V_a(i) - c_a^a\right)}{1 + \exp\left(V_a(i) - V_a(i) - c_a^a\right)} = \frac{\exp\left(-c_a^a\right)}{1 + \exp\left(-c_a^a\right)}. \tag{2.6}$$

It follows that the probability of accepting jobs from the same occupation that offers the

[7] See Appendix B for details.

same wage is constant across wage levels:

$$p_{ai}^{ai} = p_{aj}^{aj} \tag{2.7}$$

for all i and j. Section 3 discusses how I utilize this structure for identification.

3 Identification

The core idea behind identification is to match job-to-job and non-employment-to-job hazard rates to the structural parameters of the model. Here I unfold the main identification argument in a sequential fashion. The identification strategy presented here builds on Arcidiacono, Gyetvai, Jardim, and Maurel (2020) which models the search behavior of employed and unemployed individuals in a nonstationary environment. I build on this approach by bringing in various forms of occupational heterogeneity. This section demonstrates how incorporating this occupational heterogeneity extends this framework. In contrast with this companion paper, however, I do not consider nonstationarity on the non-employment-to-job side for exposition's sake: combining the two approaches would be an interesting extension.

3.1 Hazards

The identification of hazards is separate from the structural parameters of the model. I model the hazards of job-to-job, job-to-non-employment, and non-employment-to-job transitions in a competing risk model with multiple spells. Since the model does not accommodate heterogeneity in tenure on the current job, hazard rates are assumed to be constant over time; that is, I assume durations follow an exponential distribution. Identification of the hazards in this setting is well understood[8] and, thus, is not discussed here.

[8]See, e.g., Cox (1959, 1962); Tsiatis (1975).

3.2 Structural parameters

The backbone of my identification argument is that, by definition, the hazard rate of transitioning from job (a, i) to (b, j) is

$$h_{ai}^{bj} = \lambda_a^b f^b(j) p_{ai}^{bj} \tag{3.1}$$

where λ_a^b is the arrival rate of offers from b to a, f^b is the distribution of offered wages in b, and p_{ai}^{bj} is the probability of accepting an offer (b, j) when the current job is (a, i), conditional on receiving that offer.

The key identification argument is that fast offers yield more transitions at high wages while strong preferences spur churn at all wage levels. If workers know that offers from a certain occupation arrive frequently, they can rest assured that if they reject a low-wage offer now, a high-wage one will arrive soon. Conversely, if workers have strong preferences to work in a certain occupation, they would accept an offer from it at any wage level and they would not leave once they enter. Therefore, variation in the hazard rates across occupations at various wage levels separate offers from choices. I now expand on this high-level argument.

Offered wage distribution: Consider the hazard of moving from (a, i) to (a, i):

$$h_{ai}^{ai} = \lambda_a^a f^a(i) p_{ai}^{ai}. \tag{3.2}$$

As noted before in Equation 2.7, the offer-contingent probability to switch from a given job to the same job does not vary with the level of wages. Therefore, taking the ratio of the hazards of switching into the same job, the arrival rate and the switching probabilities cancel:

$$\frac{h_{ai}^{ai}}{h_{aj}^{aj}} = \frac{f^a(i)}{f^a(j)}. \tag{3.3}$$

Since the distribution of offered wages integrates to 1, it follows that[9]

$$f^a(i) = \frac{h_{ai}^{ai}}{\sum_w h_{aw}^{aw}}. \tag{3.4}$$

Note that the structure of choice probabilities allows for the identification of offered wages within, but not across, occupations. For example, wages in managerial job offers are drawn from the same distribution, regardless of whether the offer goes to technicians or machine operators. This result emerges from the fact that $p_{ai}^{bi} \neq p_{aj}^{bj}$ for $a \neq b$.

Offer arrival rates (J2J): Consider the log odds of accepting a job (b, j) while being employed in (a, i), denoted by ϖ_{ai}^{bj}:

$$\varpi_{ai}^{bj} \equiv \log\left(\frac{p_{ai}^{bj}}{1 - p_{ai}^{bj}}\right) = \log\left(\frac{h_{ai}^{bj}}{\lambda_a^b f^b(j) - h_{ai}^{bj}}\right) \tag{3.5}$$

using the fact that $p_{ai}^{bj} = h_{ai}^{bj} / \left(\lambda_a^b f^b(j)\right)$.[10] The only unknown parameter determining the odds is the offer arrival rate, as the hazard of switching jobs and the offered wage distribution have already been identified.

I start by identifying the arrival rates of offers from the same occupation as the current one, λ_a^a. Using the value of employment in Equation 2.3, the log odds of accepting an offer from the current occupation is

$$\varpi_{ai}^{aj} = V_a(j) - V_a(i) - c_a^a. \tag{3.6}$$

Consider a pair of log odds, one associated with moving across jobs that pay different wages within the same occupation and another associated with the reverse move. Taking two such

[9]Summing Equation 3.3 over wages yields $\frac{\sum_w h_{aw}^{aw}}{h_{ai}^{ai}} = \frac{\sum_w f^a(w)}{f^a(i)} = \frac{1}{f^a(i)}$.

[10]From Equation 3.1: $\frac{p_{ai}^{bj}}{1-p_{ai}^{bj}} = \frac{h_{ai}^{bj}/\left(\lambda_a^b f^b(j)\right)}{1-h_{ai}^{bj}/(\lambda_a^b f^b(j))} = \frac{h_{ai}^{bj}}{\lambda_a^b f^b(j) - h_{ai}^{bj}}$.

pairs, it holds that[11]

$$\varpi_{ai}^{aj} + \varpi_{aj}^{ai} = \varpi_{ak}^{a\ell} + \varpi_{a\ell}^{ak} = -2c_a^a. \tag{3.7}$$

The offer arrival rates from the current occupation are readily identified from Equation 3.7:

$$\lambda_a^a = \frac{\left[h_{ai}^{aj} f^a(i) + h_{aj}^{ai} f^a(j)\right] h_{ak}^{a\ell} h_{a\ell}^{ak} - \left[h_{ak}^{a\ell} f^a(k) + h_{a\ell}^{ak} f^a(\ell)\right] h_{ai}^{aj} h_{aj}^{ai}}{f^a(i) f^a(j) h_{ak}^{a\ell} h_{a\ell}^{ak} - f^a(k) f^a(\ell) h_{ai}^{aj} h_{aj}^{ai}}. \tag{3.8}$$

Now I turn to identifying the arrival rates of offers from different occupations than the current one, λ_a^b. Again, using the value of employment in Equation 2.3 the log odds of accepting an offer from the current occupation can be expressed as

$$\varpi_{ai}^{bj} = V_b(j) - V_a(i) - c_a^b. \tag{3.9}$$

Taking two such pairs of occupation-wage bins, it holds that

$$\varpi_{ai}^{bj} + \varpi_{bj}^{ai} = \varpi_{ak}^{b\ell} + \varpi_{b\ell}^{ak} = -\left(c_a^b + c_b^a\right). \tag{3.10}$$

Note that these log odds contain the offer arrival rates from a to b λ_a^b as well as the reverse λ_b^a. Therefore replicating Equation 3.10 for another two pairs of bins (a', i')–(b', j') and (a', k')–(b', ℓ') identifies both arrival rates. Appendix C provides more details.

Choice probabilities (J2J): With the offer arrival rates at hand, the choice probabilities are readily identified from the definition of hazards:

$$h_{ai}^{bj} = \lambda_a^b f^b(j) p_{ai}^{bj}. \tag{3.11}$$

Switching costs (J2J): The switching costs come from the identifying equations for the

[11] Note that $\varpi_{ai}^{ai} = \varpi_{aj}^{aj}$ also holds. However, I already extracted all the information contained in offers from the same wage bins in the identification of the wage offer distribution $f^b(j)$.

Figure 3.1: Switching cost multipliers

Skill content of destination occupation

Skill content of origin occupation	5	4	3	2	2	2	2	1
5								
4	α_4^5							
3	α_3^5	α_3^4						
2	α_2^5	α_2^4	α_2^3					
2	α_2^5	α_2^4	α_2^3	α_2^2				
2	α_2^5	α_2^4	α_2^3	α_2^2	α_2^2			
2	α_2^5	α_2^4	α_2^3	α_2^2	α_2^2	α_2^2		
1	α_1^5	α_1^4	α_1^3	α_1^2	α_1^2	α_1^2	α_1^2	

offer arrival rates. Equation 3.7 readily identifies the cost of switching to a job in the same occupation:

$$c_a^a = -\frac{\varpi_{ai}^{aj} + \varpi_{aj}^{ai}}{2}. \tag{3.12}$$

However, the equivalent equations for cross-occupation switches cannot separately identify the cost of switching from occupation a to b and the reverse, b to a. Rather, identification comes from the aggregation of occupations by skill content, observed in the data. The skill content of occupations is declining by the design of the occupation classification system, presented in Table A.1

I impose a restriction of relative symmetry: for occupations $a \neq b$,

$$c_a^b = \alpha_{s_a}^{s_b} \, c_b^a \tag{3.13}$$

where s_a is the skill content of occupation a. The α multipliers are the same for commercial, industry, and machine operators. I estimate 11 separate multipliers; Figure 3.1 displays them.

Under these restrictions, Equation 3.10 can be rewritten as

$$\varpi_{ai}^{bj} + \varpi_{bj}^{ai} = -\left(1 + \alpha_{s_a}^{s_b}\right) c_a^b. \tag{3.14}$$

The same logic extends to cycles of three job switches as well:

$$\varpi_{ai}^{bj} + \varpi_{bj}^{ck} + \varpi_{ck}^{ai} = -\left(c_a^b + c_b^c + \alpha_{s_a}^{s_c}\, c_a^c\right) \qquad \text{and} \tag{3.15}$$

$$\varpi_{bi}^{aj} + \varpi_{aj}^{ck} + \varpi_{ck}^{bi} = -\left(\alpha_{s_a}^{s_b}\, c_a^b + c_a^c + \alpha_{s_b}^{s_c}\, c_b^c\right). \tag{3.16}$$

This system of equations jointly identifies the unidirectional costs as well as the α multipliers, and, thus, the entire matrix of switching costs.

Offer arrival rates (N2E): Now I turn to structural parameters for those not in employment. Consider the log odds of accepting an offer (b, j) from non-employment:

$$\varpi_N^{bj} = V_b(j) - V_N - c_N^b. \tag{3.17}$$

Combining this log odds with the log odds of accepting an offer from the same occupation and another wage, (b, i), it holds that

$$\varpi_N^{bj} - \varpi_N^{bi} = V_b(j) - V_b(i) = \varpi_{bi}^{bj} + c_b^b. \tag{3.18}$$

Since all other parameters have been identified by now, this equation identifies λ_N^b.

Choice probabilities (N2E): Given the N2E offer arrival rates, the choice probabilities fall out from the hazard definition:

$$h_N^{bj} = \lambda_N^b f^b(j) p_N^{bj}. \tag{3.19}$$

Flow utilities (J2J, N2E) and switching costs (N2E): I identify the remaining structural parameters together in a system of linear equations. The log odds of accepting a job offer are

$$\varpi_{ai}^{bj} = \frac{1}{\delta_b+\rho}\left(u_b(j) - \sum_{o,w}\lambda_b^o \log(1-p_{bj}^{ow})f^o(w)\right) - \frac{1}{\delta_a+\rho}\left(u_a(i) - \sum_{o,w}\lambda_a^o \log(1-p_{ai}^{ow})f^o(w)\right) + \frac{\delta_b-\delta_a}{(\delta_b+\rho)(\delta_a+\rho)}\left(v - \sum_{o,w}\lambda_N^o \log(1-p_N^{ow})f^o(w)\right) - c_a^b \tag{3.20}$$

and

$$\varpi_N^{bj} = \frac{1}{\delta_b+\rho}\left(u_b(j) - \sum_{o,w}\lambda_b^o \log(1-p_{bj}^{ow})f^o(w)\right) - \frac{1}{\delta_b+\rho}\left(v - \sum_{o,w}\lambda_N^o \log(1-p_N^{ow})f^o(w)\right) - c_N^b. \tag{3.21}$$

Equations 3.20 and 3.21 are linear in the unknown parameters $u_a(i)$, $u_b(j)$, v, and c_N^b.[12] Therefore writing these equations for all possible $(a,i)-(b,j)$ combinations yields a linear system of equations. However, the system cannot identify all parameters separately; therefore, I impose additional structure on them.

I parametrize $u_a(i)$ to be an occupation-specific level shift of a common log wage profile:

$$u_a(i) = \psi_a + \beta \log w_i \tag{3.22}$$

for all a and i. This parametrization aids the interpretation of the results as I can quantify the occupational differences in wage profiles in utility and in monetary terms.

Finally, I restrict c_N^b to vary with the skill content of the occupation of the offer. That is, I flexibly estimate five separate switching cost parameters out of non-employment to occupations in each skill level.

[12] As standard in the literature, I do not estimate the discount rate ρ but rather set it to 0.05.

Putting these parametrizations together, the identifying equations become

$$\kappa_{ai}^{bj} = \frac{1}{\delta_b + \rho}\,\psi_b + \frac{\log w_j}{\delta_b + \rho}\,\beta - \frac{1}{\delta_a + \rho}\,\psi_a - \frac{\log w_i}{\delta_a + \rho}\,\beta + \frac{\delta_b - \delta_a}{(\delta_b + \rho)(\delta_a + \rho)}\,\upsilon \tag{3.23}$$

and

$$\kappa_N^{bj} = \frac{1}{\delta_b + \rho}\,\psi_b + \frac{\log w_j}{\delta_b + \rho}\,\beta - \frac{1}{\delta_b + \rho}\,\upsilon - c_N^b \tag{3.24}$$

where κ_{ai}^{bj} and κ_N^{bj} collect all known terms; i.e.,

$$\kappa_{ai}^{bj} = \varpi_{ai}^{bj} + \frac{1}{\delta_b + \rho} \sum_{o,w} \lambda_b^o \log(1 - p_{bj}^{ow}) f^o(w) - \frac{1}{\delta_a + \rho} \sum_{o,w} \lambda_a^o \log(1 - p_{ai}^{ow}) f^o(w)$$
$$+ \frac{\delta_b - \delta_a}{(\delta_b + \rho)(\delta_a + \rho)} \sum_{o,w} \lambda_N^o \log(1 - p_N^{ow}) f^o(w) + c_a^b$$

and

$$\kappa_N^{bj} = \frac{1}{\delta_b + \rho} \sum_{o,w} \lambda_b^o \log(1 - p_{bj}^{ow}) f^o(w) - \frac{1}{\delta_b + \rho} \sum_{o,w} \lambda_N^o \log(1 - p_N^{ow}) f^o(w).$$

The linear system admits an exact solution, which concludes our tour de identification.

3.3 Discussion

Like all economic models, the one presented here is a stylized depiction of reality and, as such, trades off simplicity for realism. At this point, it is worth taking a step back to discuss these tradeoffs.

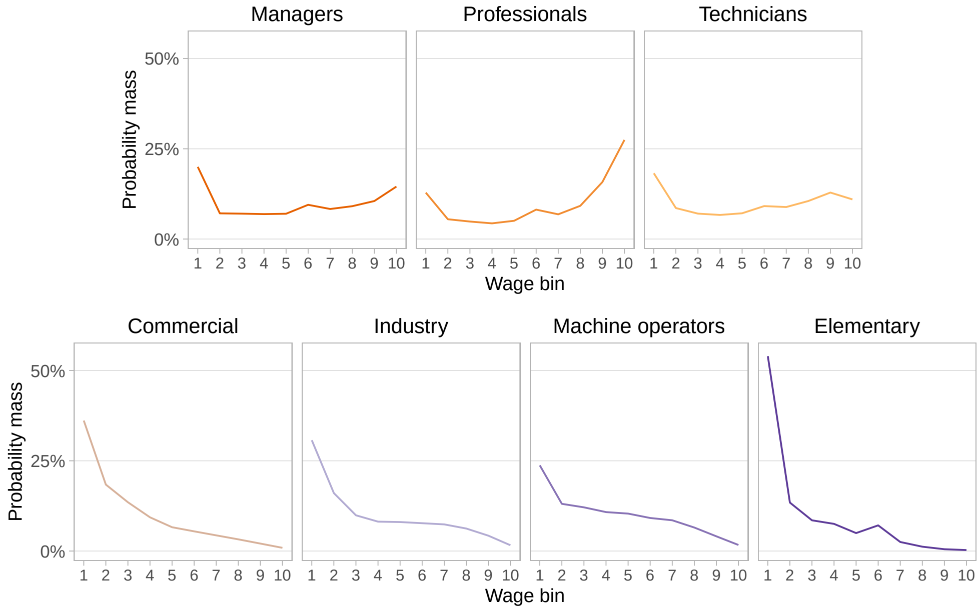
The crux of job search models is to separate "choices" from "offers"—that is, to separately identify determinants of the job search process that a worker can control from those that they cannot. In my model, offers are specific to origin and destination occupations and the wages at the destination job. Offers arrive at different rates to workers currently in different occupations; however, the offered wages are drawn from the same distribution for each destination occupation. While this feature is shared with most wage posting models (cf. Burdett and Mortensen, 1998), its implications are amplified in this setting. To the extent that workers in low-skill occupations receive high-wage offers from high-skill occupations, but they are not observed to make such transitions in the data, will seep into their preference and cost parameters.

I also remark that I omit certain sources of heterogeneity that has been in the forefront of the job search literature. Prior work has discussed the distributional implications of modeling worker-, firm-, and match-specific heterogeneity in a general equilibrium setting (Postel-Vinay and Robin, 2002; Cahuc, Postel-Vinay, and Robin, 2006; Lise, Meghir, and Robin, 2016; Bagger and Lentz, 2019). Furthermore, human capital accumulation has been shown to contribute to wage dynamics (Burdett, Carrillo-Tudela, and Coles, 2011; Bagger, Fontaine, Postel-Vinay, and Robin, 2014). Here I take a partial equilibrium approach with ex ante heterogeneity only due to occupations, and ex post heterogeneity due to the realizations of the stochastic switching costs. My goal is to assess the impact of occupations on the life cycle profiles of wage inequality; therefore I attribute these various sources of heterogeneity to occupations. Parsing the contributions to dynamic inequality of these sources is an interesting and important avenue for future research.

4 Results

I now take the model to the data. I estimate the parameters of the structural model using maximum likelihood. I impose the model structure on the hazards of job-to-job and job-to-unemployment transitions: I express the hazards as a function of the structural parameters and maximize the likelihood of these constructed hazards fitting the empirical transition patterns. I model the hazards in a competing risk framework with two-sided censoring and exponential hazards. In an additional step, I introduce unobserved heterogeneity among workers and estimate this new model using an EM algorithm (Arcidiacono and Miller, 2011).

Figure 4.1: Offered wages



Notes: Cumulative distribution functions and accepted wages in Appendix Figure E.2.

This procedure is efficient yet computationally light: the likelihood function only takes in aggregated data in each origin and destination job pair, instead of the entirety of transitions. Appendix D discusses the estimation procedure in detail.

4.1 Baseline estimates

In the model, high-skill occupations offer faster wage growth than low-skill ones. As shown on Figure 4.1, high-skill occupations offer higher wages than low-skill ones. The contrast is especially stark for professionals vs. elementary occupations: while 28 percent of professional offers come from the highest wage bin, 54 percent of offers from elementary occupations come from the lowest wage. Therefore, differential wage offers across occupations would give rise to diverging wage paths, akin to the empirical patterns. However, these forces do not impact wage trajectories in isolation: even though low-skill offers pay little, they are still rarely rejected (as shown in Appendix Figure E.2). The other mechanisms in the model shed light on the sources of this behavior.

The offer arrival rates in Table 4.1 tell two different stories about high- and low-skill

Table 4.1: Offer arrival and job separation rates

Current occupation	**Offer arrival rates** Total	Share Own	Share High	Share Low	**Job sep. rates**
Managers	2.74	17.3	25.1	74.9	0.21
Professionals	3.96	14.3	32.1	67.9	0.24
Technicians	3.04	17.1	24.5	75.5	0.25
Commercial	1.16	48.1	7.3	92.7	0.43
Industry	0.95	60.0	4.4	95.6	0.34
Machine operators	1.85	33.2	5.0	95.0	0.33
Elementary	1.20	47.9	9.7	90.3	0.99
Out of labor force	0.93	–	12.9	87.1	–

Notes: Offer arrival rates denote share of offers within a worker's current occupation, arriving from one's own/high-skill/low-skill occupations. Job separation rates are annual. All offer arrival rates in Appendix Figure E.3.

Table 4.2: Aggregate switching costs

Current occupation	**Switching costs** Own	High-skill	Low-skill
Managers	0.06	0.08	0.53
Professionals	0.06	0.05	0.70
Technicians	0.05	0.03	0.24
Commercial	0.10	0.20	1.18
Industry	0.03	1.00	0.28
Machine operators	0.04	2.07	0.48
Elementary	0.86	5.68	1.07
Out of labor force	–	0.56	0.28

Notes: Mean switching costs from one's current occupation to own/high-skill/low-skill occupations. All switching costs in Appendix Figure E.4.

Table 4.3: Flow utilities and option values

Occupation	**Flow utilities** β	ψ_a	Comp. diff.	**Option values** Mean	1–10
Managers	1.42	-1.49	0.26	4.27	1.30
Professionals		-1.83	0.21	4.97	1.31
Technicians		-1.21	0.32	4.87	1.28
Commercial		0.72	1.26	6.73	1.22
Industry		0.84	1.36	5.39	1.22
Machine operators		0.40	1.00	5.43	1.23
Elementary		5.07	26.52	11.52	1.15
Out of labor force		–	–	5.23	–

Notes: Compensating differentials: willingness-to-pay to switch to machine operating, relative to current wage. Mean option values: average across wage bins by occupations, relative to flow utilities. 1–10: change in option value from lowest to highest wage bin.

occupations. Workers in high-skill occupations receive 2.74–3.96 offers annually, while those in low-skill occupations receive about one less offer, ranging from 0.95–1.85 offers per year. Overall, offers come from low-skill occupations much more frequently than from high-skill ones. This pattern is shared by current high- and low-skilled workers as well as individuals out of the labor force. However, the differences are smaller for current high-skilled workers: they receive high-skill offers with a five time higher probability than current low-skilled workers (24.5–32.1 vs. 4.4–9.0 percent) and twice as frequently as those who are not working (24.5–32.1 vs. 12.9 percent). Low-skilled workers are more likely to receive an offer from the occupation they are currently working in (33.2–60.0 vs. 14.3–17.3 percent); however, conditional on offers from occupations in one's current skill level, offers are less likely to come from their own occupation. These patterns may emerge as a result of the transferability of skills across high- and low-skill occupations. On one hand, current high-skilled workers likely possess the necessary skills for high-skill occupations, thus they get high-skill offers more frequently—but they also get low-skill offers relatively frequently as they can be a good match there as well. On the other hand, current low-skilled workers rarely get the opportunity to switch to high-skill occupations.

The job separation rates in the last column of Table 4.1 worsen the outlooks of low-skilled workers. These jobs are less stable: workers in low-skill occupations separate to non-employment at much faster rates than those in high-skill ones. Elementary jobs are especially fragile: these workers lose their job once a year while only one in four managers separates to non-employment in a given year. These elementary occupations, such as janitors and helpers, substitute unemployment: they require no formal qualifications, pay little, and are typically short-term arrangements. Therefore, high-skill occupations are valuable not only due to their faster wage growth but also their higher retention rates.

Let's turn to what the data tell us about the behavior of workers once they get a job offer. Upon the offer's arrival, they learn the cost of switching from their current job to the new one, and decide whether they accept the offer. Table 4.2 displays aggregate measures of these cost realizations. Overall, the least costly switch is within one's current occupation, more costly to switch to another occupation at a similar skill level, and most costly to switch to occupations at another skill level. Moving to another skill level is associated with a five times higher cost than staying at one's current skill level. Once again, high- and low-skill occupations paint two different pictures: "downskilling" for high-skilled workers is less costly than "upskilling" for low-skilled workers. High-skilled workers have to pay a cost in the range of 0.03–0.08 to switch to another high-skill job and costs of 0.24–0.70 to switch to low-skill jobs. Low-skilled workers face much higher costs in both directions: 0.28–1.18 for another low-skill job and 0.20–5.68 for transitions to a high-skill one. These large costs help explain why low-skilled workers reject high-skill offers at high rates. Separating to non-employment is of no help to ease these transitions either: moving into a high-skill occupation from non-employment demands a cost of 0.56, while it is a third of that, 0.28, for low-skill occupations.

Finally, I discuss the value of jobs in high- vs. low-skill occupations. This value is comprised of the flow utility that workers instantaneously incur, as well as the option value of being in a certain job. Looking at the flow utilities in Table 4.3, low-skill jobs yield higher utility than high-skill ones for the same wage. I translate these differences into compensating differentials: the third column of the table shows how much a worker in a certain occupation would have to be compensated to become a machine operator. That is, the compensating differential for a worker in occupation a making wage $\bar{w}_a$), denoted by w_a^{MO}, satisfies the equation

$$\psi_a + \beta \log \bar{w}_a = \psi_{\mathrm{MO}} + \beta \log w_a^{\mathrm{MO}}. \tag{4.1}$$

Current high-skilled workers would accept a machine operating job for a fraction of their current pay, while current low-skilled workers would need much higher wage offers. For instance, professionals would take that job for 21 percent of their current wage but industry workers would need 36 percent higher pay. The results are especially stark for elementary workers who would accept an offer that pays 27 times more than their current job. These estimates likely reflect the occupational wage distribution: elementary occupations pay very little compared to professional ones.

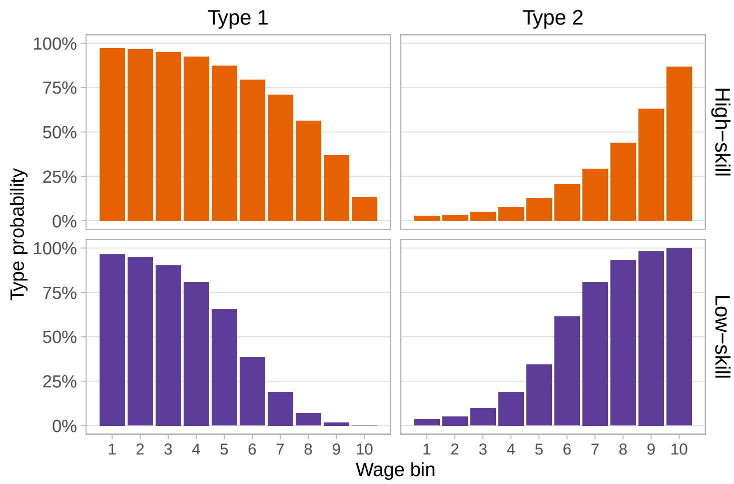
The value of a certain job is not only in how much it pays but also what future wages it affords to its workers. The option value in the last two columns of Table 4.3 quantifies these values. They are determined by two competing forces. On one hand, current low-skilled workers likely stay in other low-skill jobs which have higher flow utilities, resulting in a higher option value. On the other hand, they will not experience wage growth in these low-skill occupations, which brings the option value down. The first effect trumps the second, resulting in higher option values for low-skill occupations. Nonetheless, the wage growth channel is also substantial: the last column shows the difference in the option values of jobs in the lowest vs. highest wage bin in each occupation. Low-paying high-skill jobs have a 30 percent higher option value than high-wage high-skill jobs, precisely because workers can climb high-skill ladders faster. Low-wage jobs in low-skill occupations, however, are worth only 15–20 percent more than high-paying ones, as they have less rosy future opportunities.

4.2 Adding unobserved heterogeneity

I now introduce unobserved worker-specific heterogeneity to the model. That is, I distribute workers to two heterogeneity types with some probability π and estimate this distribution, along with all the type-specific structural parameters; see Appendix D for details about this estimation procedure.

The estimated types pick up differences in potential experience across workers. Type 1 workers, making up 64.3 percent of the sample, are 4.4 percent younger on average than Type 2 workers that make up 35.7 percent. Figure 4.2 breaks the overall figures down by wage bins and shows that types also separate low and high earners: workers in the lowest wage bin almost surely belong to Type 1 while those in the highest bin are Type 2. The contrast is starker for low-skill occupations: their highest-wage workers belong to Type 2

Figure 4.2: Aggregate type probabilities



Notes: Type probabilities by occupations in Appendix Figure E.6.

Figure 4.3: Offered wages

Notes: Type 1: solid lines. Type 2: dashed lines. Cumulative distribution functions and accepted wages in Appendix Figure E.8.

Table 4.4: Offer arrival and job separation rates with unobserved heterogeneity

Current occupation	**Offer arrival rates**								**Job sep. rates**	
	Type 1				*Type 2*					
	Total	Share			Total	Share			*Type 1*	*Type 2*
		Own	High	Low		Own	High	Low		
Managers	3.19	16.6	24.5	75.5	2.22	17.9	25.7	74.3	0.26	0.16
Professionals	4.63	13.7	31.5	68.5	3.20	15.0	33.1	66.9	0.29	0.17
Technicians	3.51	15.9	23.4	76.6	2.46	17.9	25.4	74.6	0.30	0.19
Commercial	1.29	46.0	7.7	92.3	0.96	48.8	7.2	92.8	0.54	0.29
Industry	0.98	55.1	5.0	95.0	0.89	65.3	3.9	96.1	0.44	0.23
Machine operators	2.12	31.6	5.2	94.8	1.54	34.9	4.8	95.2	0.41	0.24
Elementary	1.34	46.0	10.3	89.7	0.99	48.3	9.5	90.5	1.34	0.72
Out of labor force	0.85	–	12.2	87.8	1.06	–	13.4	86.6	–	–

Notes: Offer arrival rates: totals are annual rates; shares denote percentages of offers arriving from one's own/high-skill/low-skill occupations. Job separation rates: annual rates. All offer arrival rates in Appendix Figure E.9.

Table 4.5: Aggregate switching costs with unobserved heterogeneity

Current occupation	**Switching costs**					
	Type 1			*Type 2*		
	Own	High-skill	Low-skill	Own	High-skill	Low-skill
Managers	0.05	0.06	0.43	0.07	0.10	0.63
Professionals	0.05	0.04	0.60	0.08	0.06	0.83
Technicians	0.04	0.02	0.20	0.07	0.03	0.30
Commercial	0.08	0.17	1.04	0.12	0.24	1.38
Industry	0.02	0.85	0.24	0.03	1.16	0.34
Machine operators	0.03	1.86	0.41	0.05	2.37	0.57
Elementary	0.73	5.06	1.01	1.00	6.68	1.19
Out of labor force	–	0.67	0.34	–	0.46	0.23

Notes: Mean switching costs from one's current occupation to own/high-skill/low-skill occupations. All switching costs in Appendix Figure E.11.

Table 4.6: Flow utilities and option values with unobserved heterogeneity

Occupation	**Flow utilities**						**Option values**			
	Type 1			*Type 2*			*Type 1*		*Type 2*	
	β	ψ_a	Comp. diff.	β	ψ_a	Comp. diff.	Mean	1–10	Mean	1–10
Managers	1.14	-1.78	0.14	1.46	-1.14	0.37	5.65	1.32	3.31	1.30
Professionals		-2.18	0.10		-1.43	0.30	6.79	1.33	3.80	1.31
Technicians		-1.42	0.19		-0.92	0.43	6.17	1.30	3.83	1.29
Commercial		0.80	1.34		0.55	1.18	8.31	1.21	5.16	1.23
Industry		0.97	1.56		0.65	1.26	6.74	1.21	4.09	1.24
Machine operators		0.46	1.00		0.31	1.00	6.64	1.23	4.31	1.24
Elementary		6.06	136.20		4.02	12.58	13.59	1.13	9.86	1.17
Out of the labor force		–	–		–	–	3.25	–	7.92	–

Notes: Compensating differentials: willingness-to-pay to switch to machine operating, relative to current wage. Mean option values: average across wage bins by occupations, relative to flow utilities. 1–10: change in option value from lowest to highest wage bin.

with a 99.8 percent probability. These types apply to an uneven sample: while most Type 1 high-skilled workers earn below-median wages, they only make up 6 percent of the sample, compared to 59 percent of low-wage low-skilled workers. The contrast is even starker for low-skilled workers earning high wages: while they almost certainly belong to Type 2, they represent only 2 percent of the sample.

Turning to wage offers, Figure 4.3 attests that Type 2 workers receive higher wage offers.[13] Type 2 high-skill occupations receive a lot more offers from high wages, especially professionals at 62 percent. Workers in low-skill occupations receive almost no offers from high wages, but offers to Type 2 workers bring about slightly higher pay than for Type 1 workers.

Experience affects the rates at which workers receive offers and separate from their jobs. Type 1 workers receive 35 percent more offers in both low- and high-skill occupations than Type 2 workers, as Table 4.4 demonstrates. However, a larger share of offers come from Type 2 workers' current occupation and other occupations of a similar skill content. At the same time, Type 1 workers separate more often from their jobs while Type 2 workers face more stable employment. These patterns suggest that Type 2 workers are more securely attached to the labor market but are also locked in to their current occupations, relative to Type 1 workers.

This lock-in pattern continues with the costs of switching. Type 2 workers face higher costs across the board, as shown in Table 4.5. Otherwise, the patterns are similar to those without unobserved heterogeneity: the least costly transitions are within one's current occupation, and transitions to occupations at another skill level are the most costly ones. Current high-skilled workers face lower costs to switch to low-skill occupations than the reverse. Within low-skill occupations, Type 2 workers face costs of 0.24–6.68 to switch to high-skill occupations, while Type 1 workers face 50 percent lower costs, ranging 0.20–0.60.

[13] Even though Section 3 demonstrates that offered wages are flexibly identified, I impose additional parametrizations across types in the estimation. I model flexible wage offers for one type and shift them by a type-specific logit multiplier for the remaining types. See Appendix D for details.

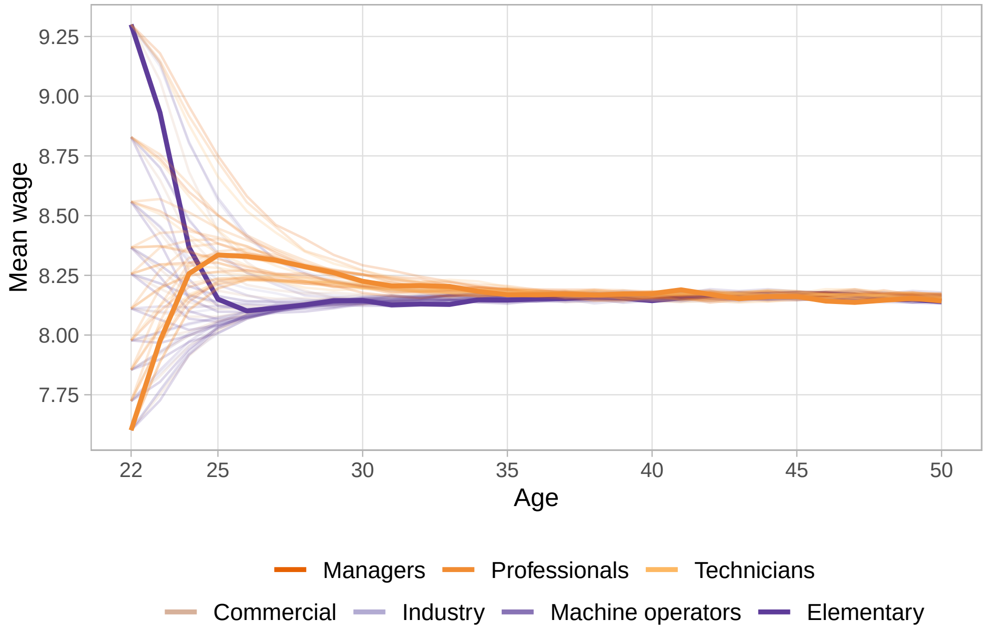
Preferences also exhibit this heterogeneity, as Table 4.6 illustrates. On one hand, Type 2 workers value wages more than Type 1, 1.46 compared to 1.14. On the other hand, Type 1 workers incur 50 percent larger non-wage amenities than Type 2 on average. Taken in tandem, the latter effect trumps the former: compensating differentials are more substantial for Type 1 than for Type 2 workers in absolute terms. Type 1 high-skilled workers would accept a machine operator offer for 10–19 percent of their current wage while this amount is 20 percentage points higher for Type 2 high-skilled workers. Similarly, Type 1 low-skilled workers would need much higher compensation than Type 2; e.g., Type 1 commercial workers require 14 percentage points higher wages than the 18 percent amount for Type 2 commercial workers.

A final piece demonstrating the effects of experience on the results is the option values resulting from these estimates. Type 1 workers' option values range between 5.65–6.79 times their flow utility in high-skill occupations and 6.64–13.59 in low-skill ones, whereas these figures are substantially lower for Type 2 workers, 3.31–3.80 in high-skill and 4.09–9.86 in low-skill occupations. Curiously, the option values of high-skill jobs for Type 1 are 6 percent higher on average than for low-skill ones for Type 2 (5.65–6.79 compared to 4.09–9.86), further pointing at the value of high-skill occupations for early-career workers. Furthermore, the option value gains from going from the lowest to the highest wages are larger for Type 1 high-skill occupations but lower for low-skill jobs. These findings suggests that in early stages of one's career, being in a low-paying high-skill job is more beneficial; but as time goes by, being locked in to a low-skill occupation yields a higher relative earnings potential.

5 Explaining Life Cycle Wage Inequality

With the structural estimates at hand, this final section relates occupational mobility to the life cycle profile of wage inequality. I simulate careers starting from the empirical distribution of initial jobs, and compare wages at various points of the life cycle. I show that workers

Figure 5.1: Ex ante wage profiles



Notes: Simulated wage profiles from baseline hazard estimates. Mean log wages by age across 2,500 simulations for each initial occupation-wage pair. Thick orange line: ex ante wage profile of a worker starting in a professional job in the lowest wage bin. Thick purple line: ex ante wage profile of a worker starting in an elementary job in the highest wage bin.

who start their careers in high-skill jobs have higher ex ante expected wages over their lifetime than low-skilled workers, regardless of their initial wage. Then I demonstrate that this occupational heterogeneity explains 96 percent of the empirical life cycle profile of wage inequality. Remarkably, the occupationally heterogeneous model fits the rapidly increasing wage inequality immediately upon entering the labor market. I find that heterogeneous offer arrival and job separation rates, flow utilities, and switching costs contribute equally to the explanatory power of occupations.

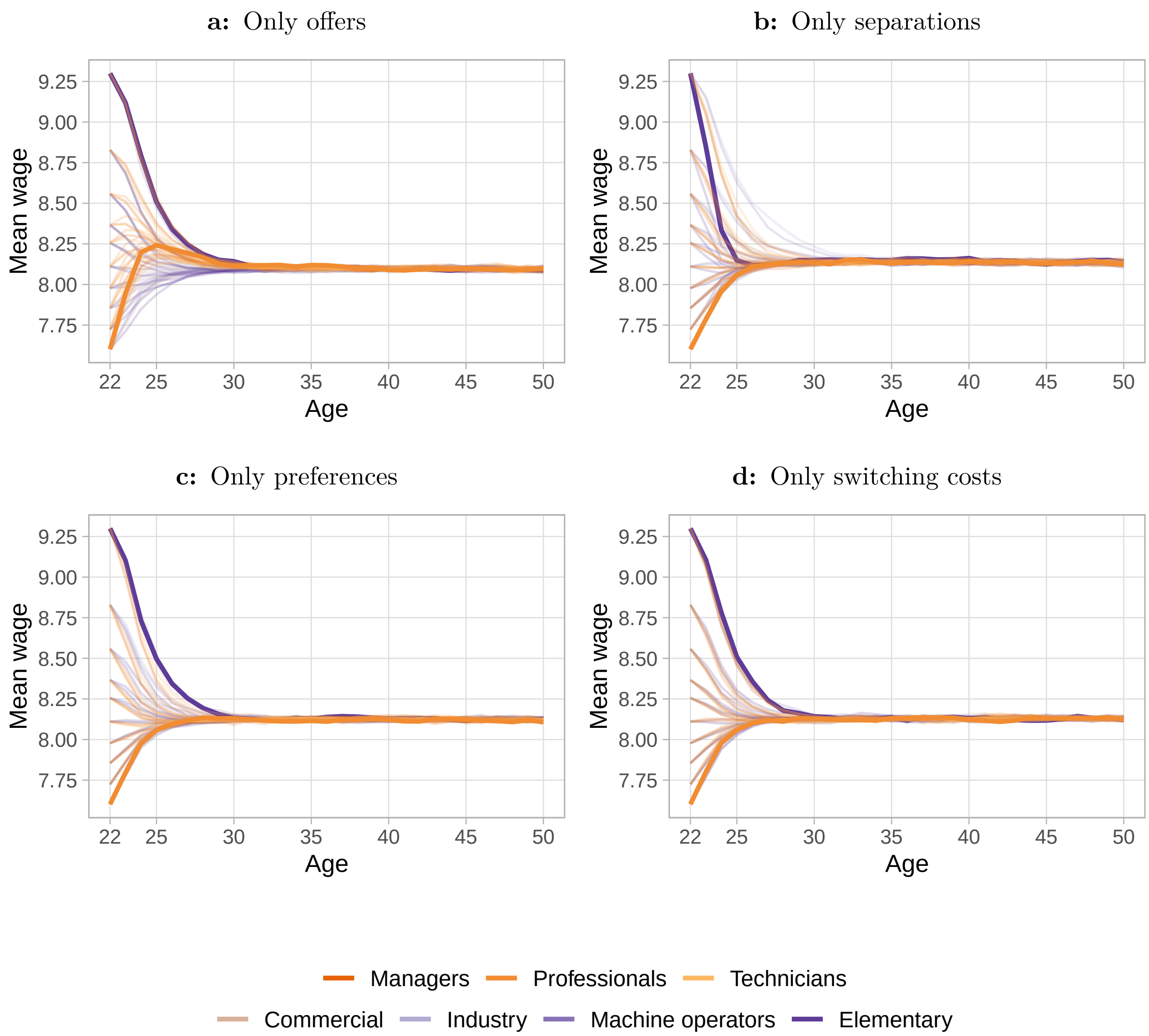
I start by demonstrating that occupations offer ex ante differing wage profiles, mirroring the descriptive evidence presented in Section 1. I simulate a large number of workers starting their careers at all possible wage levels in each occupation. Then I draw job spells from the competing hazard estimates and piece them together, thus creating lifetime careers. I record the occupational transitions and wage paths along these simulated careers. Finally, I compile ex ante wage profiles by averaging out these paths at each point of the life cycle.

Figure 5.1 shows these mean wage profiles by initial jobs in high- vs. low-skill occupations at all wage levels. Specifically, the thick orange line tracks the ex ante expected wages of a 22 year old worker starting their career in a low-wage professional job. Similarly, the purple line follows the ex ante wage evolution of a 22 year old starting in a high-wage elementary job. Within three years, by age 25, the (initially) professional worker already expects to earn higher wages. The (initially) elementary worker sharply loses their initial wage advantage: in expectation, they separate from their job and never catch up. These patterns are not unique to these two hypothetical workers: regardless of their initial wages, high-skilled workers (orange lines) always converge to the eventual steady-state mean from above while low-skilled workers (purple lines) converge from below.

Which source of heterogeneity drives these patterns? I answer this question by repeating the above simulation for a number of auxiliary models that incorporate only one heterogeneity source at a time. I estimate a variant of the baseline model that features only heterogeneous offer arrival rates and wage offer distributions. A second variant features only heterogeneous job separation rates. A third version incorporates only occupation-specific non-wage amenities. A final fourth version allows only costs of switching to vary across workers' current and offered occupations.

Figure 5.2 suggests that all of these channels contribute to the overall patterns. Occupationally heterogeneous offers generate a rapid increase in the ex ante wage profile for high-skilled workers, especially for low-wage ones as the orange line on the top left panel shows. At the same time, initially low-skilled workers do not lose their advantage quite as quickly as in the overall results: that sharp drop is instead generated by job separations, as the top right panel demonstrates. However, none of these two sources of heterogeneity creates the level of dispersion among workers who start making the same wage. That dispersion comes from non-wage amenities: the bottom left panel shows that low-skilled workers hold on to their jobs, conditional on current wages, thus they lose their wage advantage

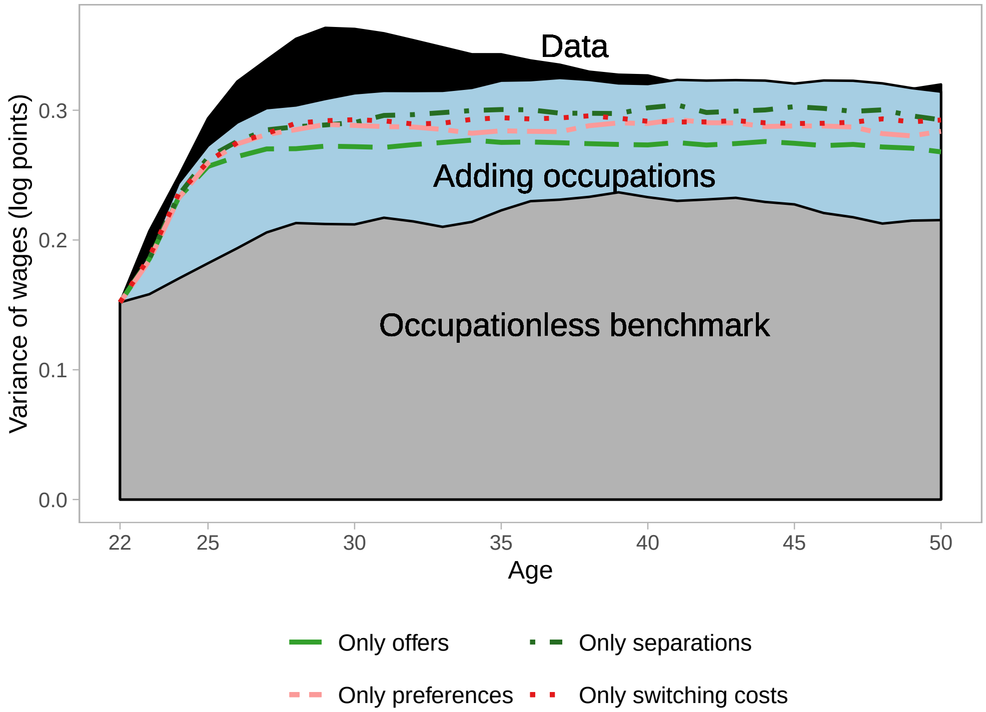
Figure 5.2: Ex ante wage profiles by heterogeneity source



Notes: Simulated wage profiles from counterfactual hazard estimates. Mean log wages by age across 2,500 simulations for each initial occupation-wage pair. Thick orange line: ex ante wage profile of a worker starting in a professional job in the lowest wage bin. Thick purple line: ex ante wage profile of a worker starting in an elementary job in the highest wage bin.

slower than high-skilled workers. Finally, heterogeneous switching costs alone do not seem to cause a big difference. Keep in mind, however, that these ex ante profiles capture only aggregate wage dynamics across groups but they do not account for dispersion among individual wages. Switching costs serve exactly this purpose in the model: they create differing acceptance behavior among identical workers. As I turn to discussing wage inequality, I will

Figure 5.3: Life cycle profiles of wage inequality



Notes: Variance of log wages by age across simulated careers from empirical distribution of initial jobs. See Appendix Table E.1 for measures of model fit.

show that occupationally heterogeneous costs are instrumental to account for the empirical wage dispersion.

Next, I link these diverging wage profiles to the life cycle profile of wage inequality. I repeat the simulation procedure outlined above, but instead of a hypothetical starting distribution, I start from the empirical distribution of jobs at age 22. Then, just like before, I draw and piece together job spells, thus creating lifetime careers from age 22 to 50. I record the occupational transitions and wage paths along these simulated careers. Finally, I calculate the within-group variance across paths at each age, by initial jobs. I repeat this exercise for estimates of the four restricted models, as well as an occupationless benchmark which shuts down all sources of occupational heterogeneity.

Figure 5.3 contrasts the life cycle profiles of wage inequality in these alternative model specifications to the empirical life cycle wage dispersion. Upon entering the labor market, the inequality of wages sharply rises in the data (black ribbon). The full model in light blue closely matches this explosive increase. After that, the empirical dispersion reaches its peak and smoothly arrives at its plateau of 0.32 log points. The inequality in the full model does not climb as high but it eventually plateaus at the same level. In contrast, the occupationless benchmark features only a modest increase at the beginning of the life cycle and eventually flattens out at a much lower level, 0.22 log points. Overall, inequality in the full model matches 96 percent of the empirical dispersion while that of the occupationless benchmark fits only 67 percent of the variation.

Why does the full model perform so well? Which added mechanism enhances its performance most: occupation-specific offers, separations, preferences, or switching costs? The dashed, colored lines in Figure 5.3 show the model fit in each of these four scenarios. All four sources of heterogeneity widen the life cycle inequality of wages. Occupation-specific offers bring about an additional 26 percent of variation, fitting 84 percent of the observed dispersion on average. Occupational job separation rates do even better, bringing the model fit up to 90 percent. Preferences fit 87 percent of the variation, while costs can account for 88 percent. Remarkably, each of these occupational sources fit the data well in isolation, both on average terms and the dynamics of life cycle inequality. When incorporating more than one at a time, the interplay between these sources would provide a fit between the measure in isolation and the overall figures.

This variance decomposition is similar in spirit to Taber and Vejlin (2020). The notable difference, however, is that here each source of heterogeneity is specific to workers' current occupations and the occupations of the job offers they receive. My results are not comparable to those from their approach for two key reasons. First, they estimate a baseline specification where wage inequality results from heterogeneous skills prior to entering the labor market, firm type-specific offer arrivals, non-wage amenities, and human capital accumulation. The model presented here shares some but not all of these features, incorporates other mechanisms, and throughout all of that focuses on occupational heterogeneity. Second, they follow a top-down approach, in that they start from the full model and shut down channels one by one. I follow a bottom-up approach instead: since each of the model mechanisms fit the data remarkably well on their own, this bottom-up approach highlights the power of occupations in fitting the empirical patterns.

6 Conclusion

This paper argues that occupations play a key role in explaining the life cycle profile of wage inequality. I unfold this argument in four steps. First, I document that high-skill occupations offer steep wage growth to early-career workers while low-skill occupations feature flat wage profiles. Second, I build a structural model of mobility in which workers can switch occupations to earn higher wages if they have the opportunity. Third, I estimate the model using Hungarian administrative data that follow workers' occupational histories and I uncover substantial occupational heterogeneity in the determinants of occupational job flows. Fourth, I show that this occupational heterogeneity fully explains the inequality of wages over the life cycle.

I gain insights from a model of job transitions which flexibly captures occupation-specific offer arrival rates, job separation rates, wage offers, non-wage amenities, and switching costs. In the model, workers are employed in one of numerous occupations and incur utility from their wage and non-wage amenities. At any given instance, they may receive a job offer from another occupation and wage level; upon the offer arrival, they decide whether to accept it. Their decision is shaped by comparing the value of their current job to the counterfactual value they would incur, should they accept the offer, as well as a net amenity cost associated with the act of switching jobs.

To estimate my structural model, I express its parameters in terms of the probability of accepting a job offer from any occupation that pays a certain wage, conditional on the current occupation and wage of the worker. I estimate the structural model on a rich data set compiled from Hungarian administrative records. These data have several main advantages over other, commonly used datasets; most importantly, they approximate the date of employer and occupational transitions and they contain administrative information on workers' occupations.

Taking my model to these rich data, I document several findings. I find that occupations that are associated with high skill levels (such as professionals or technicians) offer higher wages than low-skill occupations (e.g., machine operators and industry occupations). Job offers from high-skill occupations arrive more frequently than low-skill offers, and they offer higher wages. Yet, workers in low-skill occupations turn down these offers as it is costly to switch from a low-skill occupation to a high-skill one. Furthermore, compensating differentials for workers in low-skill occupations are substantially larger than for those in high-skill occupations.

I conclude the paper by linking the structural estimates to dynamic wage inequality. Simulating workers' careers from the empirical distribution of initial jobs at age 22, I find that workers who start out in low-paid high-skill occupations quickly surpass their peers who start out in high-paid low-skill jobs. On aggregate, occupational transitions explain 94 percent of the empirical life cycle profile of wage inequality.

References

Abowd, J. M., F. Kramarz, and D. N. Margolis (1999). High Wage Workers and High Wage Firms. *Econometrica 67*(2), 251–333.

Antonovics, K. and L. Golan (2012). Experimentation and Job Choice. *Journal of Labor Economics 30*(2), 333–366.

Arcidiacono, P., A. Gyetvai, E. Jardim, and A. Maurel (2020). Conditional Choice Probability Estimation of Continuous-Time Job Search Models. Working Paper.

Arcidiacono, P., V. J. Hotz, A. Maurel, and T. Romano (2020). *Ex Ante* Returns and Occupational Choice. *Journal of Political Economy*, forthcoming.

Arcidiacono, P. and R. A. Miller (2011). Conditional Choice Probability Estimation of Dynamic Discrete Choice Models with Unobserved Heterogeneity. *Econometrica 79*(6), 1823–1867.

Bagger, J., F. Fontaine, F. Postel-Vinay, and J.-M. Robin (2014). Tenure, Experience, Human Capital, and Wages: A Tractable Equilibrium Search Model of Wage Dynamics. *American Economic Review 104*(6), 1551–1596.

Bagger, J. and R. Lentz (2019). An Empirical Model of Wage Dispersion with Sorting. *Review of Economic Studies 86*(153-190).

Bonhomme, S., T. Lamadon, and E. Manresa (2019). A Distributional Framework for Matched Employer Employee Data. *Econometrica 87*(3), 699–739.

Burdett, K., C. Carrillo-Tudela, and M. G. Coles (2011). Human capital accumulation and labor market equilibrium. *International Economic Review 52*(3), 657–677.

Burdett, K. and D. T. Mortensen (1998). Wage Differentials, Employer Size, and Unemployment. *International Economic Review 39*(2), 257–273.

Busch, C. (2017). Occupational Switching and Wage Risk. Working Paper.

Cahuc, P., F. Postel-Vinay, and J.-M. Robin (2006). Wage Bargaining with On-the-job Search: Theory and Evidence. *Econometrica 74*(2), 323–364.

Card, D., A. R. Cardoso, J. Heining, and P. Kline (2018). Firms and Labor Market Inequality: Evidence and Some Theory. *Journal of Labor Economics 36*(S1), S13–S70.

Card, D., J. Heining, and P. Kline (2013). Workplace Heterogeneity and the Rise of West German Wage Inequality. *Quarterly Journal of Economics 128*(3), 967–1015.

Carrillo-Tudela, C. and L. Visschers (2020). Unemployment and Endogenous Reallocation over the Business Cycle. IZA Discussion Paper 13307.

Cox, D. R. (1959). The Analysis of Exponentially Distributed Life-Times with Two Types of Failure. *Journal of the Royal Statistical Society 21*(2), 411–421.

Cox, D. R. (1962). *Renewal Theory.* Methuen and Co.

DellaVigna, S., A. Lindner, B. Reizer, and J. F. Schmieder (2017). Reference-dependent Job Search: Evidence from Hungary. *Quarterly Journal of Economics 4*, 1969–2018.

Groes, F., P. Kircher, and I. Manovskii (2015). The U-Shapes of Occupational Mobility. *Review of Economic Studies 82*, 659–692.

Jarosch, G. (2015). Searching for Job Security and the Consequences of Job Loss. Working Paper.

Jolivet, G., F. Postel-Vinay, and J.-M. Robin (2006). The Empirical Content of the Job Search Model: Labor Mobility and Wage Distributions in Europe and in the US. *European Economic Review 50*, 877–907.

Kambourov, G. and I. Manovskii (2009a). Occupational Mobility and Wage Inequality. *Review of Economic Studies 76*(2), 731–759.

Kambourov, G. and I. Manovskii (2009b). Occupational Specificity of Human Capital. *International Economic Review 50*(1), 63–115.

Keane, M. P. and K. I. Wolpin (1997). The Career Decisions of Young Men. *Journal of Political Economy 105*(3), 473–522.

Lise, J., C. Meghir, and J.-M. Robin (2016). Matching, Sorting and Wages. *Review of Economic Dynamics 19*, 63–87.

McCall, B. P. (1990). Occupational Matching: A Test of Sorts. *Journal of Political Economy 98*(1), 45–69.

Miller, R. A. (1984). Job Matching and Occupational Choice. *Journal of Political Economy 92*(6), 1086–1120.

Neal, D. (1999). The Complexity of Job Mobility among Young Men. *Journal of Labor Economics 17*(2), 237–261.

Pilossoph, L. (2014). Sectoral Shocks and Move Unemployment. Working Paper.

Postel-Vinay, F. and J.-M. Robin (2002). Equilibrium Wage Dispersion with Worker and Employer Heterogeneity. *Econometrica 70*(6), 2295–2350.

Rosen, S. (1986). The Theory of Equalizing Differences. In *Handbook of Labor Economics*, Volume 1, Chapter 12, pp. 641–692.

Siow, A. (1984). Occupational Choice Under Uncertainty. *Econometrica 52*(3).

Sorkin, I. (2018). Ranking Firms Using Revealed Preference. *Quarterly Journal of Economics*, 1331–1393.

Sullivan, P. (2010). A Dynamic Analysis of Educational Attainment, Occupational Choices, and Job Search. *International Economic Review 51*(1), 289–317.

Sullivan, P. and T. To (2014). Search and non-wage Job Characteristics. *Journal of Human Resources 49*(2), 472–507.

Taber, C. and R. Vejlin (2020). Estimation of a Roy/Search/Compensating Differential Model of the Labor Market. *Econometrica 88*(3), 1031–1069.

Tsiatis, A. (1975). A Nonidentifiability Aspect of the Problem of Competing Risks. *Proceedings of the National Academy of Sciences 72*(1), 20–22.

Wiczer, D. (2015). Long-term Unemployment: Attached and Mismatched? Federal Bank of St. Louis Research Division Working Paper 2015-042A.

Appendices

A Data

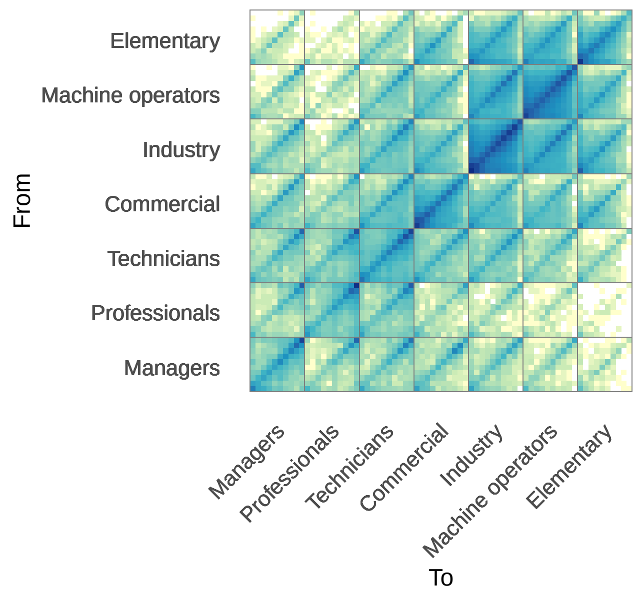
Table A.1: Skill content of occupations

Occupation	Skill level	Educational requirement
Managers	5	College or post-secondary specialist ed.
Professionals	4	College
Technicians	3	Post-secondary specialist ed.
Commercial	2	Secondary ed.
Industry	2	Secondary ed.
Machine operators	2	Secondary ed.
Elementary	1	No formal qualifications

Job-to-job transition matrix

Each block represents transitions across occupations and, due to the occupation classification system, the skill content of occupations is loosely decreasing left-to-right and bottom-to-top. Furthermore, wages within each block are increasing left-to-right and bottom-to-top; therefore, within blocks the diagonal represents transitions which do not involve a wage change, above the diagonal are wage cuts and under it are wage bumps. The overwhelming majority of occupation blocks contain at least one transition; the further the jump in terms of occupational skill content, the more likely empty cells occur. Many transitions happen at the same occupation and a similar wage rate, implying only a firm switch. Furthermore, conditional on an occupation switch (i.e., in off-diagonal blocks) wages are more likely to stay at their previous level.

Figure A.1: Number of observations by transition cells



Notes: Log number of transitions in each origin and destination occupation-wage pair. Within each occupation-to-occupation block, wages are increasing left-to-right and bottom-to-top. White cells are empty.

B Model

B.1 Stochastic switching costs interpreted as preference shocks

I assume that switching costs are ex ante unknown and drawn from a known distribution upon the arrival of job offers, which underlies the dispersion in offer acceptance even among observably equivalent workers. From the model's perspective, this assumption is equivalent to having fixed switching costs and a preference shock associated with the offer. To see this, consider rewriting Equation 2.1 as

$$\left(\sum_o \lambda_a^o + \delta_a + \rho\right) V_a(i) = u_a(i) + \delta_a V_u + \sum_o \lambda_a^o \, \mathbb{E}_{w,\varepsilon} \max\left[V_o(w) - c_a^o + \varepsilon, V_a(i)\right]. \tag{B.1}$$

Here ε is a preference shock, drawn when a job offer arrives. Assuming that ε is logistically distributed with mean zero is equivalent to the distributional assumption above.

In terms of interpretation, stochastic switching costs depict ex ante uncertainty regarding amenities associated with transitioning from occupation a to b. A negative cost realization may be interpreted as a desirable amenity the worker is willing to pay for. On the other hand, preference shocks may be interpreted as ex ante unknown non-pecuniary aspects of the job offer. These seemingly competing interpretations are, in fact, interchangeable.

B.2 Value functions expressed in terms of CCPs

From Equation 2.1,

$$\begin{aligned}
\left(\sum_o \lambda_a^o + \delta_a + \rho\right) V_a(i) &= u_a(i) + \delta_a V_u + \sum_o \lambda_a^o \, \mathbb{E}_{w,\tilde{c}} \max\left[V_o(w) - \tilde{c}_a^o, V_a(i)\right] \\
&= u_a(i) + \delta_a V_u + \sum_o \lambda_a^o \, \mathbb{E}_{w,\tilde{c}} \left\{V_a(i) + \max\left[V_o(w) - V_a(i) - \tilde{c}_a^o, 0\right]\right\} \\
\left(\delta_a + \rho\right) V_a(i) &= u_a(i) + \delta_a V_u + \sum_o \lambda_a^o \, \mathbb{E}_{w,\tilde{c}} \max\left[V_o(w) - V_a(i) - \tilde{c}_a^o, 0\right]
\end{aligned}$$

Since $\tilde{c}_a^o \sim \text{Logistic}(c_a^o)$, the Emax term can be expressed as

$$\mathbb{E}_{w,\tilde{c}} \max\left[V_o(w) - V_a(i) - \tilde{c}_a^o, 0\right] = \mathbb{E}_w \left\{\log\left(1 + \exp(V_o(w) - V_a(i) - c_a^o)\right)\right\}.$$

Now note that the log term looks eerily similar to the probability to reject an offer (c.f. Equation 2.5):

$$1 - p_{ai}^{ow} = \frac{1}{1 + \exp\left(V_o(w) - V_a(i) - c_a^o\right)}.$$

Plugging this back to the value function yields

$$(\delta_a + \rho)\, V_a(i) = u_a(i) + \delta_a V_u + \sum_o \lambda_a^o\, \mathbb{E}_w\{-\log(1 - p_{ai}^{ow})\}.$$

Finally, integrating out the offered wage pmf yields Equation 2.3:

$$(\delta_a + \rho)\, V_a(i) = u_a(i) + \delta_a V_u - \sum_{o,w} \lambda_a^o \log(1 - p_{ai}^{ow}) f^o(w).$$

The same logic applies to the value of non-employment.

C Identifying Structural Parameters

Offer arrival rates within occupations: From Equation 3.7,

$$\varpi_{ai}^{aj} + \varpi_{aj}^{ai} = \varpi_{ak}^{a\ell} + \varpi_{a\ell}^{ak} \tag{C.1}$$

$$\frac{h_{ai}^{aj}}{\lambda_a^a f^a(j) - h_{ai}^{aj}} \frac{h_{aj}^{ai}}{\lambda_a^a f^a(i) - h_{aj}^{ai}} = \frac{h_{ak}^{a\ell}}{\lambda_a^a f^a(\ell) - h_{ak}^{a\ell}} \frac{h_{a\ell}^{ak}}{\lambda_a^a f^a(k) - h_{a\ell}^{ak}} \tag{C.2}$$

Therefore

$$\lambda_a^a = \frac{\left[h_{ai}^{aj} f^a(i) + h_{aj}^{ai} f^a(j)\right] h_{ak}^{a\ell} h_{a\ell}^{ak} - \left[h_{ak}^{a\ell} f^a(k) + h_{a\ell}^{ak} f^a(\ell)\right] h_{ai}^{aj} h_{aj}^{ai}}{f^a(i) f^a(j) h_{ak}^{a\ell} h_{a\ell}^{ak} - f^a(k) f^a(\ell) h_{ai}^{aj} h_{aj}^{ai}} \tag{C.3}$$

Offer arrival rates across occupations: From Equation 3.10,

$$\begin{pmatrix} \varpi_{ai}^{bj} + \varpi_{bj}^{ai} \\ \varpi_{ai'}^{bj'} + \varpi_{bj'}^{ai'} \end{pmatrix} = \begin{pmatrix} \varpi_{ak}^{b\ell} + \varpi_{b\ell}^{ak} \\ \varpi_{ak'}^{b\ell'} + \varpi_{b\ell'}^{ak'} \end{pmatrix} \tag{C.4}$$

$$\begin{pmatrix} \dfrac{h_{ai}^{bj}}{\lambda_a^b f^b(j) - h_{ai}^{bj}} \dfrac{h_{bj}^{ai}}{\lambda_b^a f^a(i) - h_{bj}^{ai}} \\ \dfrac{h_{ai'}^{bj'}}{\lambda_a^b f^b(j') - h_{ai'}^{bj'}} \dfrac{h_{bj'}^{ai'}}{\lambda_b^a f^a(i') - h_{bj'}^{ai'}} \end{pmatrix} = \begin{pmatrix} \dfrac{h_{ak}^{b\ell}}{\lambda_a^b f^b(\ell) - h_{ak}^{b\ell}} \dfrac{h_{b\ell}^{ak}}{\lambda_b^a f^a(k) - h_{b\ell}^{ak}} \\ \dfrac{h_{ak'}^{b\ell'}}{\lambda_a^b f^b(\ell') - h_{ak'}^{b\ell'}} \dfrac{h_{b\ell'}^{ak'}}{\lambda_b^a f^a(k') - h_{b\ell'}^{ak'}} \end{pmatrix} \tag{C.5}$$

The arrival rates $(\lambda_a^b, \lambda_b^a)'$ are the solution to this system of equations; analytically,

$$\lambda_a^b = \frac{A_2 B_3 - A_3 B_2}{A_2 B_1 - A_1 B_2} \qquad \text{and} \quad \lambda_b^a = \frac{A_3 B_2 - A_2 B_3}{A_3 B_1 - A_1 B_3} \tag{C.6}$$

where

$$A_1 = f^b(\ell) f^a(k) h_{ai}^{bj} h_{bj}^{ai} - f^b(j) f^a(i) h_{ak}^{b\ell} h_{b\ell}^{ak}, \quad B_1 = f^b(\ell') f^a(k) h_{ai'}^{bj'} h_{bj'}^{ai'} - f^b(j') f^a(i') h_{ak'}^{b\ell'} h_{b\ell'}^{ak'},$$

$$A_2 = f^b(\ell) h_{b\ell}^{ak} h_{ai}^{bj} h_{bj}^{ai} - f^b(j) h_{bj}^{ai} h_{ak}^{b\ell} h_{b\ell}^{ak}, \quad B_2 = f^b(\ell') h_{b\ell'}^{ak'} h_{ai'}^{bj'} h_{bj'}^{ai'} - f^b(j') h_{bj'}^{ai'} h_{ak'}^{b\ell'} h_{b\ell'}^{ak'},$$

$$A_3 = f^a(k) h_{ak}^{b\ell} h_{ai}^{bj} h_{bj}^{ai} - f^a(i) h_{ai}^{bj} h_{ak}^{b\ell} h_{b\ell}^{ak}, \quad B_3 = f^a(k') h_{ak'}^{b\ell'} h_{ai'}^{bj'} h_{bj'}^{ai'} - f^a(i') h_{ai'}^{bj'} h_{ak'}^{b\ell'} h_{b\ell'}^{ak'}.$$

Flow utilities and switching costs: The overdetermined linear system in Equation 3.20 can be written in matrix form as

$$\boldsymbol{\kappa} = A\boldsymbol{\theta}. \tag{C.7}$$

Here

$$\boldsymbol{\kappa} = \{\kappa_{ai}^{bj}\}_{a,i,b,j} \tag{C.8}$$

with

$$\kappa_{ai}^{bj} = \varpi_{ai}^{bj} + \frac{1}{\rho+\delta_b}\left(\sum_{o,w}\lambda_b^o \log(1-p_{bj}^{ow})f^o(w)\right) - \frac{1}{\rho+\delta_a}\left(\sum_{o,w}\lambda_a^o \log(1-p_{ai}^{ow})f^o(w)\right); \tag{C.9}$$

$$\boldsymbol{\theta} = \begin{pmatrix} \{u_a(i)\}_{a,i} \\ V_u \\ \{c_a^b\}_{a,b} \end{pmatrix}, \tag{C.10}$$

and A is a design matrix that assigns $\pm\frac{1}{\rho+\delta_a}$ to $u_a(\cdot)$, $\frac{\delta_b-\delta_a}{(\rho+\delta_b)(\rho+\delta_a)}$ to V_u, and -1 to $c_\cdot^\cdot$.

Using the log parametrization $u_a(i) = \psi_a + \beta \cdot \log w_i$, the matrix representation becomes

$$\boldsymbol{\kappa} = A_l\, \boldsymbol{\theta}_l. \tag{C.11}$$

Here $\boldsymbol{\kappa}$ is unchanged from before,

$$\boldsymbol{\theta}_l = \begin{pmatrix} \{\psi_a\}_a \\ \beta \\ V_u \\ \{c_a^b\}_{a,b} \end{pmatrix}, \tag{C.12}$$

and A_l is a design matrix that assigns $\frac{1}{\rho+\delta_b}$ or $-\frac{1}{\rho+\delta_a}$ to $\psi_\cdot$, $\frac{1}{\rho+\delta_b}\log w_j - \frac{1}{\rho+\delta_a}\log w_i$ to β, $\frac{\delta_b-\delta_a}{(\rho+\delta_b)(\rho+\delta_a)}$ to V_u, and -1 to $c_\cdot$.

D Estimation Procedure

D.1 Baseline specification

Let h_{ai}^{bj} denote the hazard of moving from job (a,i) to (b,j). The cumulative hazard of switching from (a,i) to (b,j) at time t is

$$H_{ai}^{bj}(t) = \int_0^t h_{ai}^{bj}\, du = h_{ai}^{bj}\, t. \tag{D.1}$$

It then follows that the overall survival function can be written as the product of the destination-specific cumulative hazards:

$$S_{ai}(t) = \prod_{b,j} \exp\left(-H_{ai}^{bj}(t)\right) = \prod_{b,j} \exp\left(-h_{ai}^{bj}\, t\right). \tag{D.2}$$

Assume that we possess data on N individuals with S_n spells each. Specifically, we observe data $\{t_s, a_s, i_s, b_s, j_s\}_{s\in S_n}$ for all $S_n \in \{S_1, \dots, S_N\}$. That is, for spell s we know its length t_s, and the occupation and wage rate in the origin job (a_s, i_s) and in the destination job (b_s, j_s). The likelihood of observing these data for individual n is

$$L_n = \prod_{s=1}^{S_n} L_{ns}\left(h_{ai}^{bj}\right) \tag{D.3}$$

where $L_{ns}\left(h_{ai}^{bj}\right)$ is the likelihood contribution of individual n's spell s; i.e.,

$$L_{ns}\left(h_{ai}^{bj}\right) = \prod_{a,i}\prod_{b,j}\left[\left(h_{ai}^{bj}\right)^{\mathbb{1}(b_s=b,j_s=j)} \exp\left(-h_{ai}^{bj}\, t_s\right)\right]^{\mathbb{1}(a_s=a,i_s=i)}. \tag{D.4}$$

I impose the structure of the model on this likelihood using Equation 3.1. With it, I can write the likelihood as

$$L_n(f,\lambda,p) = \prod_{s=1}^{S_n} L_{ns}\left(\lambda_a^b f^b(j) p_{ai}^{bj}\right). \tag{D.5}$$

Going further, I get the CCPs by iterating the value function to a fixed point within the routine: in the nth iteration,

$$\left(\sum_o \lambda_a^o + \delta_a + \rho\right) V_a(i)^{(n)} = u_a(i) + \delta_a V_0^{(n-1)} + \sum_o \lambda_a^o \left(V_a(i)^{(n-1)}\right) \\ + \sum_{b,j} \lambda_a^b \log\left(1 + \exp\left(V_b(j)^{(n-1)} - V_a(i)^{(n-1)} - c_a^b\right)\right) f^b(j). \quad \text{(D.6)}$$

From here, I calculate the CCPs as

$$p_{ai}^{bj} = \frac{\exp(V_b(j) - V_a(i) - c_a^b)}{1 + \exp(V_b(j) - V_a(i) - c_a^b)}. \quad \text{(D.7)}$$

This way I can express the CCPs in terms of the flow utilities and switching costs. Assembling all terms, I ultimately write the likelihood as

$$L_n(f, \lambda, u, c) = \prod_{s=1}^{S_n} L_{ns}\left(\lambda_a^b f^b(j) p_{ai}^{bj}(u_a(i), c_a^b)\right). \quad \text{(D.8)}$$

Therefore, the parameter estimates are

$$(\hat{f}, \hat{\lambda}, \hat{u}, \hat{c}) = \arg\max_{f,\lambda,u,c} L_n(f, \lambda, u, c) = \arg\max_{f,\lambda,u,c} \prod_{s=1}^{S_n} L_{ns}\left(\lambda_a^b f^b(j) p_{ai}^{bj}(u_a(i), c_a^b)\right). \quad \text{(D.9)}$$

D.2 Introducing unobserved heterogeneity

In a next step, I add unobserved heterogeneity to the mix. I model workers to be one of a pre-specified, discrete number of types with some probability. Following Arcidiacono and Miller (2011), I estimate this heterogeneous model using an EM algorithm.

With unobserved heterogeneity, the full likelihood of observing the data modifies to

$$L_n = \sum_{r=1}^{R} \pi_r \left(\prod_{s=1}^{S_n} L_{nsr}\left(h_{ai}^{bj}(r)\right)\right) \quad \text{(D.10)}$$

where π_r is the population probability of type r and $L_{nsr}\left(h_{ai}^{bj}(r)\right)$ is the likelihood contribution of individual n's spell s given that she is type r; i.e.,

$$L_{nsr}\left(h_{ai}^{bj}(r)\right)=\prod_{a,i}\prod_{b,j}\left[\left(h_{ai}^{bj}(r)\right)^{\mathbb{1}(b_s=b,j_s=j)}\exp\left(-h_{ai}^{bj}(r)\,t_s\right)\right]^{\mathbb{1}(a_s=a,i_s=i)}. \tag{D.11}$$

The expected loglikelihood is

$$\log L=\sum_{n=1}^{N}\sum_{r=1}^{R}q_{nr}\left(\sum_{s=1}^{S_n}\log L_{nsr}\left(h_{ai}^{bj}(r)\right)\right) \tag{D.12}$$

where q_{nr} is the posterior probability that individual n is type r. I implement the following EM algorithm:

0. Initialize posterior probabilities $\{q_{nr}^{(0)}\}_{n,r}$, then average them out by the first occupation in each individuals' job history to retrieve population probabilities:

$$\pi_{ar}^{(0)}=\frac{1}{N_a}\sum_{n\in a}q_{nr}^{(0)}. \tag{D.13}$$

   Here N_a is the number of individuals whose first observed occupation was a, and $n\in a$, abusing notation for the sake of brevity, denotes those individuals.

1. **M-step.** Taking posterior probabilities $\{q_{nr}^{(m-1)}\}_{n,r}$ as given, estimate hazard rates $h_{ai}^{bj}(r)^{(m)}$ by maximizing the expected likelihood in Equation D.12.

2. **E-step.** Taking hazard rates $\left\{h_{ai}^{bj}(r)^{(m)}\right\}_{r,a,i,b,j}$ as given, update posterior probabilities as

$$q_{nr}^{(m)}=\frac{\pi_{a_1r}^{(m-1)}\left(\prod_{s=1}^{S_n}L_{nsr}\left(h_{ai}^{bj}(r)^{(m)}\right)\right)}{\sum_r\pi_{a_1r}^{(m-1)}\left(\prod_{s=1}^{S_n}L_{nsr}\left(h_{ai}^{bj}(r)^{(m)}\right)\right)}. \tag{D.14}$$

   a_1 indicates the first occupation in individual n's job history. Renew population probabilities by occupation as

$$\pi_{ar}^{(m)} = \frac{1}{N_a} \sum_{n \in a} q_{nr}^{(m)}. \tag{D.15}$$

Repeat steps 1 and 2 until convergence.

D.2.1 Additional parametrizations

I impose additional structure on the wage offer distributions. Wages for Type 1 workers in occupation a are drawn from the distribution $f^a(\cdot, 1)$, parametrized as

$$f^a(\cdot,1) = \begin{pmatrix} \exp(\gamma_1)/(1+\sum_j \exp(\gamma_j)) \\ \exp(\gamma_2)/(1+\sum_j \exp(\gamma_j)) \\ \vdots \\ \exp(\gamma_{W-1})/(1+\sum_j \exp(\gamma_j)) \\ 1/(1+\sum_j \exp(\gamma_j)) \end{pmatrix} \tag{D.16}$$

where the γ parameters are unrestricted. Wages for Type $r \neq 1$ workers then are

$$f^a(\cdot,r) = \begin{pmatrix} \exp(\theta_r\gamma_1)/(\theta_r+\sum_j \exp(\theta_r\gamma_j)) \\ \exp(\theta_r\gamma_2)/(\theta_r+\sum_j \exp(\theta_r\gamma_j)) \\ \vdots \\ \exp(\theta_r\gamma_{W-1})/(\theta_r+\sum_j \exp(\theta_r\gamma_j)) \\ \theta_r/(\theta_r+\sum_j \exp(\gamma_j)) \end{pmatrix} \tag{D.17}$$

where θ_r is the type-specific logit shifter. If $\theta_r > 1$, the distribution shifts to the right.

E Additional Results

E.1 No unobserved heterogeneity

Figure E.1: Hazard rates

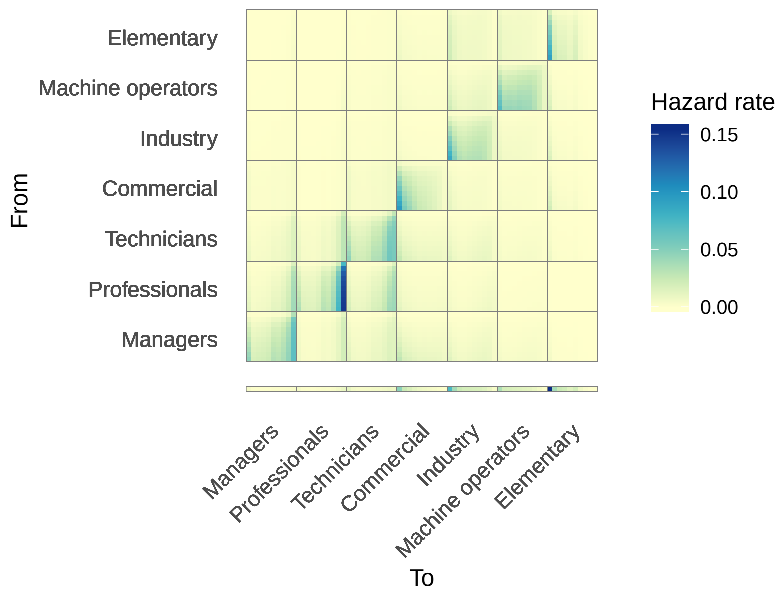
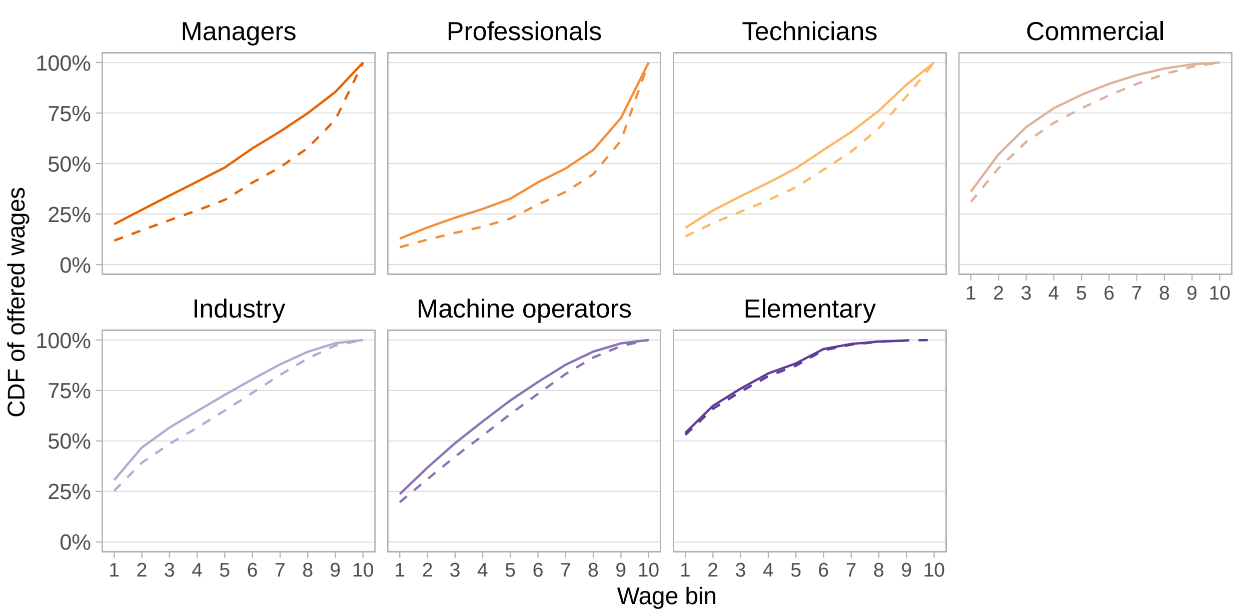
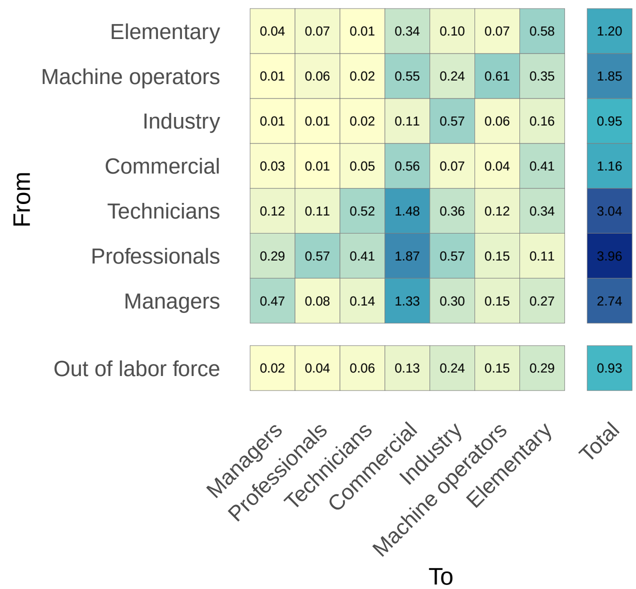


Figure E.2: Offered vs. accepted wages



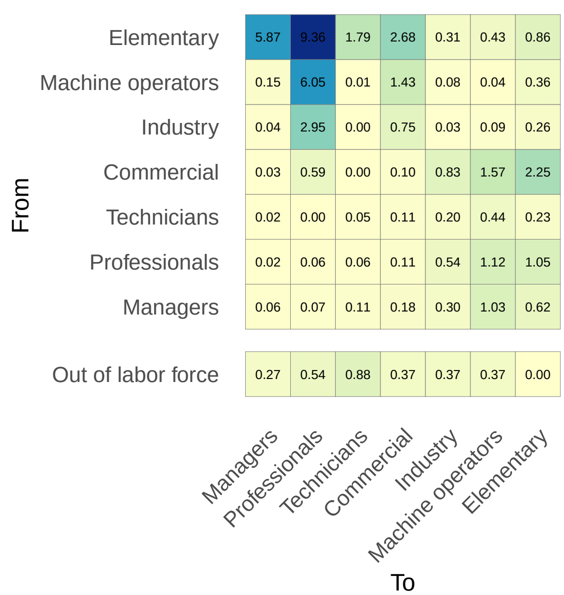
Notes: Solid lines are offered wage CDFs. Dashed lines are accepted wage CDFs. Back to Figure 4.1.

Figure E.3: Offer arrival rates



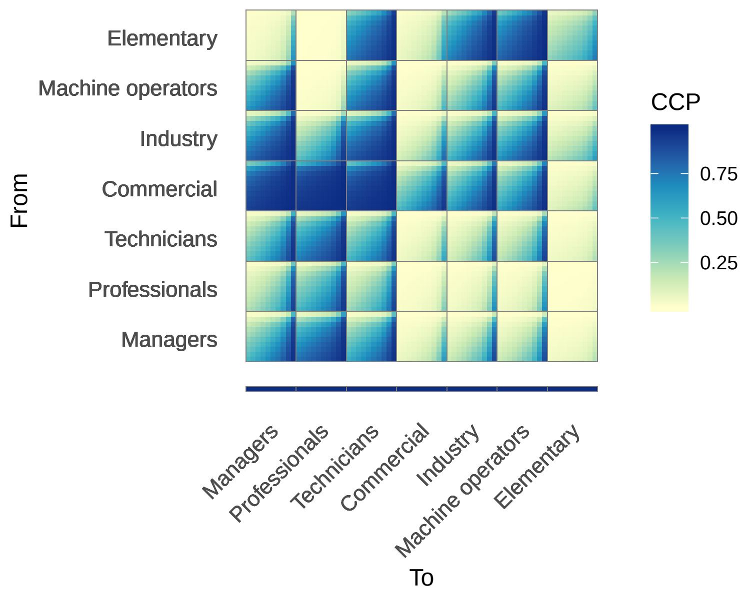
Notes: Back to Table 4.1.

Figure E.4: Mean switching costs



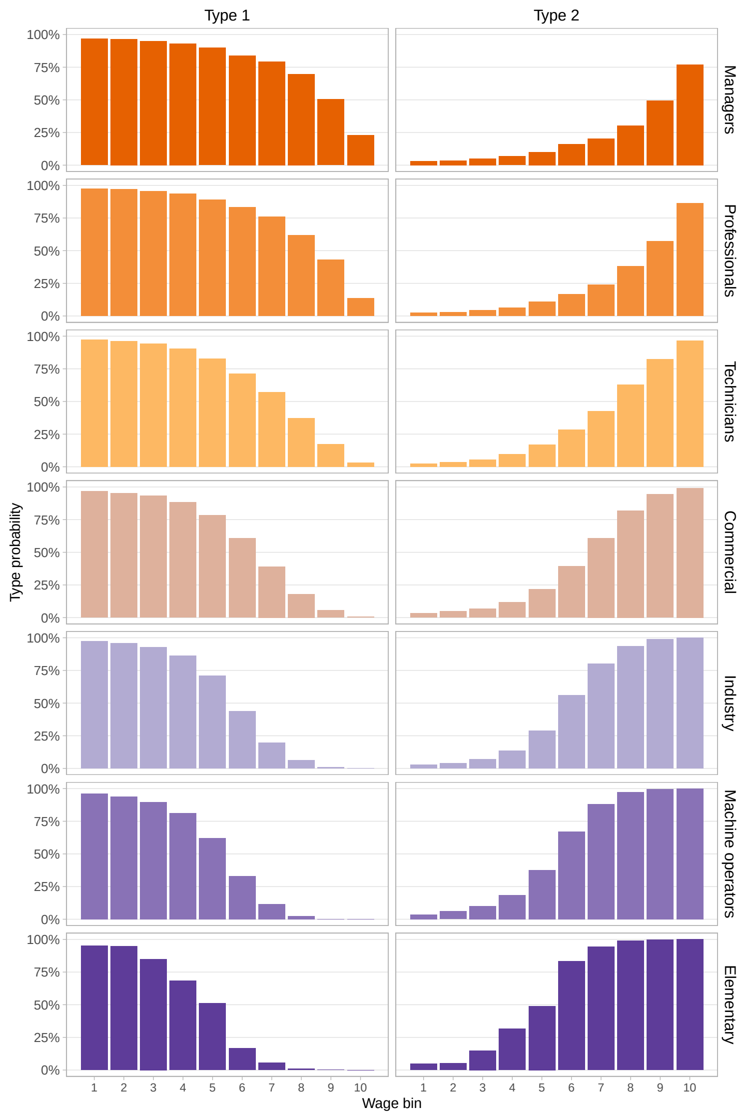
Notes: Back to Table 4.2.

Figure E.5: Conditional choice probabilities of accepting offers



E.2 Unobserved heterogeneity: two types

Figure E.6: Type probabilities



Notes: Back to Figure 4.2.

Figure E.7: Hazard rates

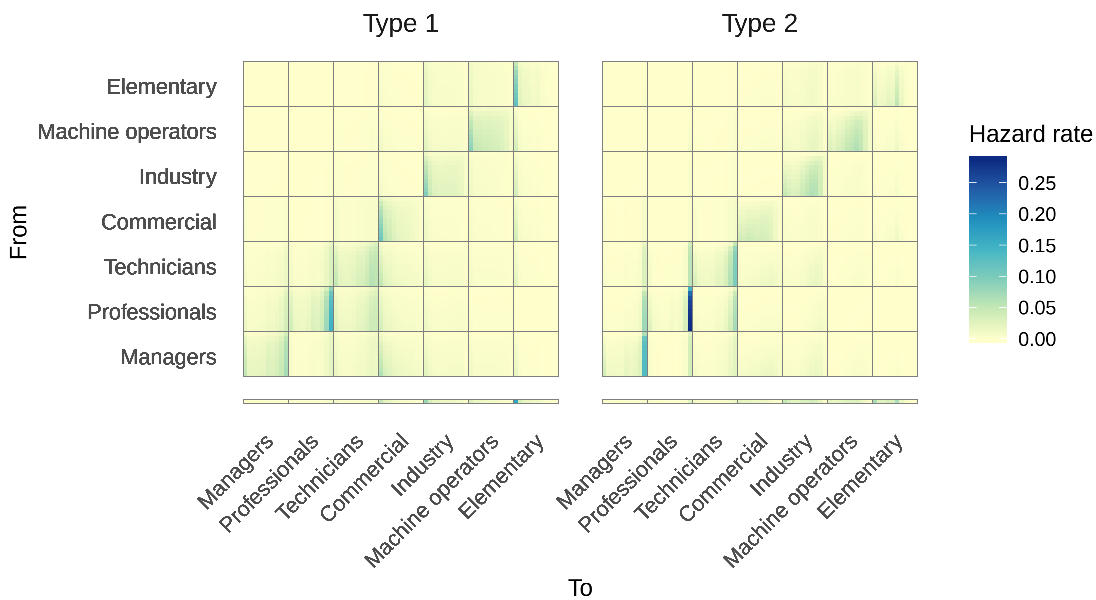
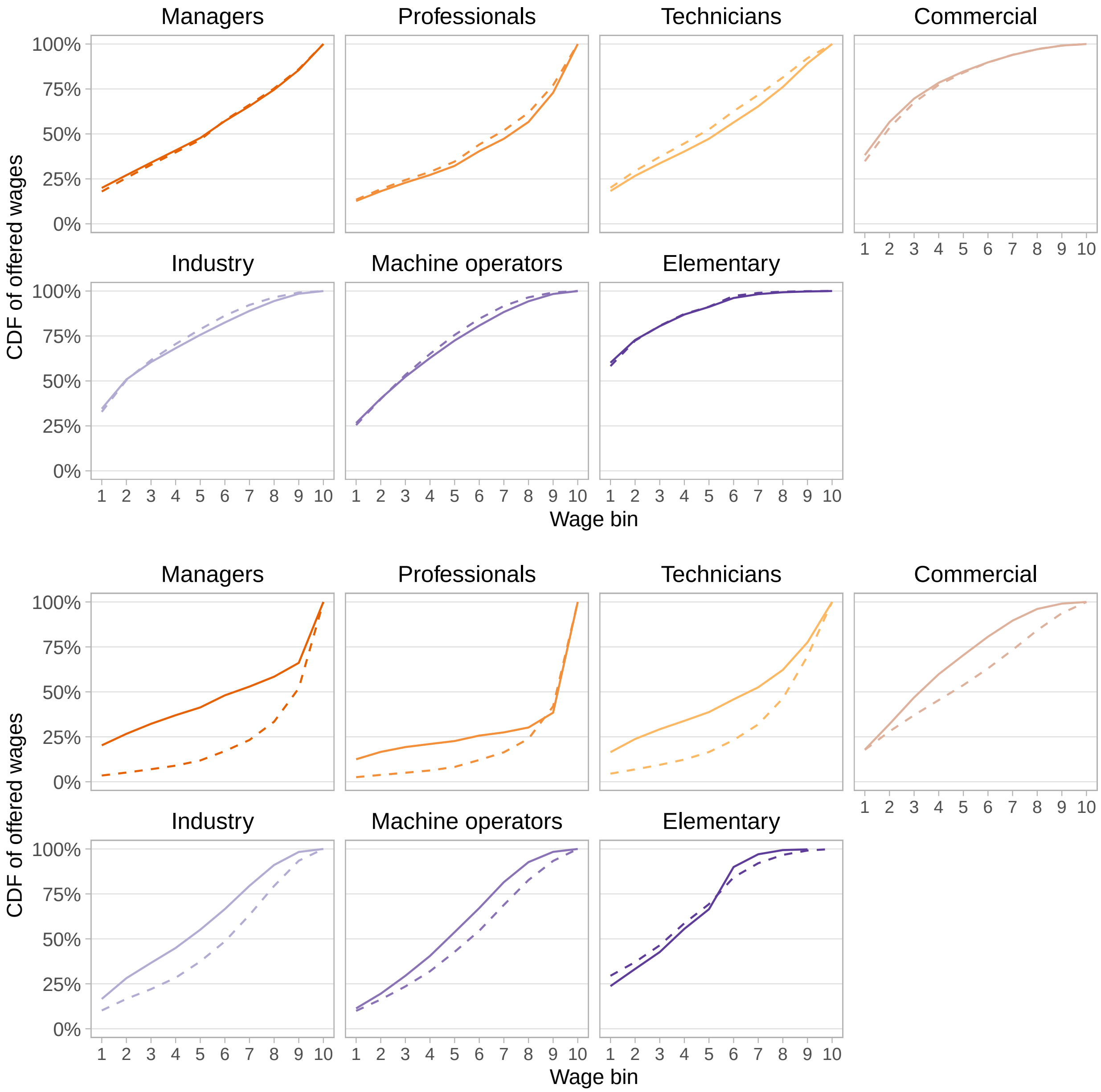


Figure E.8: Offered vs. accepted wages



Notes: Solid lines are offered wage CDFs. Dashed lines are accepted wage CDFs. Back to Figure 4.3.

Figure E.9: Offer arrival rates

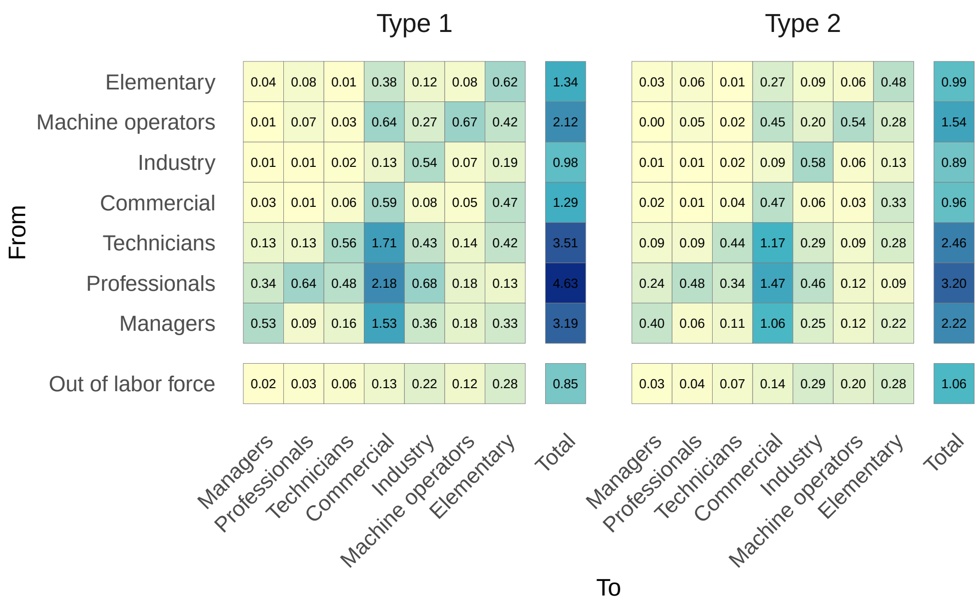


Figure E.10: Conditional choice probabilities of accepting offers

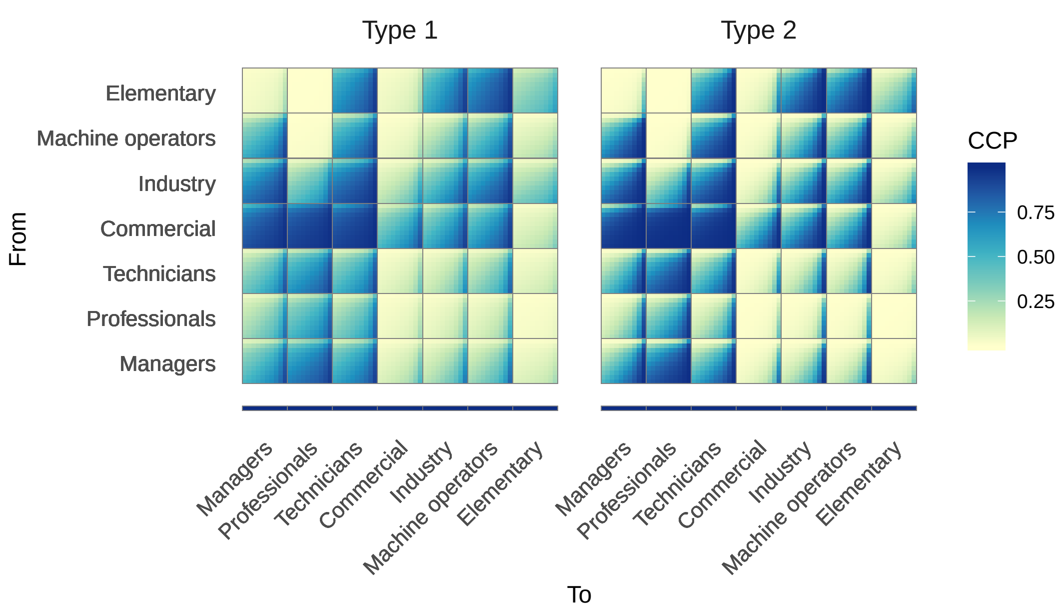
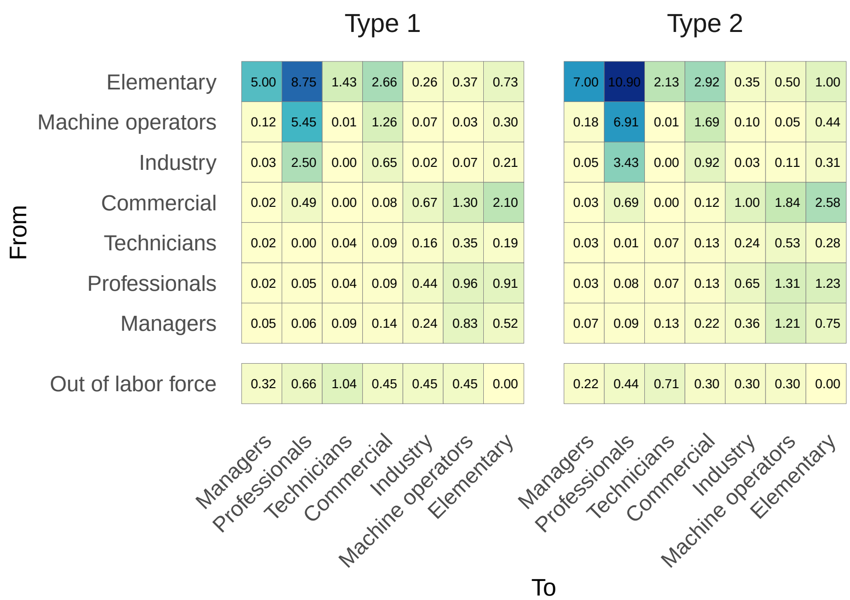


Figure E.11: Mean switching costs



Notes: Back to Table 4.5.

E.3 Career simulations

Table E.1: BIC of counterfactual model estimates

Estimates	BIC
Baseline	13,356,714.90
Occupationless benchmark	14,501,118.10
Only opportunities	13,455,938.67
Only preferences	14,409,294.71
Only switching costs	14,190,990.82

Notes: Bayesian Information Criteria: $\text{BIC} = -2\log\hat{L} + k\log N$ where $\hat{L}$ is the likelihood of the model, k denotes the number of parameters and N denotes the sample size. Back to Figure 5.3.