Multidimensional Human Capital and Wage Dynamics: Importance of Cognitive, Manual, and Interpersonal Skills

Soojeong Jung*

January 2025

Click here for latest version

Abstract

This paper studies how cognitive, manual, and interpersonal task-specific human capital shape the wage dynamics over the early career. The paper develops a model of multidimensional human capital accumulation and wage progression, where occupations differ in task intensities. Task intensities are constructed based on the job descriptors of O*NET. Workers, with different initial skill levels, accumulate task-specific human capital from experience, which is transferable across occupations. Potential worker selection effects are addressed through a Markov model of occupation choice with unobserved worker types. The model is estimated using employment histories from the National Longitudinal Survey of Youth 1997. Decomposition exercises show that 63% of wage growth is explained by the accumulation of cognitive human capital, 25% by manual human capital, while only 9% comes from interpersonal human capital. The significantly smaller contribution of interpersonal skills is attributed to their slower rate of accumulation. Cross-sectional wage variation across workers is primarily explained by initial skills and occupation-level wage differences, with cognitive human capital accumulation accounting for the rest.

*University of Wisconsin-Madison, soojeong.jung@wisc.edu. I am extremely grateful to my advisors, Chris Taber, John Kennan and Corina Mommaerts for invaluable advice, guidance and support.

1 Introduction

Wage growth varies widely across occupations, contributing to wage dispersion over time. A key factor behind these disparities is the nature of tasks performed in different occupations, which also shapes how workers accumulate skills through experience on the job. The significance of task-specific human capital compared to other types of human capital is well documented in the literature (Gathmann and Schönberg (2010)), but the dynamics of task-specific skills and wage evolution are understudied. This paper offers a perspective on how each skill type distinctly influences wage dynamics and how the skills accumulate over one's career. To do so, it is crucial to understand how different tasks contribute to wage and affect human capital accumulation differently across various occupations.

This paper answers two main questions: How does the accumulation of multidimensional human capital influence wage dynamics, and to what extent do different skills—cognitive, manual, and interpersonal—drive wage growth and variation?

In this paper, I develop a model that incorporates multi-dimensional human capital accumulation over the life-cycle. Human capital is task-specific, which has cognitive, manual, and interpersonal dimensions. When entering labor market, workers are different in 3-dimensional initial human capital, captured by unobserved worker types. Human capital accumulates based on tasks workers perform on the job, and is transferable when the occupation changes. However, the multidimensional human capital will be valued differently in the new occupation. This allows for wage differences across occupations given the same level of human capital, determined by how much each occupation requires each skill dimension.

The model addresses self-selection in occupation choices. Specifically, I approximate the occupation choice probabilities by using a Markov model, with multinomial logit specification. In the model, workers choose occupation based on previous occupation, experiences in each task, and the unobserved worker type. This modeling approach enables a direct examination of worker selection, incorporating both worker heterogeneity and experience as key determinants.

I estimate the model using wage and occupational panel data of workers born between 1980 and 1984 from National Longitudinal Survey of Youth 1997. Task intensities, which capture occupational task requirements, are constructed using the O*NET database.

The accumulation rate of task-specific human capital quantifies the profitability of the task experience in different occupations. It is identified from comparisons of on-the-job wage growths across occupations with varying task content. The three-dimensional initial skill level distribution is identified from within worker mean wage differences in different occupations with varying task use.

Using the estimated model, two linear decompositions are conducted to quantify the effects of each skill on wage growth over time and cross-sectional wage variation. Over the first ten years since labor market entry, mean wage level increases by 0.32 log point. The accumulation of cognitive human capital drives 63% of the increase. Manual skill accumulation contributes another 25% to wage growth. The accumulation of interpersonal skills contribute only 9%. The small contribution of interpersonal skill accumulation is due to the slower rate of accumulation of interpersonal skills compared to the other two skills. In other words, workers do not experience significant wage increases from gaining additional experience in interpersonal tasks.

At ten years since labor market entry, the explained variation in wages across workers is primarily explained by initial skills and occupation-specific factors, which account for about 80%. The rest 20% of wage dispersion across workers can be attributed to human capital accumulation, with cognitive skills making the largest contribution. Manual and interpersonal skill accumulation combined contribute to only 2% of variation across worker. Moreover, the observed occupation assignments matter for wage. If instead of observed occupation, workers choose wage-optimal occupations, then more workers tend to choose occupation that is manual heavy and contribution of manual skill would be bigger.

These findings emphasize the central role of cognitive human capital in both wage growth and wage dispersion. Workers with stronger initial cognitive skill and experience in cognitively demanding occupations tend to earn higher wages over time, thereby contributing to greater wage inequality. In contrast, manual skills yield modest but positive returns, while interpersonal tasks have little impact on wages, suggesting that task-specific skill accumulation across different dimensions of human capital has varying effects on shaping one's wage over the life-cycle.

This paper makes contributions to the literature on human capital and wage dynamics. The huge literature on returns to specific skills and heterogeneous human capital addresses the transferability of human capital, highlighting the importance of task-specific human capital compared to occupation- and firm-specific human capital (e.g.,(Poletaev and Robinson (2008), Kambourov and Manovskii (2009), Gathmann and Schönberg (2010), Pavan (2011)).[1] My paper builds on this approach by treating human capital as multidimensional and task-specific, and contributes by focusing on the relative importance of skills in shaping wage and quantifying the unequal impacts of these skills. Keeping the task dimensions at cognitive, manual, and interpersonal—neither too few nor too many—aligns well with previous literature. For instance, Deming (2017) emphasizes the role of social skills in wage determination. I supplement by highlighting that while initial skills significantly influence wages, interpersonal skills develop more slowly, resulting in cognitive and manual skills becoming more critical as workers gain experience. This finding aligns with Lise and Postel-Vinay (2020b), which notes a slower adjustment speed for interpersonal skills compared to cognitive or manual skills.

Another distinctive contribution of this paper is the sample used; by analyzing the wage dynamics of the millennial U.S. cohort with NLSY 97 data, this study provides an updated view, as well as capturing the wage dynamics of both women and men. Additionally, the approach to addressing worker selection diverges from using instruments (Kambourov and Manovskii (2009), Gathmann and Schönberg (2010)) or fully structural approaches (Pavan (2011)) by using a Markov model, offering a simpler yet effective way to account for

[1]See Sanders and Taber (2012) for an excellent review of the literature.

worker selection (Sauer and Taber (2021) uses similar Markov model to study wage growth of women).

My paper also contacts with the literature using the task content of occupations. This paper is different in focus, by examining the life-cycle effects of accumulated skill and initial skill, rather than the cost of mismatch between worker skill and occupational requirement (Lise and Postel-Vinay (2020b), Guvenen et al. (2020)), or changes in returns to skill over time (Roys and Taber (2019)). Also, the initial skills are identified in a different way. Rather than using proxies of initial skill in cognitive, manual, and interpersonal skill from the survey, the initial skill distribution is estimated from wage data using MLE, and specifically from wage change upon occupational changes. This approach more reliably captures how effectively a worker will perform specific tasks in a job setting.

While Adda and Dustmann (2023) provides a comprehensive analysis of multidimensional transferable skills and their contribution to wage growth through multiple channels—including pre-labor market choices and their interplay with wage growth dynamics—my research offers a distinct perspective. I focus on task-specific human capital accumulation across cognitive, manual, and interpersonal dimensions, rather than modeling routine-manual and cognitive-abstract sector experiences to be transferable. In addition, I examine not only how wages grow, but also how wage variation emerges among workers due to accumulation of human capital. Lastly, my paper focuses on the U.S. cohort born between 1980 and 1984, offering insights into a younger and more diverse population, whereas Adda and Dustmann (2023) examines men from West Germany born between 1960 and 1972.

The rest of the paper is organized as follows. Section 2 explains data and occupational task intensity measure. Section 3 explains the model. Section 4 explains loose identification and discuss estimated coefficients. Section 5 explains wage variance and wage growth decomposition exercises. Section 6 concludes.

2 Data

2.1 NLSY 97

The employment and wage data are derived from the National Longitudinal Survey of Youth 1997 (NLSY97). My sample includes only the cross-sectional sample and excludes the oversample. I use the surveys starting from 1997 to 2019.[2] In addition, entries are excluded under several conditions. Labor market entry is defined as the month the worker received the final degree or the month following the final enrollment in college or school. I exclude observations prior to an individual's labor market entry.[3] The entire work history of individuals whose occupation is not matched to the O*NET occupation code for some period (the link to the O*NET database is described below) is excluded. Veterans and workers whose final education is a graduate degree are excluded from the sample.

I construct yearly wage, experience, and occupation variables using the weekly employment event history, the employer roster, the variables constructed by the NLSY team. Starting from the week of labor market entry, I record the occupation and wage at the main job on a yearly basis.[4] Summary statistics of the analyzed sample is presented in Table 1

2.2 O*NET and Task Intensities

To determine the proportion of skills required for occupations, I use the O*NET database, administered by the U.S. Department of Labor. O*NET provides detailed description on tasks and activities associated with different occupations. I classify skills into cognitive,

[2]The survey was conducted annually from 1997 to 2010 and biannually from 2011 to 2019. However, the survey asks for complete employment histories for periods since the last survey, which are reflected in the weekly work history and employer roster files.

[3]This cleaning approach excludes part-time jobs during college and full-time jobs between high school graduation and college from being considered as work experience. However, if these jobs contributed to skill development before entering the labor market, their impact will be reflected in the initial skill level.

[4]If a worker is not employed exactly one year after but has held a job at some point during the past year, I record the most recent main job. Thus, if a worker is considered not working in the yearly panel, it means that the worker did not have a job for the entire year. I define the primary job as the job with the longest hours, by filling in wages and hours from the employer job roster into the event history.

Table 1: Summary Statistics

Variable	Mean	Standard deviation
Employed	0.899	0.301
Potential experience	7.946	5.114
Log(wage)	1.893	0.627
Final education: College	0.298	0.457
Female	0.461	0.499
Occupation: White-Collar	0.245	0.430
Occupation: Manual	0.212	0.408
Occupation: Service	0.293	0.455
Occupation: Technical	0.150	0.357
Number of individuals	4041	
Number of observations	48605	

Notes: Summary statistics for the pooled data of yearly variables. For details on occupation categories, see subsection 2.2.

manual and interpersonal. The first step is to follow the categorization of Lise and Postel-Vinay (2020a), so that each of the 197 descriptors of O*NET 18.1 Database are assigned into one of the three groups.[5]

Given that each skill category has multiple descriptors assigned, I first reduce the dimensionality of the descriptors into one, for each skill category. I use principal component analysis separately for each skill category, running the analysis three times in total similar to Roys and Taber (2019), to obtain the three factor scores for each occupation. Following this, I shift each factor score to ensure the non-negativity of the measure, by subtracting minimum value of each factor separately.[6] Finally, I adjust these measures by dividing each by 10 to make the scale manageable. A vector of the three (transformed) first principal components for each skill represents the task intensity of each skill category for the occupation.

[5] Although I employ the categorization provided in their replication package, Lise and Postel-Vinay (2020b) do not assign descriptors directly into these three skill categories when constructing the skill intensity measure. Instead, they conduct a single PCA with three factors, subsequently recombining these factors so that they can be interpreted as representing cognitive, manual, and interpersonal skills. This approach captures the covariance patterns between skills. Instead of using their score, I choose to construct three-dimensional factor scores that are easier to interpret.

[6] The non-negativity is required because I assume that human capital level increases by task intensity. I am interpreting task intensity as amount of tasks used on occupation. Even though negative measure will still preserve the relative comparison, shifting the scores make this interpretation easier.

Table 2: Task Intensities for Occupations in the NLSY 97 Sample (SOC 2010 Codes)

Variable	Obs	Mean	SD	Min	Max
γ^C	484	1.67	0.66	0.00	3.15
γ^M	484	1.04	0.60	0.01	2.34
γ^I	484	1.03	0.39	0.00	1.98

Notes: Task intensities (γ^κ) are constructed for cognitive, manual, and interpersonal skills. See text for construction methods. The table includes occupations present in the NLSY 97 sample, matched to 484 SOC 2010 codes.

Occupational Task Intensity at the 3-Digit Level Next, I match O*NET SOC 2010 codes to 2002 Census codes in NLSY 97. Table 2 summarizes task intensity measures of 484 occupations taken by workers in the NLSY 97 sample.[7] The dimensions of the tasks exhibit different ranges and variability in their intensities across occupations.

Aggregated Occupations Based on Task Intensity Finally, I group the 3-digit occupations into four categories based on the vectors of three-dimensional task intensity of each occupation using K-means clustering. After grouping the occupations, I take the average of the task intensities of each skill (manual, cognitive, and interpersonal) across the occupations within a group to represent the task intensities of each skill of each group.[8][9] I name the

[7]NLSY 97 uses 2002 Census codes for occupation and O*NET 18.1 Database uses SOC 2010 codes. According to the crosswalk provided by Beard et al. (2022), 2002 4-digit Census occupation codes are the same as 10 multiplied by the 2000 3-digit Census occupation codes. For the crosswalk between 2000 Census occupation codes and SOC 2010 codes, I use the crosswalk done by Carl Sanders in Sanders (2012), which can be found in the replication file of Lise and Postel-Vinay (2020b). Census also provides crosswalk here: https://www2.census.gov/programs-surveys/demo/guidance/industry-occupation/occ2000t.pdf. Some 2002 3-digit Census codes correspond to the same SOC 2010 codes. Thus, to be accurate, the constructed task intensities varies at the SOC codes level, not the Census codes level.

[8]Note that PCA reduces the dimensionality tasks from 197 descriptors to three categories—cognitive, manual, and interpersonal—while K-means clustering groups *occupations*, preparing the data for the model. Although occupations could alternatively be aggregated by using K-means clustering based on the 197 O*NET descriptors, I find it more intuitive to base aggregated on the three-dimensional task intensity. This is because I need the task intensity measures constructed for each aggregated occupation regardless.

[9]The choice of $K = 4$ categories is based on several considerations. First, with skills being three-dimensional, at least three occupational categories are necessary for identifying both initial skills and skill accumulation rates. Additionally, selecting $K = 4$ provides a balance between reducing the within-cluster sum of squares and maintaining a manageable number of clusters. I used the elbow method suggested by Thorndike (1953) to examine the within-cluster sum of squares and similar statistics, following Makles (2012), for various cluster counts. Since occupation choices are modeled using a multinomial logit form,

aggregated occupations: White Collar, Manual, Service and Technical. Table 3 presents four aggregated occupation groups and their group-level task intensities in cognitive (γ^C), manual (γ^M), and interpersonal (γ^I) dimensions. Each group represents a different mix of task requirements, as illustrated by example occupations. The White-Collar category (Group 1) exhibits high cognitive task intensity ($\gamma^C = 2.36$) and moderate interpersonal task intensity ($\gamma^I = 1.39$), with examples including Regulatory Affairs Managers and Registered Nurses. Group 2, Manual occupations, shows higher manual intensity ($\gamma^M = 1.67$) with less emphasis on cognitive and interpersonal tasks, represented by roles such as Construction Laborers and Carpenters. Service occupations (Group 3) require balanced but relatively low levels across all task types, with examples like Cashiers and Customer Service Representatives. Finally, Technical occupations (Group 4) demonstrate a blend of task intensities, with a notable balance between cognitive, manual, and interpersonal dimensions, including occupations like Automotive Master Mechanics and Electricians.

Notably, cognitive and interpersonal task intensities are positively correlated across these groups, indicating that occupations requiring more cognitive tasks often also demand greater interpersonal skills. This could potentially be a problem in separately identifying interpersonal human capital accumulation rate from cognitive, but as is the case in Manual and Service occupations, they are not perfectly correlated.

3 Model

In this section, I introduce a model of wages and occupation choices that reflects the contribution of human capital accumulated in three different skill dimensions on one's wage profile. The model accounts for unobserved types of workers which capture the difference in the initial level of human capital in each skill dimension and how workers act differently with regard to their occupation choices. There are four occupations to choose from which are defined in the previous section.

having too many clusters would increase the number of parameters by a lot, which I aim to avoid.

Table 3: Representative Occupations and Task Intensity Across Aggregated Occupations

Aggregated Occupation	γ^C	γ^M	γ^I	**Example Occupations**
1 (White-Collar)	2.36	0.40	1.39	Regulatory Affairs Managers Executive Secretaries and Executive Administrative Assistants Registered Nurses Elementary School Teachers, Except Special Education
2 (Manual)	1.15	1.67	0.68	Driver/Sales Workers Laborers and Freight, Stock, and Material Movers, Hand Construction Laborers Construction Carpenters
3 (Service)	1.05	0.73	0.78	Cashiers Retail Salespersons Cooks, Fast Food Customer Service Representatives
4 (Technical)	2.00	1.30	1.26	First-Line Supervisors of Retail Sales Workers Automotive Master Mechanics Electricians Food Service Managers

Notes: Occupations are grouped into clusters based on similarities in task profiles. Task intensity scores represent the mean values across occupations within each cluster. The examples listed correspond to the four most frequent 3-digit SOC 2010 codes. Occupation names are given by the author.

Overview Time is discrete and starts at $t = 1$. In the initial period, workers begin with heterogeneous initial human capital levels (H_1^C, H_1^M, H_1^I). The initial skill distribution is discrete and characterized by time-invariant unobserved worker types.[10] In each period, workers make an occupational choice, after which wages are realized. Task-specific human capital then accumulates according to the task intensities required in their chosen occupation.

At period t, workers enter with three-dimensional human capital, represented by (H_t^C, H_t^M, H_t^I), their previous occupation O_{t-1}, and their designated worker type. Workers choose their occupation and wages for the period are determined. After wages are realized, human capital accumulates, resulting in an updated human capital level $(H_{t+1}^C, H_{t+1}^M, H_{t+1}^I)$ that workers carry into the next period. That is, workers enter period $t+1$ with their newly accumulated human capital, occupation from t (O_t), and their worker type which is time-invariant. This process continues iteratively.[11] The following subsections provide a more detailed explana-

[10] Workers are aware of their own types. The types are only unobserved to the econometrician.

[11] In the data, workers differ in the number of observations. The primary reason is that individuals less educated enter the labor market at an earlier age than others. For estimation, I take the observed number of periods for each worker as given. When assessing model fit or using the estimated model to decompose wages, I either simulate for the actual number of observed periods for each individual or simulate all individuals

tion of each component of the model.

3.1 Wage Model

Wage Determination The log wage for worker i who chose occupation o at time t is determined as follows.

$$\begin{aligned}\log(w_{iot}) &= \alpha_o + \gamma_o^C H_{it}^C + \gamma_o^M H_{it}^M + \gamma_o^I H_{it}^I + \epsilon_{it} \\ &= \alpha_o + \gamma_o' H_{it} + \epsilon_{it} \qquad (1)\end{aligned}$$

An occupation-specific intercept α_o captures the wage differences at the occupation level, irrespective of individual skill level. The key component of wages is the three-dimensional task-specific human capital, H_{it}. For a given bundle of task-specific skills, the value of those skills varies by occupation, where the value of each task-specific skill is higher in occupations that heavily utilize the task.

This implies that task-specific human capital is transferable across occupation. In other words, workers do not lose their human capital accumulated when they switch occupations. However, their skill levels (or human capital level) will be valued differently in different occupations, depending on how much each occupation values each skill. For instance, a worker with high cognitive skill will earn a higher wage in a cognitive skill-heavy occupation. In addition, this wage difference will be greater than for a worker with lower cognitive skill.

The model allows that task-specific human capital is transferable across occupations upon switching. Specifically, workers tend to earn higher wages in occupations that heavily utilize the tasks in which they have substantial skills. For example, a worker with high cognitive human capital will receive a higher wage in an occupation with a strong cognitive focus than in one without it.

Log wages are linear in human capital, and human capital accumulation is modeled as

over an extended period. I will clarify the specific approach taken in each section. An alternative would be to model the distribution of working periods.

a concave function of experience which will be discussed in subsection 3.3. Finally, an error term ϵ_{it}, which is assumed to follow a normal distribution $N(0, \sigma^2)$, captures unexplained factors that influence wages.

Note that the only parameter to be estimated in this equation is α_o; the task intensities γ represent occupation-specific characteristics which are directly taken from O*NET the data as explained in the section on task intensity. The law of motion for human capital H is simulated in the model which will be explained shortly.

3.2 Initial Skill and Worker Type

Unobserved worker types capture two key dimensions of heterogeneity among workers. The first is 3-dimensional initial skill H_1, and the other is the overall pattern in occupation choices. Initial skills will be described below and the occupation choice patterns will be described in subsection 3.4.

The initial skill bundle H_1 is a multidimensional attribute of each worker, representing initial levels of cognitive, manual, and interpersonal skills that contribute to earning potential within different tasks. It reflects the pre-market skills, but if there exist persistent characteristics of individuals in specific tasks (such as attitude, adaptability, and problem-solving capacity which remain stable over time), these will also be loaded into initial skill vector. The initial skill vector consists of three components: cognitive skill (H_1^C), manual skill (H_1^M), and interpersonal skill (H_1^I).

Workers are categorized into three discrete types based on their initial skill profiles, each type defined by a unique combination of (H_1^C, H_1^M, H_1^I) values. Thus, each worker type represents a distinct skill profile that influences earnings potential by aligning with the intensities of occupational tasks.

Table 4 summarizes the three worker types, where each type c is defined by its own values of (H_1^C, H_1^M, H_1^I) and associated probability mass, capturing the proportion of workers with a given ability profile. These types allow wage outcomes to depend not only on human

Table 4: Probability Mass and Initial Skill Level for Worker Types

Type	(H_1^C, H_1^M, H_1^I)	Probability Mass
$c = 1$	$(0, 0, 0)$	p_1
$c = 2$	$(\psi_2^C, \psi_2^M, \psi_2^I)$	p_2
$c = 3$	$(\psi_3^C, \psi_3^M, \psi_3^I)$	$1 - p_1 - p_2$

capital accumulation but also on the initial skills.

Since both occupation intercepts and the three-dimensional initial skills are time-invariant, three location normalizations are required for identification. I set the cognitive, manual, and interpersonal skill levels of the first type to zero.[12] This normalization allows the initial skill levels for other types to have negative values. Since skill levels affect the log of wages rather than wage levels, the exponential of initial skill scales the impact of accumulated skill on wage.

In the model, the initial skill H_1 influences wages through its interaction with the task requirements of each occupation. Specifically, a worker with a high cognitive skill (H_1^C) earns a higher mean wage compared to a worker with a lower H_1^C in the same occupation, as H_1^C enhances the return on tasks that are cognitively intensive. This means that higher H_1^C values lead to larger wage differences in occupations where cognitive skills are crucial.

Moreover, the wage gain from moving to a more cognitively intensive occupation would equal $\Delta\gamma_o^C H_1^C$, where $\Delta\gamma_o^C$ represents the increase in cognitive intensity between two occupations. Thus, as workers with high cognitive skill (H_1^C) switch to occupations that place greater emphasis on cognitive tasks, their wages rise in proportion to both their initial skill

[12] Although one type is normalized as the base, the model cannot uniquely identify which of the three types serves this role. This implies two alternative sets of support parameters. For example, if Type 2 serves as the base, Type 2 would have zero initial skills, while the initial skill bundle for Type 1 would adjust to $(-\psi_2^C, -\psi_2^M, -\psi_2^I)$, and Type 3 would become $(\psi_3^C - \psi_2^C, \psi_3^M - \psi_2^M, \psi_3^I - \psi_2^I)$. This flexibility is possible because adjustments to the occupation intercepts in the wage equation and the coefficients on type dummies in the occupation equation can yield the same occupation choice probabilities and wages across occupations. To address this, I impose a constraint requiring the initial cognitive skill levels of Types 2 and 3 to be non-positive, ensuring that the type with the highest cognitive skill level is designated as the base type. Additionally, the model may encounter a labeling issue where it interchanges the labels for Types 2 and 3. I resolve it by reassigning the labels based on the initial cognitive skill level after estimation, and adjusting probabilities and type specific coefficients accordingly.

level and the change in cognitive task intensity. This structure applies similarly to the manual (H_1^M) and interpersonal (H_1^I) dimensions, allowing the model to capture how variations in initial skills translate into wage differences across occupations based on occupational demands. This point will be revisited in identification section.

It is important to note that, unlike in some previous studies (such as Deming (2023) and Lise and Postel-Vinay (2020b)), initial skills in this model are not derived from data proxies. Instead, I infer initial skills using initial wage and wage variation upon occupation switches. This to better capture the non-time-varying component of task-specific productivity. Skill measures available in NLSY 97 for manual and interpersonal dimensions include ASVAB scores and personality measures. They reflect general tendencies, but do not necessarily indicate how effectively that worker will perform specific tasks in a job setting. By allowing wage data to reveal worker productivity in tasks, the model tries to measure the overall effectiveness of a worker in task performance.

3.3 Human Capital Accumulation

Human capital accumulation occurs within each task dimension and depends on the current occupation's task requirements. For each $\kappa \in \{C, M, I\}$, the accumulation is represented as follows.

$$H^\kappa_{i,t+1} = \begin{cases} H^\kappa_{i,t} + \beta^\kappa \gamma^\kappa_{o(it)} \exp(-\lambda E_{it}) & \text{if } o(it) > 0 \\ H^\kappa_{i,t} - d(H^\kappa_{i,t} - H^\kappa_{i,1}) & \text{if } o(it) = 0 \end{cases} \tag{2}$$

When worked, human capital in each task dimension—cognitive (H^C), manual (H^M), and interpersonal (H^I)—grows proportionally to the task intensity of the occupation the worker worked. Growth diminishes with overall experience, represented by an exponential decay function $\exp(-\lambda E_{it})$, where λ indicates the rate of diminishing returns and E_{it} is general experience. β represents the accumulation rate of each task-specific skill. It quantifies the rate at which kill increases, in proportion to task use.

When a worker does not work, task-specific human capital depreciates at a rate d, leading to a reduced stock of human capital for each dimension when transitioning to a period of unemployment. Only the human capital gained through labor market experience is subject to depreciation, as the initial skill level reflects the amount of skill that is gained prior to labor market.[13][14]

3.4 Occupation Choice

The discrete occupation choice is considered as an auxiliary model to account for worker selection. I am not fully structurally modeling this part by specifying the underlying utility from wage, preference, friction and so on. Rather, I use a Markov model with multinomial logit specification to approximate the occupation choice probability function. Workers are allowed to choose occupation based on previous occupation, experiences in each task, and the unobserved worker type. By modeling this way, I am able to directly address worker selection based on worker heterogeneity and experience.

There are several reasons for this approach. First, the primary focus of this paper is the wage model, while the reason for occupation choices per se is not my focus. Therefore, it is more crucial for the model to accurately capture occupation choice patterns than to pinpoint the precise impact of each variable on those choices. This simplification supports decomposition purposes while accounting for self-selection, without overcomplicating the model. Also, even if the model were structural and dynamic, given the same state variables, the choices would still summarize as a function of those variables. Thus, coefficients in the multinomial logit can be interpreted as reduced-form parameters, capturing the relationship between state variables and occupational outcomes, while allowing for some discrepancy due to the assumed additive effects of variables.[15][16]

[13]Robustness check on the way.

[14]When a worker is employed in an occupation that does not require a particular skill, that skill does not depreciate. However, in practice, none of the four aggregated occupations has zero task intensity in any skill.

[15]This discrepancy reflects the functional form simplification, yet even a structural model would likely contain some degree of misspecification error.

[16]The choice of a multinomial logit form here is based on functional form convenience rather than the

The choice set includes five options: not working (occupation 0) and four occupations (Number of occupations $K = 4$). The probability of worker i selecting occupation k at time t, is given by the following equation.

$$P(O_{it} = k | X_{it}) = \frac{\exp(X_{it}' \delta_k)}{1 + \sum_{k'=1}^{K} \exp(X_{it}' \delta_{k'})} \tag{3}$$

O_{it} represents the occupation choice at time t, and X_{it} is a vector of explanatory variables. This vector X_{it} includes indicators for the worker's previous occupation, experience in each task in cognitive, manual, and interpersonal dimensions, and dummy variables for unobserved worker type. X_{it} is specified as

$$\begin{aligned} X_{it} = \Big(& \mathbb{1}\{O_{i,t-1} = 1\}, \mathbb{1}\{O_{i,t-1} = 2\}, \cdots, \mathbb{1}\{O_{i,t-1} = K\}, \\ & \tilde{E}_{it}^{C}, \tilde{E}_{it}^{M}, \tilde{E}_{it}^{I}, \mathbb{1}\{\text{type}_i = 2\}, \mathbb{1}\{\text{type}_i = 3\}, 1 \Big)^{\top} \end{aligned}$$

where $\mathbb{1}\{O_{i,t-1} = j\}$ represents an indicator variable for the worker's previous occupation j. $\tilde{E}_{it}^{C}$ is adjusted cognitive experience, defined as

$$\tilde{E}_{it}^{C} = \frac{H_{it}^{C} - H_{i1}^{C}}{\beta_{H}^{C}}$$

which represents the cumulative sum of cognitive task intensities across all periods worked so far, adjusted for concavity and depreciation. The adjusted manual and interpersonal experience measures are defined similarly. $\mathbb{1}\{\text{type}_i = c\}$ indicates for the worker's type. This specification accounts for worker heterogeneity in overall mobility patterns while also capturing differences stemming from experience.[17]

assumption of the Independence of Irrelevant Alternatives (IIA). This choice is not restrictive in itself, as any probability can, in theory, be represented in logit form. The restriction arises from the assumption that the effects of X are additive. Additionally, the model's simplicity aligns with the study's goal of quantifying the contributions of each variable, rather than performing counterfactual exercises that rely on changing model parameters to observe dynamic effects. For this reason, when I perform a counterfactual exercise of changing occupational assignment, I look at the contemporaneous effect.

[17]An alternative approach could involve using only H values without type dummies, but this would be more restrictive, as it imposes a constant slope on both initial and accumulated skills.

The parameters δ_k capture occupation-specific effects on choice probabilities, allowing for varying influences across occupational alternatives.

At $t = 1$, the tasks experience is equal to zero. Thus, the explanatory vector X simplifies, including only the type dummies and an intercept term. The coefficients for this initial period are estimated separately.

Note that the parameters in wage equation do not influence the occupation model, and vice versa. This approach reflects that the occupation choice model does not rely on the mean wage as a foundation. By not restricting occupation choices solely to wage-based determinants, the model captures non-wage-driven mobility patterns and thereby fit the data more accurately. Later, I test whether the observed choices align with wage maximization. Additionally, the model remains flexible, with the only connection between wage and occupation model being the shared unobserved type.

How Current Occupation Affects Wages The current occupation o influences wages through three main channels as shown in Equation 5: an occupation-specific intercept α_o, the value of accumulated task-specific human capital weighted by task intensities γ_o, and the value of initial skills.

First, a change in occupation affects wages directly through the intercept α_o, which represents a fixed wage level for each occupation, independent of worker-specific factors. Shifting to an occupation with a higher (or lower) α_o results in an immediate adjustment in wages, reflecting the intrinsic wage differences across occupations.

Second, each occupation has distinct task intensities—γ_o^C, γ_o^M, and γ_o^I—that determine how cognitive, manual, and interpersonal human capital contribute to wages. A worker's accumulated human capital in each dimension, represented by $\sum_{\tau=2}^{t} \Delta H_{i\tau}^C$, $\sum_{\tau=2}^{t} \Delta H_{i\tau}^M$, and $\sum_{\tau=2}^{t} \Delta H_{i\tau}^I$, is weighted by the corresponding task intensity of current occupation, γ_o. Therefore, an occupation change may shift the importance of these dimensions, influencing wages based on the worker's accumulated human capital. For instance, if the new occupation o'

emphasizes cognitive skills more heavily ($\gamma_{o'}^{C}$ is high), a worker with substantial accumulated cognitive human capital will experience a relatively higher wage.

Third, initial skill H_{1i} contributes to wages in a way that depends on the task intensity parameters γ_o of the chosen occupation. This means that the effect of intrinsic worker ability on wages is occupation-dependent, as γ_o determines how much initial skill in each task dimension (cognitive, manual, or interpersonal) is valued in that occupation. A worker with a high initial skill in cognitive tasks (${H_1}_i^C$) would earn a premium in an occupation with high cognitive intensity (γ_o^C), while the same (${H_1}_i^C$) would contribute less in an occupation where cognitive skills are less emphasized. Thus, the wage impact of occupation o changes through three pathways: the occupation-specific intercept α_o, task-specific human capital accumulation weighted by γ_o, and the interaction of initial skill H_{1i} with task intensities γ_o.

4 Identification and Estimation

The model estimates several key parameters that capture the relationship between task-specific skill and wages. This section discusses those key parameters along with loose identification, and reports the parameter estimates.

Informal Identification of Human Capital Accumulation Rate The first set of key parameters are the accumulation rates of task-specific human capital, β. It quantifies how quickly skill in each task grow as experience in that task increases. For example, the accumulation rate of cognitive skill captures how profitable cognitive task experience is. Thus, β is mainly identified by comparing on-the-job wage growth across occupations with varying task intensities. By comparing the wage trajectories of workers who accumulate experience in occupations with different task profiles, the model identifies the contribution of task-specific human capital to wage growth, thus capturing the role of β in shaping wage dynamics across occupations.

Informal Identification of Initial Skill Distribution Another important set of parameters refers to the distribution of initial skills, represented by H_1. Identification of H_1 is achieved through within-worker wage comparisons as individuals transition between occupations with varying task profiles. If a worker experiences a higher wage gain when moving to a more cognitive skill-heavy occupation, the model infers that the worker possesses high initial skill in the cognitive dimension. This is upon the assumption that type is time-fixed. Thus, identifying the level of initial skill H_1 relies on variation in task intensities across occupations like β, but additionally requires enough occupation switches by workers (to be precise, enough variation of task intensity in those switched occupations). See Appendix B for an illustration of identification through a simple example.

The model is estimated using maximum likelihood estimation (MLE).[18] The likelihood function follows.

4.1 Likelihood Function

To construct the likelihood function, I begin by formulating the conditional likelihood given a worker's type.

Likelihood for Period $t = 1$, conditional on worker type The conditional likelihood accounts for the probabilities of observed wages and occupational choices, given the underlying worker type. Let $\Theta = (\alpha, \beta, \sigma_\epsilon, H_1, d, \lambda, \delta)$. The probability of observing O and w for individual i in period $t = 1$ for worker type c is the product of the probability of occupation choice and the probability of wages from the normality assumption:

[18] I find the maximum by trying 20 different starting points and use MATLAB (fminsearch).

$$
\begin{aligned}
L_{i1}(\Theta, O_{i1}, w_{i1}; H_{1i} = H_{1c}) = & \left(\frac{1}{1 + \sum_{k'=1}^{K} \exp(X'_{itc}\delta_{k'})}\right)^{I(O_{i1}=0)} \\
& \times \prod_{k=1}^{K} \left(\frac{\exp(X'_{itc}\delta_k)}{1 + \sum_{k'=1}^{K} \exp(X'_{itc}\delta_{k'})}\right)^{I(O_{i1}=k)} \\
& \times \prod_{k=1}^{K} \left(\frac{1}{\sqrt{2\pi\sigma_\epsilon^2}} \exp\left(-\frac{(w_{i1} - \gamma'_k H_{1c})^2}{2\sigma_\epsilon^2}\right)\right)^{I(O_{i1}=k)}
\end{aligned}
$$

Likelihood for Period $t \geq 2$, conditional on worker type For period 2 or later, The probability of observing O, w for individual i in period $t = 2$ given unobserved type of worker is:[19]

$$
\begin{aligned}
L_{it}(\Theta, O_{it}, w_{it}; H_{1i} = H_{1c}) = & \left(\frac{1}{1 + \sum_{k'=1}^{K} \exp(X'_{itc}\delta_{k'})}\right)^{I(O_{it}=0)} \\
& \times \prod_{k=1}^{K} \left(\frac{\exp(X'_{itc}\delta_k)}{1 + \sum_{k'=1}^{K} \exp(X'_{itc}\delta_{k'})}\right)^{I(O_{it}=k)} \\
& \times \prod_{k=1}^{K} \left(\frac{1}{\sqrt{2\pi\sigma_\epsilon^2}} \exp\left(-\frac{(w_{it} - \alpha_k - \gamma'_o H_{it})^2}{2\sigma_\epsilon^2}\right)\right)^{I(O_{it}=k)}
\end{aligned}
$$

Combined Likelihood for All Periods Thus, the combined likelihood for worker i considering all T_i observations and three possible worker types is:

$$
\begin{aligned}
L_i(\Theta | O_i, w_i) = & p_1 \prod_{t=1}^{T_i} L_{it}(\Theta, O_{it}, w_{it}; c = 1) \\
& + p_2 \prod_{t=1}^{T_i} L_{it}(\Theta, O_{it}, w_{it}; c = 2) \\
& + (1 - p_1 - p_2) \prod_{t=1}^{T_i} L_{it}(\Theta, O_{it}, w_{it}; c = 3)
\end{aligned}
$$

[19] When computing likelihoods, I take T_i for each individual as given.

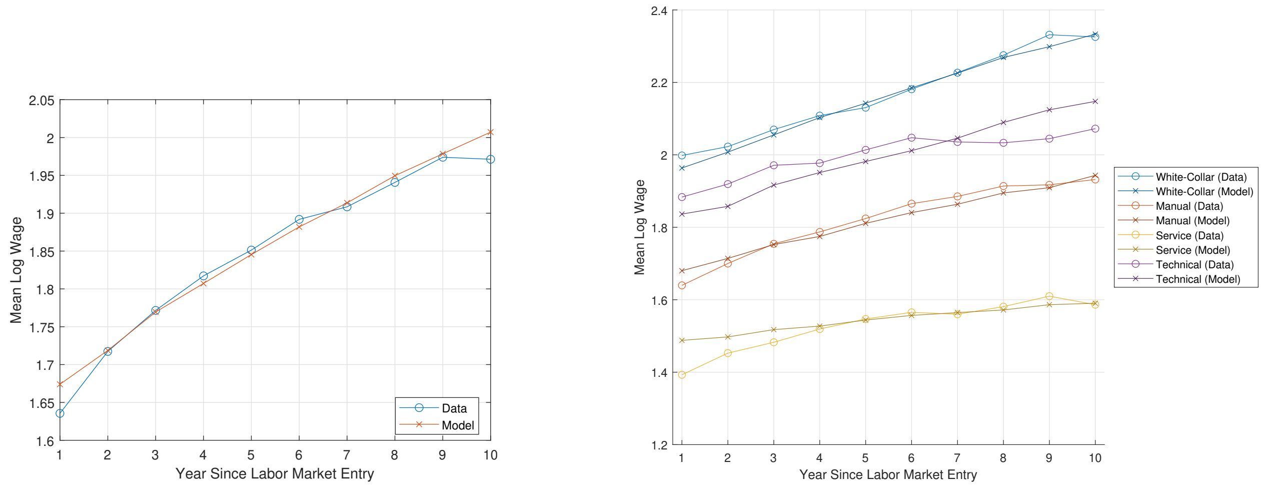
Combined Log-Likelihood for All Individuals and Periods The final likelihood function used for estimation combines all individuals and periods.

$$\log L(\Theta|O, w, X) = \sum_{i=1}^{n} \log L_i(\Theta, O, w, X) \tag{4}$$

4.2 Model Fit

To evaluate the fit of my model, I simulate data for a sample of $N = 50,000$ individuals based on the estimated parameters. First, I draw each individual's unobserved type based on estimated probabilities for each type, where the probability distribution is $p(c = 1) = \hat{p_1}$, $p(c = 2) = \hat{p_2}$, and $p(c = 3) = 1 - \hat{p_1} - \hat{p_2}$. Then, for each simulated individual, I draw independent and identically distributed shocks for wages and occupation choices. Based on these shocks, occupation choices and wage trajectories are simulated for 14 periods (the median number of observations for workers with high school degree). Finally, I compare the moments computed from the simulated data to the empirical moments from observed data.

Moments used for evaluation are presented in graphs. Figure 1 illustrates the alignment of model-generated moments with the data. Figure 1a shows that the average wage level over the first ten years in the labor market fits well, though the model overestimates wages in periods 1 and 10. The wage profiles by occupation, as depicted in Figure 1b, illustrate the evolution of average wages across occupations over time. Specifically, for each year since labor market entry, the figure presents the mean log wage of workers within each occupation. While the model's wage profiles match reasonably well for most occupations, the fit is less accurate for Technical occupation. This likely arises because technical occupation have high task intensity but slower wage growth. In the model, occupations with similar task intensities are assumed to experience similar wage growth due to human capital accumulation (except for worker composition changes), which restricts the fit for technical occupation. Additionally, starting wages in Service occupations are overestimated, whereas initial wages in White-Collar occupations are underestimated.



(a) Wage profile

(b) By occupation

Figure 1: Fit of the model (wage). The data wage is directly calculated from observed data, while the model wage is simulated using the estimated model with $N = 50,000$. The figures depict wage profiles over years since labor market entry. Panel (a) presents the average logged wage of workers at each year, and Panel (b) illustrates the average logged wage of workers within each occupation at each year.

The fit of the model for the occupation choices is shown in Figure 2. Each circle represents the fraction of workers in the corresponding occupation at a given period, with periods labeled by numbers on the graph. The model demonstrates a strong fit to the data, owing to its flexible specification. By incorporating distinct coefficients for both unobserved worker types and accumulated human capital in the wage and occupation models, the framework effectively captures the data.

4.3 Estimates

This section reports the estimates, with empirical confidence intervals obtained through block bootstrapping for 2000 times.

Wage and Parameters Table 5 provides estimates for each occupation's intercepts and the variance of wage shock, as described in Equation 1.

The intercepts are fairly similar across occupations, with White-Collar and Manual occupations having higher intercepts than Service occupations. Although the standard errors are

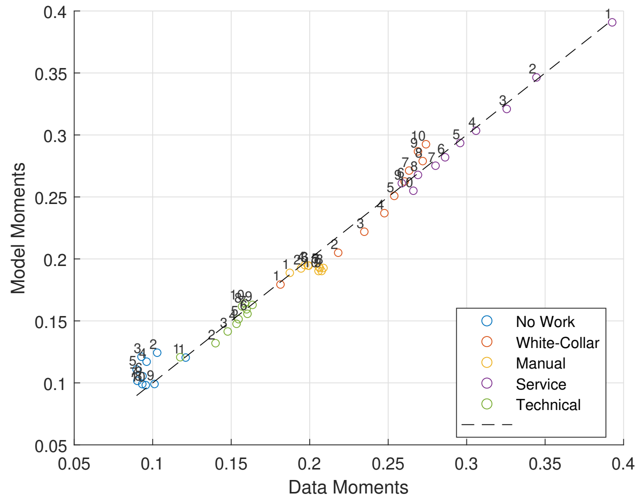


Figure 2: Fit of the model (occupation choice). Data moments are calculated directly from observed data, while model moments are simulated using the estimated model with $N = 50,000$. The dashed line represents the 45-degree line, indicating perfect alignment between model moments and data moments. Each circle represents the fraction of workers in a given occupation (including non-employment) from year 1 to year 10, with periods labeled by their corresponding numbers in the figure.

large, the intercept for Technical occupations is slightly lower than that for Service, despite the average wage level being higher in Technical occupations as in Figure 1b.[20] The standard deviation of the wage shock, σ_ε, is estimated at 0.43, suggesting moderate wage variability. This represents the variation due to factors outside the model, such as firm-specific human capital. In the model,

Human Capital Accumulation Parameters Table 6 displays estimates for the depreciation rate d and the concavity parameter λ from Equation 2, which influence the dynamics of human capital accumulation. The depreciation rate is estimated to be 0.09, aligning closely with the findings of Fan et al. (2024). Meanwhile, the concavity parameter is estimated at 0.07, as my sample consists of workers in the early stages of their careers—up to about 15 years since labor market entry—during which wages tend to be almost linear in experience.

The estimates for β indicate a relatively high rate of accumulation for cognitive and

[20]This is not contradictory; in models such as Gibbons and Waldman (1999), where workers sort into occupations based on ability, higher-paying occupations tend to have lower intercepts, as these positions attract higher-ability individuals, while lower-paying occupations have higher intercepts, explaining selection by workers with lower ability.

Table 5: Wage Parameters

Occupation Intercepts α	Estimate	Confidence Interval
White-Collar	1.22	[0.21, 1.84]
Manual	1.23	[-0.37, 1.52]
Service	1.02	[0.31, 1.38]
Technical	0.97	[-0.21, 1.47]
σ_ε	0.43	[0.41, 0.45]

Notes: The 95% confidence intervals are estimated using bootstrapped samples.

Table 6: Human Capital Accumulation Parameters

	Estimate	Confidence Interval
Depreciation Rate d	0.09	[0.05, 0.30]
Concavity λ	0.07	[0.05, 0.10]
Skill Accumulation Rate β		
Cognitive	0.0114	[0.0001, 0.0154]
Manual	0.0130	[0.0100, 0.0158]
Interpersonal	0.0040	[0.0000, 0.0353]

Notes: The 95% confidence intervals are estimated using bootstrapped samples.

manual human capital but a slower pace for interpersonal human capital. While workers differ in their initial levels of interpersonal skill, they do not experience significant wage increases from gaining additional experience in interpersonal tasks. However, several points are worth noting.

First, the magnitude of β alone is less meaningful, as β is interacted with γ in the human capital accumulation and the scale of γ differs across task dimensions. For example, manual task intensities have smaller scale than cognitive does; thus, even if skill accumulation rate is similar, the human capital will be growing faster in cognitive dimension. In the decomposition exercises, I focus on the combined effects of β, γ, and H. Second, the accumulation rate of cognitive and interpersonal skill is less precisely estimated. Also, as β is constrained to be non-negative, and since the estimated parameter is close to the boundary, the bootstrap standard errors may not be consistent, as shown by Andrews (2000).

Table 7: Heterogeneity Parameters

	CogInit	**ManInit**	**IntInit**	**Prob**
Type 1	0	0	0	0.20
CI	-	-	-	[0.03, 0.65]
Type 2	-0.61	0.10	1.38	0.56
CI	[-2.50, -0.11]	[-1.14, 1.10]	[-0.34, 4.41]	[0.03, 0.65]
Type 3	-1.12	0.20	2.65	0.24
CI	[-3.63, -0.44]	[-1.36, 1.24]	[0.10, 5.79]	[0.02, 0.36]

Notes: **CogInit**: initial cognitive skill level, **ManInit**: initial manual skill level, **IntInit**: initial interpersonal skill level for each type. **Prob**: Probability mass assigned to the type. The 95% confidence intervals are estimated using bootstrapped samples.

Type Parameters The type support parameters, shown in Table 7, are less precisely estimated. As discussed at the beginning of the section on identification, the estimation of type parameters requires enough variation in task intensities and as well as number of individuals who switch occupations, making them harder to be identified compared to the β parameters. The estimates suggest different levels of initial skills for each type, with Type 1 having the highest initial cognitive skill and lowest manual and interpersonal skill, Type 2 moderate initial cognitive skill, and Type 3 lowest initial cognitive skill and highest interpersonal skill. However, interpreting these results is challenging; for instance, Type 1 workers are expected to earn lower wages in cognitive skill-intensive occupations compared to Type 2 or 3 workers when the estimates are plugged in for $t = 1$, due to lower initial interpersonal skill and high correlation between interpersonal and cognitive task intensity. This likely stems from poor identification, so in the decomposition exercise, I report the combined effects of cognitive and interpersonal initial skills rather than separating them.

In terms of initial manual skill, there is less variation across types compared to other skills, which aligns with the intuition that manual skills are rather accumulated and initial skill matters less than in cognitive or interpersonal dimension. The last column of Table 7 shows the probabilities for each type, with Type 2 being the most common at 56%.

The coefficients for the multinomial logit model are reported in Table A1.

5 Decomposition Results

In this section, I apply the estimated model to quantify the impact of each task-specific skill on wage growth and wage variance. The decomposition is clarified through the rewritten wage equation:

$$\begin{aligned}\log(w_{iot}) = \; & \alpha_o + \gamma_o^C \left(H_{1i}^C + \sum_{\tau=2}^{t} \Delta H_{i\tau}^C \right) + \gamma_o^M \left(H_{1i}^M + \sum_{\tau=2}^{t} \Delta H_{i\tau}^M \right) \\ & + \gamma_o^I \left(H_{1i}^I + \sum_{\tau=2}^{t} \Delta H_{i\tau}^I \right) + \epsilon_{it} \end{aligned} \tag{5}$$

This representation emphasizes that wages consist of three components: an occupation-specific factor, initial skill, and accumulated skill.

5.1 Wage Growth

I begin by analyzing the contribution of each component to wage growth over time. I compare the average wage level at year 10 to the year of labor market entry, and decompose the differences into accumulated human capital components, initial human capital and occupation level components.

$$\begin{aligned}\mathbb{E}[\Delta \log w] = \; & \underbrace{\mathbb{E}\left[\Delta(\gamma^C \textstyle\sum \Delta H^C)\right]}_{\textbf{Accumulated Cog Skill}} + \underbrace{\mathbb{E}\left[\Delta(\gamma^M \textstyle\sum \Delta H^M)\right]}_{\textbf{Accumulated Man Skill}} + \underbrace{\mathbb{E}\left[\Delta(\gamma^I \textstyle\sum \Delta H^I)\right]}_{\textbf{Accumulated Int Skill}} \\ & + \underbrace{\mathbb{E}\left[\Delta\left(\alpha_o + \gamma^C \psi^C + \gamma^M \psi^M + \gamma^I \psi^I\right)\right]}_{\textbf{Occupation Intercept \& Initial Skill}} \end{aligned} \tag{6}$$

$\Delta \log w$ represents the change in log wages, with each term capturing the contribution of accumulated cognitive, manual, and interpersonal human capital ($\sum \Delta H^C$, $\sum \Delta H^M$, $\sum \Delta H^I$) and the influence of initial skills (ψ^C, ψ^M, ψ^I) as weighted by occupation-specific factors γ. Because the initial skill is time-fixed in the model, the ability contribution can-

Table 8: Wage Growth Decomposition

Component	**Contribution (%)**
Human Capital Accumulation:	
Cognitive	63.00
Manual	25.32
Interpersonal	8.58
Initial Skill + Occupation Intercepts	3.09
Total	**100**

Notes: The logged wage growth between $t = 10$ and $t = 1$ is broken down into the contributions from accumulated human capital (cognitive, manual, and interpersonal) and the combined effects of initial skills and occupation intercepts.

cels out for a worker keeping the same occupation. However, when workers move across occupations, the value of initial skill as well as occupation intercepts change.

The expectations in these decompositions are estimated by sample averages of simulated wages, with $N_{\text{sim}} = 2000$.[21]

The difference in mean log wages between $t = 10$ and $t = 1$ ($\overline{\log w_{10}} - \overline{\log w_1}$) is 0.32. Table 8 breaks down this change. 63.00% of the wage growth is due to cognitive human capital accumulation, and 25.32% is due to manual, only 8.58% is due to interpersonal human capital accumulation. The remaining 3.09% is attributed to the combined effects of cognitive, manual, and interpersonal initial skill, along with occupation intercepts.

As expected, wage growth is primarily driven by accumulated human capital, but with cognitive skills playing the most substantial role. The contributions from accumulated manual skill is modest and that of and accumulated interpersonal human capital are significantly smaller, due to slower growth of interpersonal skills.

Although initial skill levels remain fixed over time, the value of initial skills changes as the occupation changes, thus contributing to wage growth. This suggests that occupation

[21] For each simulation, I draw the same number of workers (4,041 workers) as in the data. First, each worker's unobserved type (from types 1 to 3) is sampled based on their Bayesian posterior probabilities. Next, I simulate the occupation, wage panel, and human capital levels for the observed number of periods for each worker in the data. Finally, I calculate averages using 4,041 workers per simulation across N_{sim} simulations.

mobility plays a role in wage gains as workers transition to occupations that better align with their initial skill sets. Consequently, while accumulated human capital is the primary driver, occupational shifts enhance wage growth by allowing workers to capitalize on their initial skills within more suitable occupational contexts.

5.2 Wage Variance

The purpose of the decomposition is to analyze how different factors contribute to observed wage variation across agents at 10 years since labor market entry. Following Equation 5, the variance of log wages, $var(\log w)$, is decomposed into parts associated with occupation-specific intercepts, accumulated task-specific human capital, and initial human capital, as follows:

$$
\begin{aligned}
var(\log w) = \underbrace{cov(\log w, \alpha_o)}_{\textbf{Occupation Intercept}} &+ \underbrace{cov(\log w, \gamma^C H_1^C)}_{\textbf{Initial Cog Skill}} + \underbrace{cov(\log w, \sum \Delta H^C)}_{\textbf{Accumulated Cog Skill}} \\
&+ \underbrace{cov(\log w, \gamma^M H_1^M)}_{\textbf{Initial Man Skill}} + \underbrace{cov(\log w, \sum \Delta H^M)}_{\textbf{Accumulated Man Skill}} \\
&+ \underbrace{cov(\log w, \gamma^I H_1^I)}_{\textbf{Initial Int Skill}} + \underbrace{cov(\log w, \sum \Delta H^I)}_{\textbf{Accumulated Int Skill}} \\
&+ \underbrace{var(\epsilon)}_{\textbf{Error}} \qquad (7)
\end{aligned}
$$

The decomposition highlights the contribution of occupation intercepts to wage variance, reflecting the role of occupation-specific factors. Additionally, each skill dimension—cognitive, manual, and interpersonal—has distinct contributions to wage variance through both accumulated human capital ($\sum \Delta H^C$, $\sum \Delta H^M$, $\sum \Delta H^I$) and initial human capital level (H_1^C, H_1^M, H_1^I). These components capture how differences in accumulated skills and innate abilities in each dimension affect wage outcomes.

In practice, because of the previously mentioned identification issue in initial skill level, I

Table 9: Cross-sectional Wage Variance Decomposition

Component	**Contribution (%)**
Occupation Intercept	4.31
Skill Accumulation:	
Cognitive	10.10
Manual	-0.12
Interpersonal	1.13
Initial Skills:	
Cognitive and Interpersonal	34.42
Manual	4.98
Error	45.17
Total Var$(w) = 0.4042$	**100**

Notes: Cross-sectional wage variation at year 10 is decomposed into components. Var(w) represents the explained variance in logged wages. Accumulation components reflect the task-specific human capital accumulated over the career, while initial components capture the impact of initial skills at labor market entry. Occupation Intercept represents fixed effects associated with occupations. Refer to the text for details on decomposition methods.

combine cognitive and interpersonal initial skill effects, and decompose the variation across workers at 10 years since labor market entry. Covariances in these decompositions are computed from simulated wages, as explained in the previous section about wage growth decomposition.

The decomposition shown in Table 9 indicates that 11.11% of wage dispersion is attributed to human capital accumulation, with the majority of this effect stemming from cognitive human capital. This finding suggests that skill acquisition in cognitive tasks plays a particularly significant role in explaining wage variance, potentially due to fast-growing cognitive skills across various occupations. In contrast, contributions from manual and interpersonal human capital appear less substantial. The covariance with manual skill accumulation, though close to zero, is negative. This indicates a potential negative relationship between wage levels and manual experience, which could be attributed to the possibility that workers with extended manual experience tend to be in lower-paying occupations. Thus differences in manual and interpersonal skill accumulation explain only small part of wage variation.

A larger portion of 43.71% of wage dispersion arises from the combined effects of occupation intercepts and initial skill. Even after a decade in the labor market, covariance with initial skills continue to account for a significant portion of wage dispersion. The importance of initial manual skill is much smaller than that from cognitive or interpersonal initial skills, suggesting that there is less variation in innate skill in manual dimension across workers, compared to cognitive or interpersonal abilities. The importance is similar to that of occupation intercepts, such as occupation-level characteristics that influence pay scales.

The remaining portion of the wage variance stems from unexplained factors captured in the error term. This is big, but comparable to literature. These unexplained factors may reflect the effects of firm-specific human capital, as firm switch is not considered in the model and all the wage losses not due to task requirements are attributed to error term.

To conclude, wage inequality in early careers is primarily driven by initial skill and occupation-specific factors. However, the impact of accumulated human capital, particularly in the cognitive dimension, is also significant. In contrast, manual and interpersonal human capital contribute minimally to wage variation across workers.

Effects of Observed Occupation Returning to the wage equation, the assignment of workers to different occupations plays a critical role in determining their wages. Specifically, the current occupation characteristics, denoted by α_o and occupation-specific task weights γ_o^C, γ_o^M, and γ_o^I, directly influence wage outcomes. These parameters interact with each worker's accumulated human capital and initial human capital level in cognitive, manual, and interpersonal tasks, respectively. This means that the current occupation choice affects significantly to relative importance of each tasks.

Moreover, in reality, workers' occupation choices are not always optimized to maximize wages and that is captured by the model, as it fits the occupation choices separately from

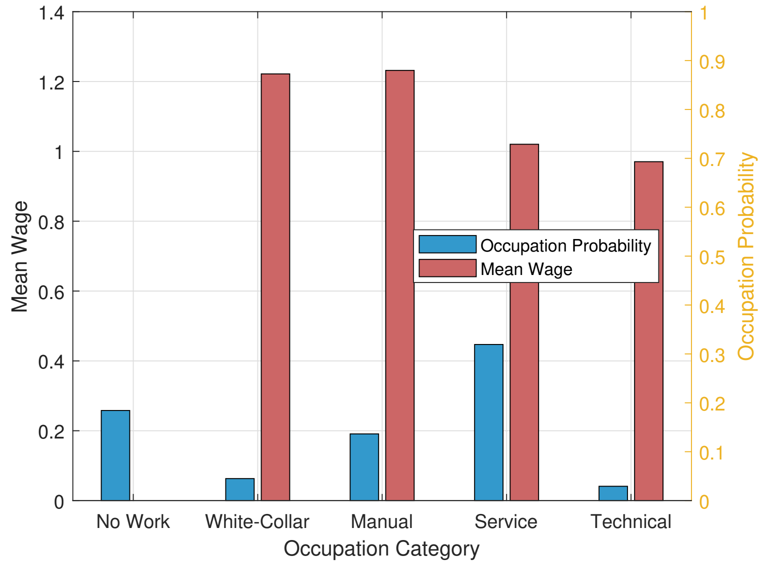


Figure 3: Occupation choice probability and mean wage of worker type 1 at period 1. The mean logged wage is plotted on the left axis, and occupation choice probabilities are shown on the right axis.

wages. For instance, Figure 3 shows that worker type 1, at $t = 1$, would earn the highest mean wage in the Manual occupation; however, their most likely chosen occupation is Service. This discrepancy suggests factors beyond wage maximization may influence occupation choice. While the model does not provide direct explanations about why, potential reasons include non-monetary preferences for certain occupations or frictions in occupational mobility, which may restrict workers' ability to transition to the highest-paying occupations. See Figure A2 for all three worker types.

Hypothetical Occupation Choice Scenarios and Wage Variation To further understand the role of observed occupation assignments in wage variation, I compare wage outcomes under two alternative scenarios: (1) a wage-optimal scenario, where workers choose the occupation that maximizes their mean wage at current period, and (2) a random scenario, where occupation choices are entirely random, independent of any factors such as previous occupation, accumulated human capital, or worker type and the choice probabilities equal the mean fractions of each occupation.

In the wage-optimal scenario, each worker myopically selects the occupation with the highest mean wage, conditional on working. This scenario assumes that workers have complete information about potential earnings in each occupation and can move freely to the highest-paying option. At $t = 10$, using the previous simulation for variance decomposition, I calculate the mean wage level for each worker based on their simulated H_{i10} levels across occupations. I then compute the probability of each worker not working at $t = 10$, given their current state variables. The wage-optimal probability of working in the highest-paying occupation is set to $1-$ (probability of not working). Using this updated occupation choice probability and the same random draw of occupation choice shocks from the baseline model, I then simulate each worker's occupation. That is, the occupation choice model is used solely to determine the probability of not working in this scenario.

The random scenario treats all workers the same and occupations are assigned randomly irrespective of worker type or previous occupation history. This setup provides a counterfactual where occupations are determined independently of worker preferences, abilities, or accumulated human capital. At $t = 10$, based on the baseline simulation for variance decomposition, I calculate the fraction of workers in each occupation (from $O = 0$ to $O = 4$). This fraction then defines the occupation choice probability for all workers in the random scenario. Using these uniform probabilities and the baseline choice shocks, I simulate each worker's occupation.

For both scenarios, when adjusting the occupation assignments at period $t = 10$, I keep the baseline level of accumulated human capital fixed at $t = 10$. This allows for a consistent comparison of hypothetical wage variation due to different occupation assignments. By analyzing wage outcomes under each scenario, I explore the extent to which observed wage inequality can be attributed to the influence of occupation assignments.

Table 10 shows the decomposition of wage dispersion components under three different scenarios: the Baseline scenario, the Random assignment scenario, and the Wage-Optimal assignment scenario. In the Baseline scenario, workers' occupation choices are based on the observed data; in the Random scenario, occupation choices are independent of individual

Table 10: Wage Dispersion Components Under Different Occupation Scenarios, $t = 10$

Components	**Baseline**	**Random**	**Wage-Optimal**
Var(w) - Var(ε)	0.2216	0.1811	0.1748
Accumulated Cog	18.42%	13.75%	13.91%
Accumulated Man	-0.23%	0.44%	1.37%
Accumulated Int	2.08%	1.60%	1.66%
Initial Cog + Int	62.70%	66.80%	80.14%
Initial Man	9.08%	12.26%	13.85%
Occ Intercept	7.85%	5.14%	-10.87%

Notes: Cross-sectional wage variation at year 10 is decomposed into components. Var(w) - Var(ε) represents the explained variance in wages after accounting for residual noise (ε). Accumulated components reflect the task-specific human capital accumulated over the career, while initial components capture the impact of initial skills at labor market entry. Occ Intercept represents fixed effects associated with occupations. The scenarios compare baseline observed occupation assignments with random assignment, and wage-optimal assignment at period $t = 10$.

characteristics and task-specific human capital; and in the Wage-Optimal scenario, workers select occupations that maximize their mean wage.

First, both the Random and Wage-Optimal scenario reduces wage variation[22]. As expected, the average wage level is lower in the Random scenario and higher in the Wage-Optimal scenario as shown in Table A2. Second, Random assignment is significantly different from Baseline and it underestimates the importance of cognitive human capital yet overestimates importance of manual human capital. This outcome highlights the potential bias if the non-random nature of occupation choices is ignored, suggesting the importance of modeling occupation choice.

Notably, the contribution of accumulated cognitive human capital to wage variation is reduced in the Wage-Optimal scenario, accounting for 13.91% compared to 18.42% in the Baseline. This suggests that wage optimization in occupation choice lessens the role of cognitive skills in explaining wage differences.

[22]Notice that both are partial equilibrium analysis since I am assuming that workers can freely choose occupation without affecting the other workers. A closer to general equilibrium exercise will be the one in which population occupation fraction is preserved but sum of wages are maximized by a social planner. However, since this optimization requires solving integer problem, it is numerically hard to solve.

Accumulated cognitive skills and combined cognitive-interpersonal initial skills dominate wage variation across all scenarios, but the contribution of initial manual skill increases under the Wage-Optimal scenario, rising to 13.85% compared to 9.08% in the Baseline. This indicates that under a wage-maximizing occupation assignment, initial manual skills have a greater impact on wage variation, potentially reflecting increased wage opportunities for manual skills when occupation choice aligns with maximizing earnings. The covariance between occupation intercepts and wage decreases at the same time. This is likely due to workers choosing Technical occupation to maximize wages, who would have otherwise chosen service occupations in the baseline scenario. This suggests that workers may have non-wage motives or face restrictions in choosing observed occupations, which in turn reduces the importance of manual skills under those occupational choices.[23]

5.3 The Wage Impact of Skill Accumulation Across Education Levels

In this subsection, I use the model to answer the following question. How do workers with different demographics accumulate human capital differently, and what is the impact on wages?

Since education and gender, among other individual characteristics, do not enter the model directly, the differences in occupation choice and initial skill level due to these time-fixed individual characteristics are all attributed to the unobserved worker type. Table 11 shows how workers with different demographics are considered differently within model in terms of unobserved type. With the estimated parameters, I first compute for each worker, the posterior probability of being each type, using the empirical Bayes estimator. The table shows that there is substantial differences in the probability of being each type, across gender and final education level. I compare the group whose final degree is high school

[23] An informative exercise would be to examine which occupations are chosen more frequently and identify the workers who change their occupation choices under the wage-optimal scenario.

Table 11: Mean Probability of Each Type by Gender and Education Level with Observations

Group	Probability of Each Type			Observations
	C=1	C=2	C=3	
Male, High school or Less	0.22	0.59	0.19	1,458
Male, College Graduate	0.06	0.44	0.50	609
Female, High school or Less	0.33	0.60	0.07	1,182
Female, College Graduate	0.08	0.55	0.37	792

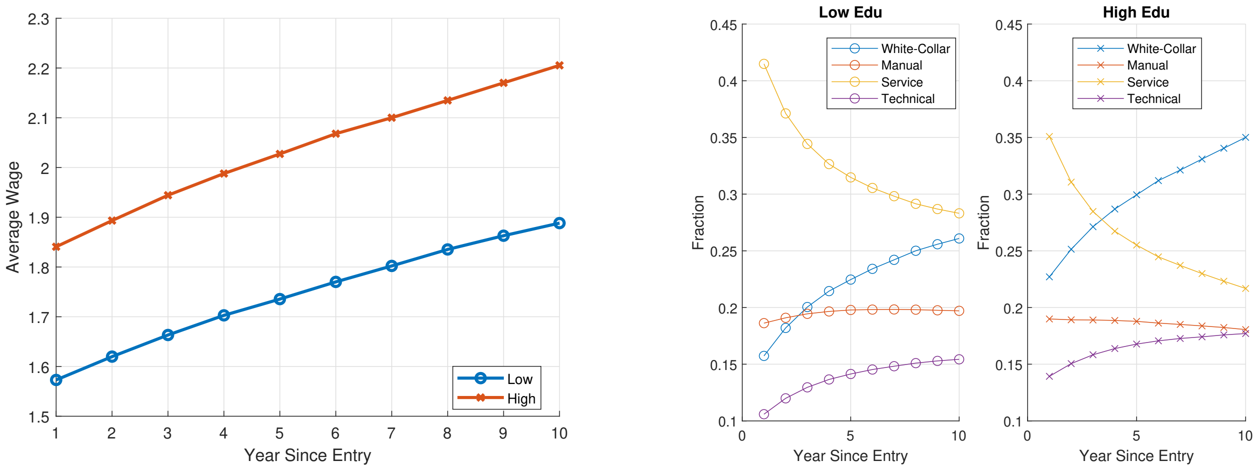
Notes: For each group defined by gender and final education level, the table displays the posterior probability of belonging to each unobserved type based on the observed data. Education levels are categorized by final degree attained: high school or less, and college (including 2-year and 4-year degrees).

or less to those who have obtained a two-year or four-year college degree.[24] While gender differences exist within both education levels, the differences across education levels is even more pronounced.

Wage Gaps Between Educational Attainment Groups I analyze the extent to which wage gaps between workers with a high school diploma or less and those with a college degree can be attributed to differences in human capital accumulation, initial skill levels, and occupational choices. In Figure 4, panel (a) displays the wage profiles for the two educational groups using simulated data. A wage gap is evident from the first year of labor market entry, and this gap increases slightly over the first ten years. Panel (b) illustrates the occupational choices of workers in each group, showing the fraction of workers in each occupation relative to the total number of workers employed at each period, averaged across simulation.

The occupational progression patterns are similar across the two groups: workers tend to transition from Service to White-Collar and Technical occupation, while the proportion in Manual occupation remains relatively stable. However, key differences emerge. Workers with a high school diploma or less are more likely to begin in Service occupations, whereas college-educated workers are more likely to start in White-Collar. After five years in the labor market, White-Collar occupations become the most common for college-educated workers.

[24] As explained in sample restriction, workers who has graduate degree are dropped from the sample.



(a) Wage Profile

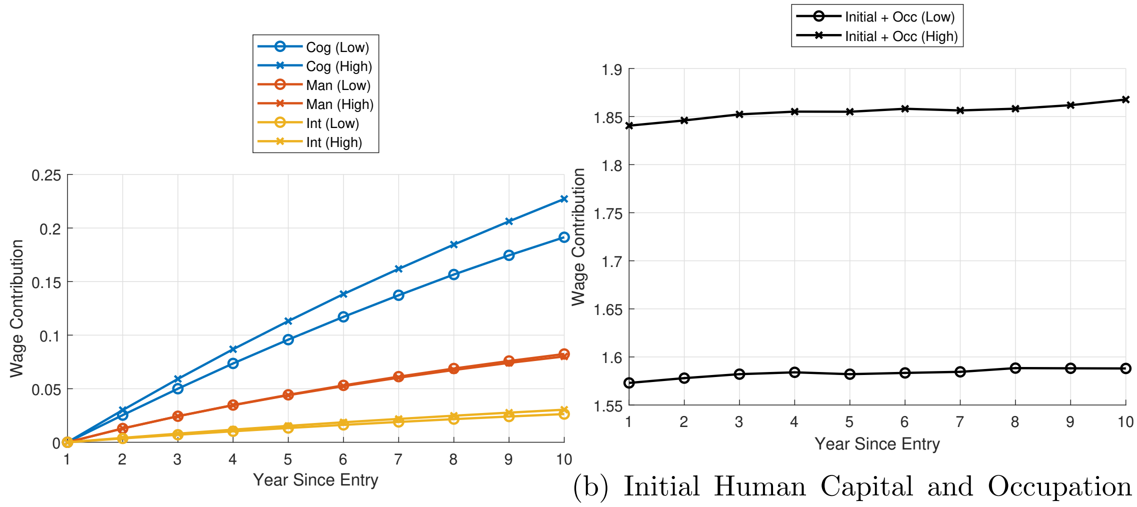
(b) Fraction of Workers in Occupation

Figure 4: Wage Profiles and Occupation Mobility by Education. Workers are divided into two groups based on their final degree: the Low group includes those with a high school diploma or less, while the High group includes those with a college degree (2-year or 4-year). Wages and occupation choices are derived from simulations ($N_{\text{sim}} = 2000$). Panel (a) displays the average logged wage across periods, while Panel (b) illustrates the fraction of workers in each occupation at each period, excluding non-employment.

Figure 5 illustrates the changes in each component contributing to wage. For simplicity, I divide wage into two parts: (1) accumulated human capital and (2) a combination of initial human capital levels and occupational intercepts. Panel (a) depicts the wage contributions from the accumulation of cognitive, manual, and interpersonal human capital, while Panel (b) shows those from initial skill levels and occupational intercepts.

Panel (b) highlights that the wage gap is prominent from the first period, reflecting differences in initial skill levels and occupational intercepts. The wage increases from year 1 to year 10 are quite similar across educational groups, driven by workers increasingly sorting into White-Collar and Technical occupations. These changes influence both the occupational intercepts and the value of initial skill levels.

Panel (a) reveals that, according to the estimated model, the wage contribution of human capital accumulation in manual and interpersonal skills is nearly identical on average, across the educational groups. However, the contribution of cognitive skill accumulation increasingly diverges over time.



(a) Human Capital Accumulation

(b) Initial Human Capital and Occupation Intercepts

Figure 5: Decomposition of Mean Wage into Accumulated Human Capital and Initial Ability Contribution. Workers are grouped by final degree: Low (high school or less) and High (2-year or 4-year college degree). Wages and occupation choices are derived from simulations ($N_{\text{sim}} = 2000$). Panel (a) illustrates wage contributions from human capital accumulation across skill dimensions. Panel (b) depicts contributions from initial skills and occupation intercepts combined.

While the overall contribution of human capital accumulation to wages is smaller compared to that of initial skill levels, the widening wage gap stems primarily from differences in the accumulation of human capital, rather than solely from variations in chosen occupations.

6 Conclusion

This study highlights the critical role of cognitive human capital accumulation in both wage variance and wage growth, suggesting that investments in cognitive skills can yield significant economic returns. Accumulated cognitive human capital consistently contributes to wage inequality and the increase in the mean level over time, underscoring its importance as a driver of wage dynamics.

In contrast, while manual human capital accumulation does support wage growth, its effect on wage variance is minimal, indicating that experience in manual tasks is less influential in differentiating workers' wages across occupations.

Interpersonal human capital, on the other hand, accumulates more slowly than the other two skills. Thus, additional experience in this dimension contributes significantly less to wage growth and also plays a smaller role in explaining the wage inequality. Although initial skills in interpersonal tasks are important and explain wage differences between workers, time spent on interpersonal tasks in the labor market has limited economic value in terms of wage gains. Indeed, initial skill and occupation-level wage differentials play a substantial role in explaining wage variance.

These findings emphasize the need for policies and interventions that consider the differential human capital assumption in different skills. In addition, the paper suggests the importance of promoting the development of cognitive skills.

References

J. Adda and C. Dustmann. Sources of Wage Growth. *Journal of Political Economy*, 131 (2):456–503, Feb. 2023. ISSN 0022-3808, 1537-534X. doi: 10.1086/721657. URL https://www.journals.uchicago.edu/doi/10.1086/721657.

D. W. K. Andrews. Inconsistency of the Bootstrap When a Parameter is on the Boundary of the Parameter Space. *Econometrica*, 68(2):399–405, 2000. ISSN 0012-9682. URL https://www.jstor.org/stable/2999432. Publisher: [Wiley, Econometric Society].

J. D. Beard, M. A. Verdeja, D. A. Bonsrah, S. D. Westfall, A. L. Steege, and M. K. Schubauer-Berigan. Crosswalks to Convert US Census Bureau Industry and Occupation Codes, 1980–2018. *Epidemiology*, 33(2):e8, Mar. 2022. ISSN 1044-3983. doi: 10.1097/EDE.0000000000001440. URL https://journals.lww.com/epidem/fulltext/2022/03000/crosswalks_to_convert_us_census_bureau_industry.20.aspx.

D. J. Deming. The Growing Importance of Social Skills in the Labor Market*. *The Quarterly Journal of Economics*, 132(4):1593–1640, Nov. 2017. ISSN 0033-5533, 1531-4650.

doi: 10.1093/qje/qjx022. URL https://academic.oup.com/qje/article/132/4/1593/3861633.

D. J. Deming. Why Do Wages Grow Faster for Educated Workers?, June 2023. URL https://www.nber.org/papers/w31373.

X. Fan, A. Seshadri, and C. Taber. Estimation of a Life-Cycle Model with Human Capital, Labor Supply, and Retirement. *Journal of Political Economy*, 132(1):48–95, Jan. 2024. ISSN 0022-3808, 1537-534X. doi: 10.1086/726232. URL https://www.journals.uchicago.edu/doi/10.1086/726232.

C. Gathmann and U. Schönberg. How General Is Human Capital? A Task-Based Approach. *Journal of Labor Economics*, 28(1):1–49, 2010. ISSN 0734-306X. doi: 10.1086/649786. URL https://www.jstor.org/stable/10.1086/649786. Publisher: [The University of Chicago Press, Society of Labor Economists, NORC at the University of Chicago].

R. Gibbons and M. Waldman. A Theory of Wage and Promotion Dynamics Inside Firms*. *The Quarterly Journal of Economics*, 114(4):1321–1358, Nov. 1999. ISSN 0033-5533. doi: 10.1162/003355399556287. URL https://doi.org/10.1162/003355399556287.

F. Guvenen, B. Kuruscu, S. Tanaka, and D. Wiczer. Multidimensional Skill Mismatch. *American Economic Journal: Macroeconomics*, 12(1):210–244, Jan. 2020. ISSN 1945-7707. doi: 10.1257/mac.20160241. URL https://www.aeaweb.org/articles?id=10.1257/mac.20160241.

G. Kambourov and I. Manovskii. Occupational Specificity of Human Capital*. *International Economic Review*, 50(1):63–115, 2009. ISSN 1468-2354. doi: 10.1111/j.1468-2354.2008.00524.x. URL https://onlinelibrary.wiley.com/doi/abs/10.1111/j.1468-2354.2008.00524.x. _eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1111/j.1468-2354.2008.00524.x.

J. Lise and F. Postel-Vinay. Data and Code for: Multidimensional Skills, Sorting, and Human Capital Accumulation, July 2020a. URL https://www.openicpsr.org/openicpsr/project/117662/version/V1/view.

J. Lise and F. Postel-Vinay. Multidimensional Skills, Sorting, and Human Capital Accumulation. *American Economic Review*, 110(8):2328–2376, Aug. 2020b. ISSN 0002-8282. doi: 10.1257/aer.20162002. URL https://pubs.aeaweb.org/doi/10.1257/aer.20162002.

A. Makles. Stata Tip 110: How to Get the Optimal K-Means Cluster Solution. *The Stata Journal*, 12(2):347–351, June 2012. ISSN 1536-867X. doi: 10.1177/1536867X1201200213. URL https://doi.org/10.1177/1536867X1201200213. Publisher: SAGE Publications.

R. Pavan. Career Choice and Wage Growth. *Journal of Labor Economics*, 29(3):549–587, 2011. ISSN 0734-306X. doi: 10.1086/659346. URL https://www.jstor.org/stable/10.1086/659346. Publisher: [The University of Chicago Press, Society of Labor Economists, NORC at the University of Chicago].

M. Poletaev and C. Robinson. Human Capital Specificity: Evidence from the Dictionary of Occupational Titles and Displaced Worker Surveys, 1984–2000. *Journal of Labor Economics*, 26(3):387–420, July 2008. ISSN 0734-306X. doi: 10.1086/588180. URL https://www.journals.uchicago.edu/doi/full/10.1086/588180. Publisher: The University of Chicago Press.

N. A. Roys and C. R. Taber. Skill Prices, Occupations, and Changes in the Wage Structure for Low Skilled Men, Nov. 2019. URL https://www.nber.org/papers/w26453.

C. Sanders. Skill Uncertainty, Skill Accumulation, and Occupational Choice. *2012 Meeting Papers*, 2012. URL https://ideas.repec.org//p/red/sed012/633.html. Number: 633 Publisher: Society for Economic Dynamics.

C. Sanders and C. Taber. Life-Cycle Wage Growth and Heterogeneous Human Capital. *Annual Review of Economics*, 4(1):399–425, Sept. 2012. ISSN 1941-1383, 1941-1391.

doi: 10.1146/annurev-economics-080511-111011. URL https://www.annualreviews.org/doi/10.1146/annurev-economics-080511-111011.

R. M. Sauer and C. Taber. Understanding women's wage growth using indirect inference with importance sampling. *Journal of Applied Econometrics*, 36(4):453–473, 2021. ISSN 1099-1255. doi: 10.1002/jae.2818. URL https://onlinelibrary.wiley.com/doi/abs/10.1002/jae.2818. _eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1002/jae.2818.

R. L. Thorndike. Who belongs in the family? *Psychometrika*, 18(4):267–276, Dec. 1953. ISSN 1860-0980. doi: 10.1007/BF02289263. URL https://doi.org/10.1007/BF02289263.

A Additional Tables and Figures

A.1 Number of Occupation Clusters

Following Makles (2012), I examine four different metrics (explained below) across various numbers of occupational groups. Figure A1 plots these metrics against the values of the number of clusters K from 1 to 10. At $K = 4$, the explained variance reaches about 80%. Although there is no significant "elbow", the within-cluster sum of squares drops at a decreasing rate at $K = 3$ or more. Overall, the results suggest that $K = 4$ is a reasonable choice. However, this clustering is based solely on occupational task requirements and does not account for factors such as the frequency of occupation selection within the sample or variations in average wage across occupations. With this in mind, I plan to extend the analysis by using a more granular set of occupational categories.

Within-Cluster Sum of Squares (WSS) The Within-Cluster Sum of Squares (WSS) measures the sum of squared distances between each data point and the centroid of the cluster it belongs to. Lower WSS values indicate tighter clusters. The formula for WSS is:

$$\text{WSS} = \sum_{k=1}^{K} \sum_{i \in C_k} ||x_i - \mu_k||^2$$

where K is the number of clusters, C_k is the set of data points in cluster k, x_i represents each data point in cluster k, and μ_k is the centroid of cluster k. Log of WSS is also computed to study the diminishing returns in WSS as K increases.

Explained Variance (η^2) η^2 measures the proportion of the total variance explained by the clustering structure. The formula for η^2 is:

$$\eta^2 = 1 - \frac{\text{WSS}}{\text{TSS}}$$

where TSS is the Total Sum of Squares, calculated as the sum of squared distances

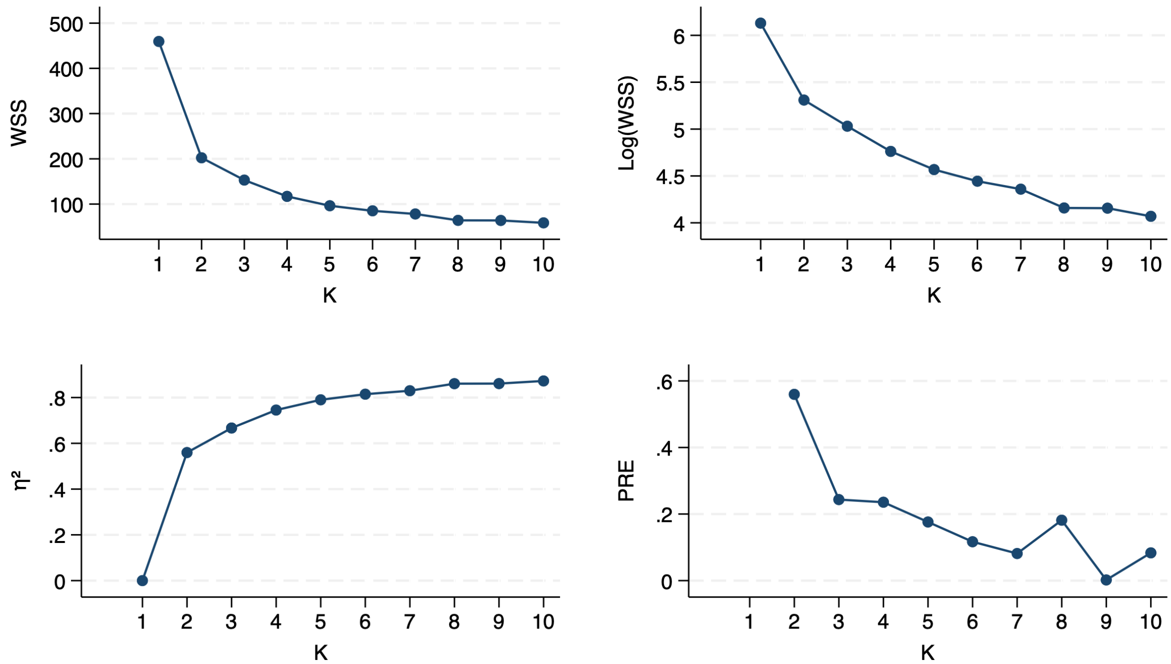


Figure A1: Comparison of different K cluster solutions. See text for statistics.

between each data point and the overall mean.

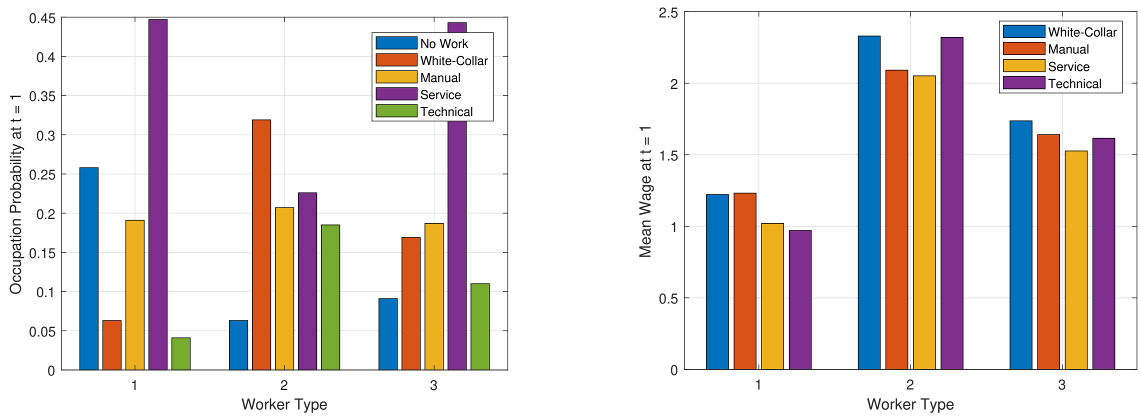
Proportion Reduction in Error (PRE) Proportion Reduction in Error (PRE) quantifies the relative reduction in error compared to a baseline model without clustering. It is defined as:

$$\text{PRE} = \frac{\text{TSS} - \text{WSS}}{\text{TSS}}$$

Table A1: Coefficients for the occupation choice multinomial logit model

	White-Collar	Manual	Service	Technical
Previous Occ = White-Collar	4.99	1.57	1.51	2.21
CI	[4.74, 5.20]	[1.28, 1.78]	[1.27, 1.68]	[1.95, 2.44]
Previous Occ = Manual	1.77	3.93	1.56	1.92
CI	[1.52, 1.97]	[3.70, 4.10]	[1.35, 1.72]	[1.66, 2.13]
Previous Occ = Service	2.05	1.54	3.20	2.05
CI	[1.85, 2.22]	[1.34, 1.70]	[3.03, 3.32]	[1.85, 2.26]
Previous Occ = Technical	2.61	1.73	1.69	5.60
CI	[2.34, 2.84]	[1.42, 1.94]	[1.43, 1.88]	[5.34, 5.84]
Cognitive Experience	0.25	0.21	-0.65	-0.01
CI	[0.10, 0.39]	[0.04, 0.35]	[-0.83, -0.55]	[-0.15, 0.13]
Manual Experience	-0.08	0.22	-0.04	0.05
CI	[-0.11, -0.05]	[0.19, 0.27]	[-0.07, -0.00]	[0.02, 0.08]
Interpersonal Experience	-0.30	-0.58	1.05	-0.00
CI	[-0.53, -0.03]	[-0.80, -0.29]	[0.89, 1.37]	[-0.21, 0.25]
Unobs type = 2	0.95	0.60	0.47	0.97
CI	[-0.54, 1.14]	[-1.00, 1.07]	[-0.22, 0.61]	[-0.96, 1.16]
Unobs type = 3	1.35	0.57	0.26	1.18
CI	[-1.18, 1.63]	[-1.34, 1.48]	[-0.43, 0.57]	[-1.86, 1.60]
Intercept	-2.91	-2.07	-1.34	-3.33
CI	[-3.26, -2.29]	[-2.88, -1.71]	[-1.51, -1.07]	[-3.79, -2.67]
(t=1) Unobs type = 2	1.70	0.96	0.85	1.86
CI	[-1.61, 2.26]	[-1.74, 1.88]	[-0.36, 1.04]	[-1.95, 2.55]
(t=1) Unobs type = 3	3.13	1.63	0.85	2.96
CI	[-17.56, 3.86]	[-2.79, 2.93]	[-0.76, 1.23]	[-13.37, 4.02]
(t=1) Intercept	-1.30	-0.35	0.60	-1.74
CI	[-2.24, 0.69]	[-1.66, 0.33]	[0.34, 1.23]	[-2.97, -0.35]

Notes: 95% Confidence intervals are calculated from bootstrapped sample.



(a) Occ P at t = 1 (b) Mean Wage at t = 1

Figure A2: Occupation Choice Probabilities and Mean Wages of each Worker Type at $t = 1$. Panel (a) represents the fraction of each type workers choosing each occupation (including non-working). Panel (b) depicts the average logged wages at each occupation, conditional on choosing the occupations.

Table A2: Detailed decomposition of wage variation across scenarios

Components	**Baseline**	**Random**	**Optimal**
Total Var(w) - Var(ε)	0.2216	0.1811	0.1748
Accumulated Cog	0.0408 (18.42%)	0.0249 (13.75%)	0.0243 (13.91%)
Accumulated Man	-0.0005 (-0.23%)	0.0008 (0.44%)	0.0024 (1.37%)
Accumulated Int	0.0046 (2.08%)	0.0029 (1.60%)	0.0029 (1.66%)
Occ Intercept	0.0174 (7.85%)	0.0093 (5.14%)	-0.0190 (-10.87%)
Initial Cog	-0.3593 (-162.19%)	-0.2899 (-160.05%)	-0.3129 (-178.98%)
Initial Man	0.0201 (9.08%)	0.0222 (12.26%)	0.0242 (13.85%)
Initial Int	0.4984 (224.89%)	0.4109 (226.85%)	0.4529 (259.12%)
Mean(w)	1.9914	1.9313	2.1197

Notes: Cross-sectional wage variation at year 10 is decomposed into components. Var(w) - Var(ε) represents the explained variance in wages after accounting for residual noise (ε). Accumulated components reflect the task-specific human capital accumulated over the career, while initial components capture the impact of initial skills at labor market entry. Occ Intercept represents fixed effects associated with occupations. The scenarios compare baseline observed occupation assignments with random assignment, and wage-optimal assignment at period $t = 10$.

B Identification: Hypothetical Scenario

Consider a worker with initial abilities ψ^C, ψ^M, and ψ^I.

Table B3: Hypothetical Scenario for Identifying Parameters

Year	Occupation	γ^C	γ^M	γ^I	E^C	E^M	E^I	Mean Wage
1	1	1	0	0	0	0	0	ψ^C
2	1	1	0	0	1	0	0	$\psi^C + \beta^C$
3	2	0	1	0	2	0	0	ψ^M
4	2	0	1	0	2	1	0	$\psi^M + \beta^M$

In this scenario:

- Wage growth from Year 1 to Year 2 identifies β^C, as it reflects changes in wages within a cognitive occupation.
- Wage growth from Year 3 to Year 4 identifies β^M, as it reflects changes within a manual occupation.
- The difference in the base wage level between Occupation 1 and Occupation 2 identifies $\psi^C - \psi^M$, capturing how initial ability is rewarded differently based on the task content of each occupation.

B.1 Identification: Less Extreme Scenario

Consider a less extreme scenario with a worker starting with initial abilities ψ^C, ψ^M, and ψ^I.

Table B4: Less Extreme Scenario for Identifying Parameters

Year	Occupation	γ^C	γ^M	γ^I	E^C	E^M	E^I	Mean Wage
1	1	1	0	0	0	0	0	ψ^C
2	1	1	0	0	1	0	0	$\psi^C + \beta^C$
3	2	0.3	0.7	0	2	0	0	$0.3\psi^C + 0.7\psi^M + 0.6\beta^C$
4	2	0.3	0.7	0	2.3	0.7	0	$0.3\psi^C + 0.7\psi^M + 0.69\beta^C + 0.49\beta^M$

In this scenario:

- Wage growth between Years 1 and 2, along with wage growth between Years 3 and 4, jointly identify β^C and β^M by leveraging variation in task intensity.
- The base wage level difference between Occupation 1 and Occupation 2 identifies ψ^C and ψ^M, as it reflects varying wage levels across occupations with a mix of cognitive and manual tasks.