PERUVIAN ECONOMIC ASSOCIATION

Lifetime Hours Inequality and Occupational Choice

César Urquizo Ubillús

Working Paper No. 208, April 2025

Lifetime Hours Inequality and Occupational Choice

César Urquizo Ubillús *

University of Pennsylvania

curquizo@sas.upenn.edu

This version: January 19, 2025

Please click here for the most recent version

Abstract

This paper explores the effect of hours worked on lifetime earnings inequality, a factor often overshadowed by the focus of the literature on wages. I argue that hours dispersion arises from individuals with heterogenous learning ability and leisure preferences selecting into occupations that reward hours worked with future wage growth at different rates. Using empirical evidence, I demonstrate strong correlations between occupational wage growth, cognitive test scores, and hours worked. Informed by this evidence, I develop and calibrate a life-cycle model of endogenous labor supply and occupational choice to disentangle the role of leisure preferences and learning ability in explaining hours worked and earnings dispersion. I find that cognitive ability is responsible for about one fourth of the variance in log hours at age 23, and leisure preferences are responsible for the remaining three fourths. Despite its seemingly small contribution to hours dispersion, cognitive ability accounts for three times as much of the variance of earnings at age 55 (31%) compared to leisure preferences (10%). Finally, I look into the normative implications of these findings, showing that when incorporating learning ability as a driver of hours dispersion, increases in tax progressivity are more effective at reducing inequality and less costly in terms of lifetime welfare.

JEL Classification : D15, J22, J24, J31
Keywords : Labor Supply, Occupational Choice, Life Cycle, Inequality

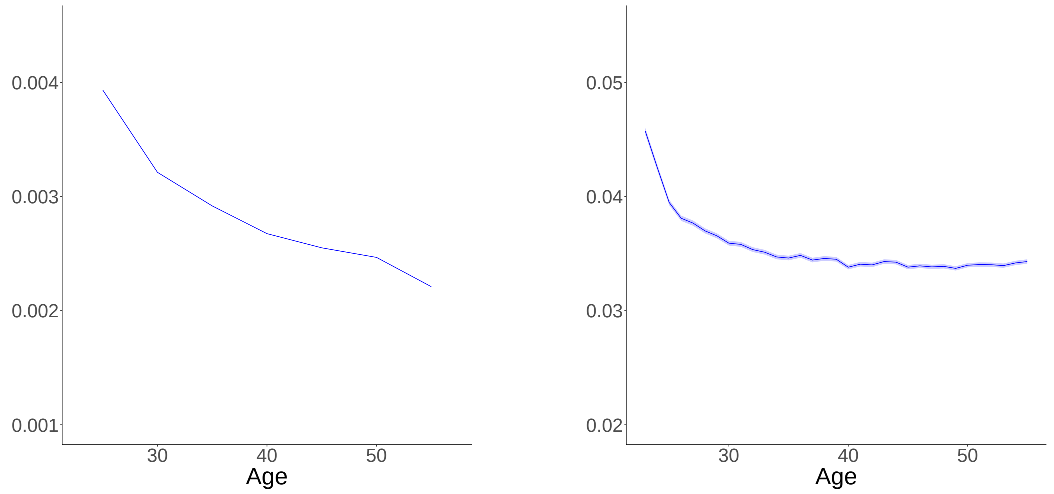
*I am grateful to my committe: Víctor Rios-Rull, Dirk Krueger, Iourii Manovskii, and Joachim Hubmer, for their feedback and support on this project. I would also like to thank Antonella Rivera, Daniel Jaar, Alberto Ramírez de Aguilar, Jordan Peeples, and participants of the Penn Student Macro Seminar, Penn Money Macro Seminar, Macroeconomics Across Time and Space Meeting, and the Research Seminar at Universidad del Pacífico for their insightful comments. All errors are my own.

1 Introduction

Earnings inequality has been widely documented to increase with age, but much of the existing literature has focused on how wages contribute to this growing disparity. The role of hours worked, however, has often been overlooked due to its modest direct contribution to earnings inequality. As shown in Figure 1, the between-occupation and cross-sectional variance of hours worked for males is higher at younger ages and declines over the life cycle.[1] This observation is important for two reasons. First because it suggests a potential role of human capital accumulation, since such a theory naturally predicts this pattern given that learning incentives disappear as people age. And second, because existing theories on the determinants of hours worked–both over the life cycle (Bick et al., 2024) and between occupations (Erosa et al., 2022)–fail to explain this pattern.

Figure 1. *Males, aged 23-55, working 20+ hours per week*

(a) Between-Occupation Variance of Log Hours **(b)** Cross-Sectional Variance of Log Hours



Source: American Community Survey 2000-2019

This raises an important question: What explains differences in hours worked between occupations and in the cross-section, and how do these differences shape earnings inequality? The literature has mainly relied on leisure preference heterogeneity to explain differences in hours worked over the life cycle and between occupations (Bick et al., 2024; Erosa et al., 2022). In this paper, I propose that learning ability in combination with occupations rewarding hours worked at different rates in terms of future wages also play an important role. By allowing for both drivers, and I can let the data speak

[1] This pattern also holds for females. However, given that a large proportion of women in their 20s temporarily exit the labor market to have children, some important assumptions required in the empirical analysis are violated if women are included in the sample. For this reason, this document will only focus on males.

about the importance of each. This theory can naturally explain the downward slopes presented in Figure 1 as well as the positive correlation between average wage growth, average hours worked and cognitive test scores by occupation reported later in this paper. Using a life-cycle model of labor supply and occupational choice, I disentangle the role of leisure preferences and learning ability in determining the dispersion in working hours. With this model, I conduct counterfactual experiments to quantify the role of each in determining life-cycle hours and earnings inequality. The normative implications of this analysis are significant: if hours dispersion is driven entirely by leisure preferences, I show that progressive taxation is less effective at reducing inequality while imposing greater welfare costs when compared to a model in which learning ability also plays a role.

I begin this paper by presenting several empirical facts regarding earnings and hours inequality over the life cycle, obtained from the American Community Survey (ACS) 2000-2019 and the National Longitudinal Survey of Youth 1997 (NLSY97). In this Section, I compute average life-cycle wage growth for the 25 major Census Occupations using ACS 2000-2019 and relying on the framework by Lagakos et al. (2018) to control for time and cohort effects when using repeated cross-sectional data. This analysis reveals two novel empirical facts. First, I provide evidence of a positive correlation between average hours worked within an occupation and average wage growth. Moreover, I find that this correlation is still present when controlling for current wages. Second, I document a positive correlation between average cognitive test scores taken early in life and average wage growth within an occupation. Both of these correlations are robust to using alternative measures of wage growth computed using panel data, allowing for controls on selection based on fixed observable and unobservable characteristics.

Informed by the empirical evidence, I develop a structural life-cyle model with endogenous labor supply and occupational choice. In the model, occupations differ in the extent to which they reward current hours worked with future wage growth, and individuals enter the labor market with heterogeneous leisure preferences, cognitive ability, and occupation-specific human capital. The central mechanism in the model is that there is complementarity between learning ability and occupation-specific human capital accumulation technology. In other words, individuals with higher cognitive ability are better positioned to benefit from the learning-by-doing mechanism in occupations that feature more significant learning opportunities.

I calibrate the model to match key empirical facts, using Generalized Method of Moments to estimate the parameters driving the key mechanisms: occupation-specific human capital accumulation technology, variance of learning ability and variance in leisure tastes. Drawing on a framework similar to that of Heckman et al. (1998), I assume that the dispersion in hours at age 55 is attributable solely to differences in leisure preferences. I use this insight to disentangle and separately estimate the variance of learning ability and leisure preferences. In addition, I use data on cognitive test scores to infer differences in the mean and variance of learning ability by occupation. Finally, using data on wages, I discipline the mean and variance of initial human capital by occupation, as well as its covariance with learning ability.

Using the calibrated model I conduct counterfactual experiments aimed at decomposing the variance in hours worked and earnings into components attributable to differences in learning ability and preferences for leisure. My findings indicate that differences in learning ability account for 24% of the variance in log hours at age 23, and 31% of the variance in log earnings at age 55. In contrast, differences in leisure preferences explain 73% of the variance in log hours at age 23, but only 10% of the variance in earnings at age 55. Notably, despite the relatively small contribution of learning ability to hours dispersion at age 23 when compared to leisure preferences, the interaction of hours, learning ability and wage growth implies that learning ability is much more important when determining earnings inequality at older ages.

Finally, I compare this model with a recalibrated version in which there is no learning ability heterogeneity, and all the variance in hours is driven by leisure tastes, to assess the effects of changes in tax progressivity on earnings and consumption inequality, output, welfare and wage growth for a cohort of individuals. The key finding of this analysis is that the source of hours dispersion significantly affects the welfare cost of reducing earnings inequality. When hours dispersion is driven entirely by differences in leisure preferences, increasing tax progressivity to reduce earnings inequality by one percentage point leads to a 2.6% reduction in lifetime welfare, measured in terms of equivalent consumption. However, when learning ability also plays a role in hours dispersion, the same reduction in earnings inequality results in only a 0.4% decrease in lifetime welfare[2]. This stark difference highlights that policies aimed at reducing inequality can have much lower welfare

[2]The intuition behind this result is that in the model with learning ability heterogeneity, top earners are mostly high ability, they are more sensitive to chances in progressivity, making the policy more effective at reducing inequality. Regarding the welfare changes, in the model with only leisure preference heterogeneity, inequality is a reflection of people's preferences, and thus, bringing the low earners closer to the top earners in terms of earnings will not have as big of an impact on the welfare of the low earners. This is discussed in more detail in Section 6.

costs when hours are influenced by factors like learning ability, which affect future wage growth, rather than solely by individual preferences for leisure.

I contribute to two main strands of literature: the study of earnings inequality and that of occupational choice. In particular, this paper provides new insights into how learning, occupational choice, and human capital accumulation influence the dynamics of hours dispersion and earnings inequality over the life cycle.

Research has long documented the rising dispersion of earnings both over the life cycle (Huggett et al., 2006) and over time (Heckman et al., 1998; Guvenen et al., 2022). Huggett et al. (2011) provided new insights using a life cycle model to decompose lifetime earnings inequality into lifetime shocks and initial conditions, showing how early-life differences shape long-term outcomes, while Heathcote et al. (2010) discussed the welfare implications of the rise in wage inequality. However, these studies focus exclusively on the contribution of wages to earnings inequality, neglecting the role of hours worked. Recent research by Bick et al. (2022) documents key features of the cross-sectional distribution of hours, while Bick et al. (2024) extends this work, emphasizing how differences in leisure preferences –both permanent and transitory[3]– can drive hours and earnings dispersion over the life cycle. [4]

I build on the emerging literature about hours inequality (Bick et al., 2024) by proposing a new theory of hours dispersion, which incorporates learning, occupational choice, and human capital accumulation as critical drivers of the cross-sectional dispersion of hours over the life cycle. Moreover, I use information on cognitive test scores to measure differences in mean and dispersion of learning ability by occupation and attribute the residual dispersion as preference heterogeneity, whereas the earlier literature loads all the variance in hours on preference heterogeneity by assumption. This theory matches not only the decline in hours dispersion reported in Figure 1, but also incorporates previously undocumented relationships between hours worked and wage growth across occupations.

In addition, I contribute to the literature on labor taxation and earnings inequality. While Badel et al.

[3]Bick et al. (2024) rely on the transitory component of leisure preference heterogeneity to match the magnitude of the autocorrelation of hours between times t and $t + s$, and its negative slope with respect to s. Although not a main focus of the theory I propose in this paper, my calibrated model is also able to generate an autocorrelation of hours that decreases with s thanks to occupation switching.

[4]While Bick et al. (2024) also allow for learning ability heterogeneity, which plays an important role in explaining lifetime earnings inequality, in their setup virtually all the variance in hours worked comes from leisure preference heterogeneity.

(2020) examines the impact of taxing top earners in a model with human capital accumulation, and Guvenen et al. (2014) explores how progressive taxation affects wage inequality, I conduct a policy experiment to evaluate the effectiveness of tax progressivity in reducing inequality and its welfare cost. Specifically, I compare outcomes in a model where all hours dispersion stems from preference heterogeneity with a model where part of the dispersion is driven by differences in learning ability. I find that increasing progressivity in the presence of learning heterogeneity is both more effective at reducing inequality and less costly in terms of welfare than when all hours dispersion is attributed to leisure preferences.

The study of occupational choice and mobility has a rich history, with its relationship to earnings inequality being a central focus since the introduction of the canonical Roy (1951) model. Early works by Jovanovic (1979) and Miller (1984) examine how individuals make occupational choices over the life cycle. Neal (1999) and Carrillo-Tudela and Visschers (2023) explore the factors that drive workers to switch occupations and the impact of these transitions on earnings.

My paper extends this research by being the first to report positive correlations between average wage growth, cognitive ability, and workweek length across occupations. I develop a model that incorporates these correlations, providing new insights into how individuals with different abilities and preferences select occupations that reward longer workweeks with higher future wages. This model also emphasizes the importance of occupation-specific human capital in shaping lifetime earnings, building on the findings of Kambourov and Manovskii (2009a,b), Kambourov et al. (2020), and Lise and Postel-Vinay (2020).

Furthermore, my work closely relates to Erosa et al. (2022), who use an extended Roy (1951) model with heterogeneity in leisure preferences and non-linear compensation varying by occupation to explain differences in workweeks and earnings across occupations. I build upon their static model by introducing a dynamic component, proposing that the non-linear compensation structure manifests as future wage growth. This addition makes learning ability a crucial factor in understanding why different occupations have different workweeks. Moreover, while Erosa et al. (2022) load all the differences in hours worked on leisure preferences–which are inferred from a model–I use data on cognitive test scores to measure differences in the mean and dispersion of learning ability between occupations. The novel dynamic dimension can explain the decreasing between-occupation dispersion of hours worked over the life cycle, and is also consistent with the empirical evidence, which suggests that after controlling for current wages, occupation average wage growth is positively correlated with the hours worked by an individual.

By integrating the dynamics of occupational choice, learning ability, and human capital accumulation into the analysis of hours dispersion and earnings inequality, my paper offers a novel perspective on the mechanisms driving these phenomena. The theory I propose can produce the two patterns observed in Figure 1, which the literature on life-cycle/occupational labor supply fails to explain. Furthermore, it highlights important policy implications, particularly regarding how changes in tax progressivity can affect output, welfare, wage growth, and earnings inequality by distorting labor supply.

The rest of this paper is structured as follows. Section 2 presents facts about between-occupation and lifecycle hours inequality. Section 3 presents the structural life-cycle model. Section 4 describes the calibration strategy and presents the estimation results. Section 5 presents the counterfactual experiments to decompose the variance of hours and earnings. Section 6 reports the results of running policy experiments with tax progressivity. Section 7 concludes.

2 Empirical results

In this section I describe the data used for documenting empirical facts about lifetime hours, wages, earnings, and occupational choice. I also document two novel empirical facts that motivate the main mechanism in my model: a positive, statistically significant, correlation between occupation-specific average wage growth and workweek, and between occupation-specific average wage growth and cognitive test scores.

2.1 Data

The empirical facts presented in this section come mainly from two data sources: American Community Survey (ACS) 2000-2019, and National Longitudinal Survey of Youth 1997 (NLSY97). Throughout this section I focus mainly on highly attached males aged 23-55, that means people that responded that their usual week throughout the past year was 20 hours or more, to avoid dealing with part-time workers that might have multiple occupations. Women are excluded from the sample because the labor supply determinants for women in their 20s are likely to also be influenced by fertility decisions, which I will not attempt to model in this paper.

I start the sample at age 23, because the average age of college graduation in the US is between 23 and 24, and I argue that the decreasing pattern in the cross-sectional variance of hours is not primarily driven by people who are working and studying simultaneouosly. Starting the sample at age 25 reduces the slope of the decline in the variance of hours over the life cycle, since the variance at 23 is about 30% higher than the variance at 55, whereas the variance at 25 is only about 20% higher. Although this decrease is not negligible, the fact that the pattern remains suggests that part-time studying cannot be entirely responsible for the reduction in the variance of hours. Moreover, it is not clear that part-time studying would be responsible for the decrease that we observe in the between-occupation variance of hours worked. The reason for cutting the sample at age 55 is because after that age, the fraction of people that retire, either totally or partially starts growing every year, which introduces a new mechanism that affects the variance of hours and is not the focus of this paper.

A final note on the sample selection is that I also excluded 3 out of the 25 Census occupations[5], due to small sample sizes or unavailability of data in one or more of the datasets that I used during the model calibration.

From ACS 2000-2019 I obtained cross-sectional data of hours worked, wages and earnings for individuals of different ages. From NLSY97, I obtain panel data of hours and wages, as well as learning ability measured by Armed Services Vocational Aptitude Battery (ASBAV). Although NLSY79 provides a longer sample than NLSY97, I picked the latter because it covers mostly the same years as my ACS sample. Therefore, I can use it for validation of some of the results I obtain from the cross-sectional data that I use to infer life-cycle behavior without having to worry about time effects making the results non-comparable. In Table 1 I provide a summary of the ACS and NLSY97 samples used in this section. For some additional descriptive statistics by occupation, see Table 12 in Appendix A.

[5]The excluded occupations are Financial Specialists, Extraction Workers, and Military Specific Occupations

Table 1. *ACS and NLSY97 sample descriptions*

Statistic	**ACS**	**NLSY97**
Observations	14,172,908	34,605
Years	2000-2019	1998-2019
Frequency	Annual	Annual
Workweek Range	21-80	21-80
Age Range	20-55	20-39
# of Occupations	22	22
Sex	Males	Males
Mean Age	41.54	26.65
Mean hours worked	43.73	42.20
Standard deviation hours worked	8.90	8.53
Mean hourly wage (2019 US$)	31.30	20.12

Additionally, and although it will not be discussed in this section, I also used the ONET data on skill requirements by occupation to compute a measure of how easily human capital can be transferred between occupation pairs. I then use this measure to construct a matrix of occupation specific human capital transferability, which I used for the structural model and will be defined later in Section 3. For the details of this process, see Appendix D.

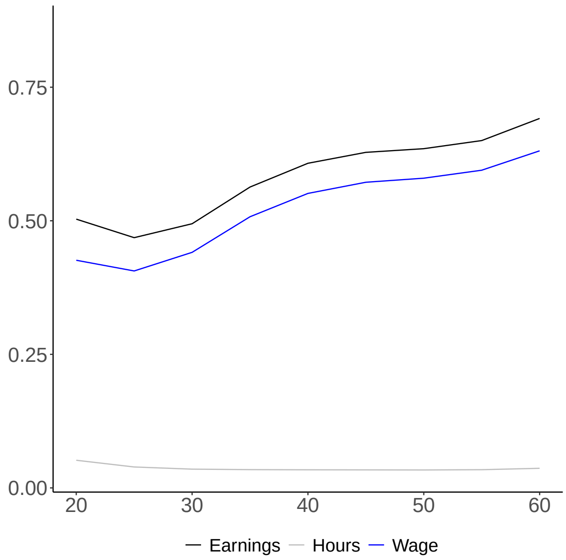
As mentioned earlier, a common concern when working with cross-sectional data to infer about life cycle behavior is that we might be confusing time, cohort and age effects. I deal with this potential problem in multiple ways. Firstly, I try computing the same statistics with ACS by restricting the sample to 2000-2010 and compare it to the 2010-2019 sample as well the full sample, finding no significant changes in the observed patterns. Secondly, I look at hours and wages behavior in the ACS full sample to NLSY97 to make sure the life-cycle behavior inferred from the richer ACS dataset is consistent with what is observed in panel data. Thirdly, to compute life-cycle wage growth from the repeated cross-sectional data I follow the methodology proposed by Lagakos et al. (2018) to control for time and cohort effects. I validate the results of this methodology by comparing the occupation-specific wage growth computed with ACS to the returns to occupational experience using NLSY97, finding a very high correlation between the two.

2.2 Decomposing the variance of earnings, wages and hours

Figure 2 shows the cross-sectional variance of the logarithm hours, wages and earnings computed using ACS 2000-2019. It has been widely documented that the dispersion of wages and earnings increases with age, but there's been considerably less attention put into the behavior of the dispersion of hours. From looking at this Figure, it might seem evident why most of the attention has been put on wages contribution to earnings dispersion, since the mechanical contribution of hours into earnings dispersion is many orders of magnitude smaller than that of wages. However, I will argue in this paper that human capital accumulation is going to make hours dispersion a much more important determinant in the variance of earnings than suggested by a simple mechanical decomposition.

Moreover, as stated in the introduction, my goal is not just to argue that hours are a more relevant contributor to earnings dispersion than previously acknowledged, but also to understand what is behind the dispersion in hours. As shown in Figure 1**(a)** the between-occupation variance of log hours drops from close to 0.004 to close to 0.002, which means there is a nearly 50% drop in the between-occupation variance of hours over the life cycle. Figure 1**(b)** on the other hand, shows the cross-sectional variance of hours worked drops from 0.47 at 23 to 0.034 at 55, with most of the drop happening before age 35. These findings are insightful because they challenge the notion that occupations can be characterized as fixed-hour packages. Moreover, they suggest that the mechanism driving the variance of hours over the life cycle do not remain constant over time.

Figure 2. *Variance of Log Hours, Wages and Earnings*



Source: American Community Survey 2000-2019

To further motivate this project, and to shed light on the potential drivers of the dispersion of hours, wages and earnings, in Table 2 I present the standard deviation of hours, wages and earnings, conditioning on 3-digit occupations (occ), learning ability measured as cognitive test scores (LA) and both, using NLSY97. To obtain this estimates, I created bins for cognitive test scores and 3-digit occupations[6]. I then computed the variance of log hours for each of the bins and I present the mean of those variances. Additionally, I restricted the sample to individuals aged between 30 and 39, when the standard deviation of hours is relatively stable. When conditioning by only by either learning ability or occupations we can see that the variance of log hours drops by about 29% and 22% respectively. When looking at wages, and earnings, the drop is of about 40% and 30% respectively. However, conditioning on both learning ability and occupation the variance of log hours drops by about 80%, and that of wages and earnings by about 86%.

[6]The results in Table 2 presented use 93 bins for occupations, which correspond to the 3-digit occupations present in the data, and 90 bins for cognitive test scores. Changing the number of bins for cognitive test scores changes the numbers in the table, but the result always remains that cognitive scores and occupation explain significantly more of the variance of hours, wages and earnings than each of them separately.

Table 2. *Standard deviation of hours, wages, and earnings by occupations and cognitive test scores*

x	$V(x)$	$V(x\|LA)$	$V(x\|occ)$	$V(x\|LA, occ)$
$Log(hours)$	0.0259	0.0184	0.0201	0.0052
$Log(wage)$	0.3844	0.2376	0.2708	0.0475
$Log(Earnings)$	0.4420	0.2694	0.3065	0.0544

Source: National Longitudinal Survey of Youth 1997

These results suggest that it is not enough to look at individuals with similar cognitive ability or in the same occupation to find people with similar labor supply. Instead, it seems to be that the interaction between these two characteristics is what determines the dispersion of hours, wages and earnings. In addition, these results further support the claim that occupations should not be seen as fixed-hour packages, as the variance of hours within occupations is still quite high when not controlling by learning ability.

2.3 Life cycle wage growth by occupation

In this section I provide some details on the computation of average wage growth, or Experience Wage Profile (EWP), by occupation for 22 of the 25 Census occupations in the ACS dataset, using the methodology proposed by Lagakos et al. (2018) to control for cohort and time effects.

The approach to compute the occupation-specific EWPs consists of estimating the regression

$$\log w_{ict} = \alpha + \theta s_{ict} + \underbrace{f(x_{ict})}_{\text{EWP}} + \gamma_t + \chi_c + \varepsilon_{ict},$$

where w_{ict} is the wage of individual i, in cohort c, at time t, s_{ict} denotes years of schooling, x_{ict} is years of potential labor experience (age - 18), γ_t is a time effect, and χ_c is a cohort effect. $f(x_{ict})$ is the Experience Wage Profile, as it tells us the increase in log wage for an additional year of labor experience. For this exercise I chose the functional form for $f(x_{ict})$ to be

$$f(x_{ict}) = \sum_{g \in G} \phi_g \mathbb{1}_{x_{ict} \in g}, \tag{1}$$

where G is a set containing 5 year age groups from 23 to 55. To identify cohort, time and age effects, I make the assumption that there are no returns to experience in the final 5 years of working life, following Heckman et al. (1998). More specifically, I follow the iterative process described in Lagakos et al. (2018):

1. Given a number of years with no experience effects y, and a depreciation rate δ, guess a trend for the time effects,
2. de-trend log wages and estimate Equation 1 with experience and cohort effects,
3. check if estimated experience effects have declined, on average by δ percent in the last y years,
4. if previous statement is true, stop, otherwise update trend in the time effects,
5. repeat until convergence.

To obtain an EWP for each occupation, I estimate this regression restricting the sample to individuals in each occupation, but using the cohort and time effects estimated using the entire U.S. sample. In Table 3 I present a summary of the results of computing the Experience Wage Profile for 22 Census occupations. The column Wage Growth shows the relevant statistic for the slope of the EWP at age 55. For a table with the specific EWPs computed for each occupation see Table 12 in Appendix A.

Table 3. *Slope of life cycle wage profile by occupation*

Statistic	**Wage Growth**
Full sample average	1.88
Maximum	2.63
Minimum	1.43
Standard deviation	0.30
Range	1.20
Range / Sample mean	0.64

As shown in Table 3 there is significant dispersion in wage growth across occupations. In the full sample average, a person at age 50 earns about 1.88 times their salary at age 20. When we look at this figure by occupations, we can see that the occupation with the highest EWP has a slope of 2.63, whereas the one with the lowest has a slope of 1.43. This means that the difference in the wage growth of the highest and lowest wage growth occupations is 120 percentage points. When we look at the standard deviation, it is also quite high, at 0.2960. This means that the dispersion in wage growth across occupations is about 57% of the sample mean.

Since this computation of wage growth is based on cross-sectional data, it is possible that workers that are in an occupation at a young age will be very different from those who are in that occupation at a later age. To address this concern, I also computed a measure of return to occupation experience using NLSY97 data, to have a notion of wage growth when controlling for individual fixed effects. Unfortunately, NLSY97 only follows people until they are 39, and sample sizes by age-occupation are considerably smaller than in ACS. This means it is not possible to follow the same methodology as in the ACS data to compute EWPs. However, I was able to compute a measure of wage growth by estimating the regression

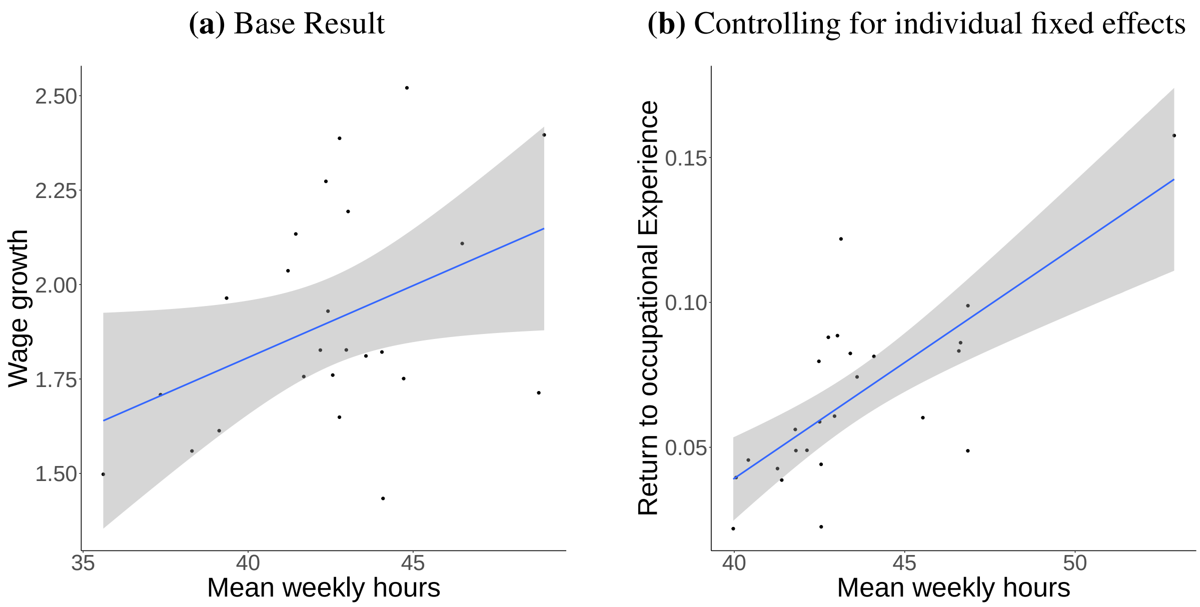
$$ln(w_{it}) = \alpha_i + \theta \; schooling_{it} + \sum_{j=1}^{J} \delta^j exp_{it}^j \times D_{it}^j + \gamma otherExp_{it} + \varepsilon_{it},$$

where i denotes an individual, t denotes time, α_i is an individual fixed effect, $schooling_{it}$ are the years of schooling, exp_{it}^j are the year of experience in occupation j, and D_{it} is an indicator variable that is 1 if an individual is working in occupation j. $otherExp_{it}$ is the experience in all other occupations. δ^j in this regression will give us information about the returns to occupation specific experience. The resulting returns to experience computed from this regression do not differ significantly from the EWPs computed using the ACS data. In fact, both approaches provide similar rankings of occupations by wage growth, and the correlation between these two measures is 0.7.

2.4 Correlation between wage growth and workweek

Using the computed slope for the occupation-specific EWPs I computed the correlation between this slope and the average workweek by occupation. The results are shown in Figure 5(**a**). As we can see, there is a positive, statistically significant correlation (0.47857) between the average wage growth of an occupation and its average workweek. This correlation is robust to controlling for fixed effects, as shown in Figure 5(**b**). This correlation is also present when I restrict the samples to the periods 2000-2010 and 2010-2019.

Figure 3. *Correlation between wage growth and workweek*



Source: American Community Survey 2000-2019, National Longitudinal Survey of Youth 1997

I also tested the robustness of this correlation to using both more coarse and finer definitions of occupations. When looking into occupations classified strictly as cognitive and not cognitive, which have the advantage of not having a lot of occupation switching between them, I still find that cognitive occupations, which have higher average wage growth also have higher average workweeks. When looking instead at the correlation of average hours worked and EWP in occupation-industry pairs, I found that the positive correlation still exists and is statistically significant. These results are provided in Appendix B.

A final approach I took to test the robustness of this finding was exploiting the individual variation in hours worked by estimating the regression

$$\log(\ell_{ij}) = \beta_0 + \beta_1 EWP_j + \beta_2 \log(w_{ij}) + \beta_3 age_{ij} + \beta_4 age_{ij} \times EWP_j + \varepsilon_{ij}, \qquad (2)$$

where EWP_j is the slope of the EWP for occupation j, w_{ij} is the wage of individual i in occupation j, and ℓ_{ij} are the weekly hours worked by individual i in occupation j. The results of this regression are shown in Table 4. As we can see, the coefficient of the EWP is positive and statistically significant, suggesting that individuals in occupations with higher average wage growth work more hours. Moreover, the coefficient of EWP, and the R^2 of the regression are larger when I restrict the sample to young workers (ages 23-30) than when I restrict it to older workers (ages 50-60), and the coefficient of the interaction between EWP and age is also negative and statistically significant.

Equation 2 also provides evidence of a positive impact of future wage growth even after controlling for current wage. This hints that relying only on different non-linear static compensation structures for different occupations, as suggested by Erosa et al. (2022), is not enough to explain the differences in hours worked.

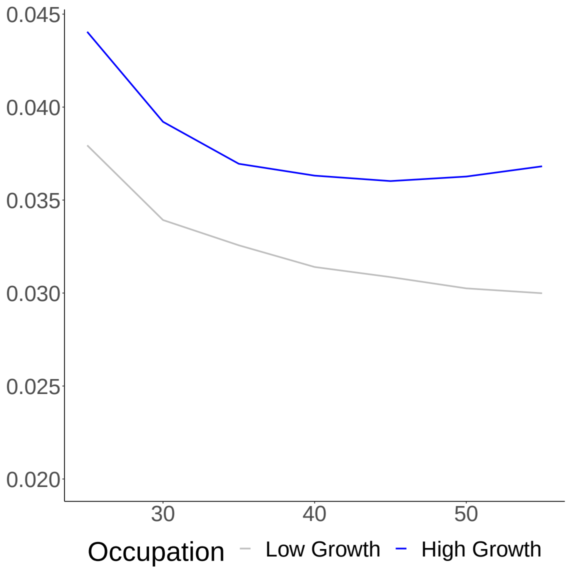
Table 4. *Regressions of log weekly hours worked on Wage Growth and log wage*

	Dependent variable:				
	$\log(hours)$				
	Full sample	Full Sample	Young Workers (23-30)	Old Workers (50-60)	Full Sample
	(1)	(2)	(3)	(4)	(5)
Wage Growth	0.171***	0.145***	0.173***	0.135***	0.265***
	(0.0004)	(0.0004)	(0.001)	(0.001)	(0.002)
$\log(wage)$		0.017***	0.023***	0.009***	0.017***
		(0.0001)	(0.0002)	(0.0002)	(0.0001)
Age					0.004***
					(0.00005)
Wage Growth × Age					−0.003***
					(0.00004)
Constant	3.518***	3.500***	3.420***	3.549***	3.324***
	(0.001)	(0.001)	(0.001)	(0.001)	(0.002)
Observations	9,867,523	9,867,523	1,751,347	2,594,844	9,867,523
Adjusted R^2	0.016	0.020	0.025	0.013	0.021
Note:					*p<0.1; **p<0.05; ***p<0.01

Source: American Community Survey 2000-2019

In addition to looking at differences in mean hours, I also tested for differences in the variance of hours by occupation type over the life cycle. When looking at the 22 census occupation it is hard to distinguish a pattern due the noise generated by having small sample sizes. However, when looking at the variance of hours over the life cycle for the top third of occupations by wage growth, and we compare it to the bottom third, there is very clearly a difference (see Figure 4). The variance of hours maintains its negative slope over the life cycle, but is consistently higher in the high growth occupations.

Figure 4. *Variance of Log Hours by occupation type*

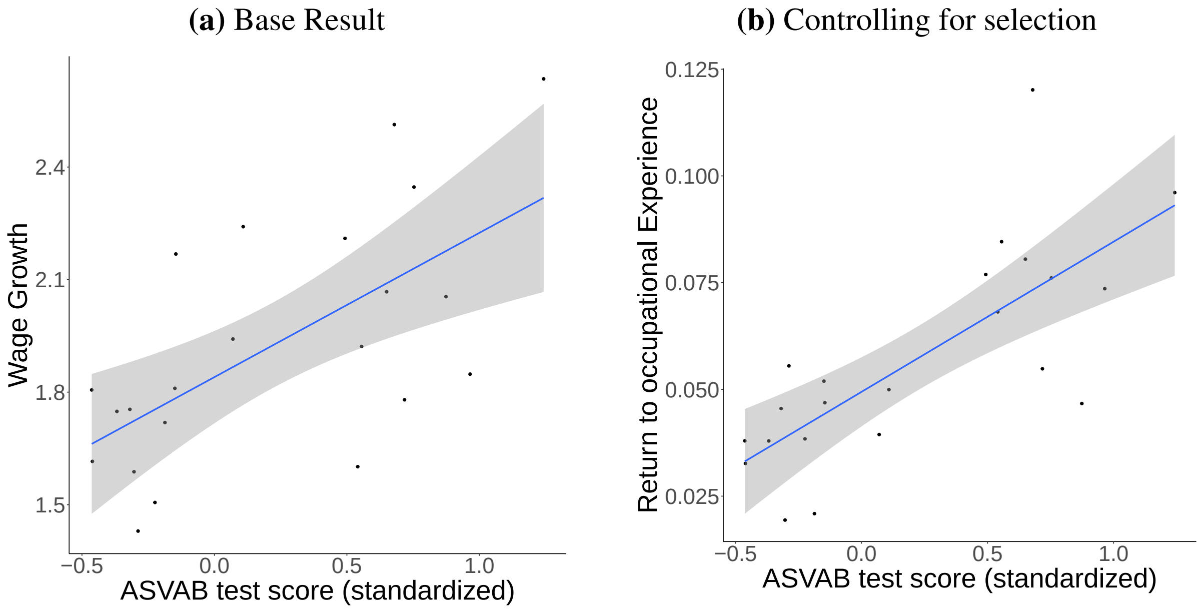


Source: American Community Survey 2000-2019

2.5 Correlation between wage growth and cognitive test scores

To end this section, I provide evidence of a positive correlation between occupation specific wage growth and average cognitive ability measured by the ASBAV test in the NLSY97 dataset. Figure 5**(a)** shows this correlation using the wage growth measure computed with ACS, whereas Figure 5**(b)** shows the correlation when using the wage growth measure computed with NLSY97, controlling for fixed effects. The correlation value is 0.64 and 0.75, respectively, and both are statistically significant.

Figure 5. *Correlation between wage growth and cognitive test scores*



Source: American Community Survey 2000-2019, National Longitudinal Survey of Youth 1997

Additionally, I also found a positive and statistically significant correlation between cognitive ability measured by the ASBAV test and hours worked when looking at workers aged 23-30. In fact, scoring one standard deviation higher on the test is associated with working an additional 1.7 percentage points hours per week. Although this number might seem small, a 1.7 percentage point longer week could accumulate to reflect substantial differences in earnings over the life cycle. More interestingly, I find that this relation is heterogenous with respect to occupation-specific wage growth. If we take, for example, the effect of having a one standard deviation higher test score on the log of hours worked, I find that in the Medical Occupations it is 0.04, whereas in Farming, Fishing and Forestry Occupations it is 0.00. That is, the effect of cognitive test scores on hours worked is only important when individuals are in occupations that have high average wage growth.

The results in Table 5 suggest that cognitive ability is also an important factor in determining hours worked at the individual level at younger ages in occupations that feature high learning opportunities.

Table 5. *Regression of hours worked on wage level and cognitive test scores, workers aged 23-30*

	Dependent variable:			
	$\log(hours)$			
	(1)	(2)	(3)	(4)
Wage Growth	0.267***		0.232***	0.184***
	(0.014)		(0.015)	(0.015)
Test Score		0.017***	−0.069***	−0.078***
		(0.002)	(0.019)	(0.019)
$\log(wage)$				0.049***
				(0.003)
Wage Growth × Test Score			0.057***	0.060***
			(0.013)	(0.013)
Constant	3.325***	3.694***	3.371***	3.302***
	(0.019)	(0.002)	(0.020)	(0.021)
Observations	11,824	11,824	11,824	11,824
Adjusted R^2	0.030	0.008	0.034	0.052
Note:				*p<0.1; **p<0.05; ***p<0.01

Source: National Longitudinal Survey of Youth 1997

3 Model

In this Section, I develop a structural model of labor supply and occupational choice that incorporates the correlation between learning ability, hours worked and wage growth by occupation. The model is designed to capture the dynamics of human capital accumulation occupation that differ in learning-by-doing technology across the life cycle.

3.1 Model Setup

In this economy workers start working at age a_0 and live for a finite number of periods T. Worker i characterized by their vector of occupation-specific human capital $\boldsymbol{h}_i$, current occupation j, learning ability parameter ξ_i and a leisure preference parameter ψ_i, which differ across workers. They all are hand-to-mouth and share the same Frisch elasticity of labor supply $1/\eta$, and relative risk aversion coefficient σ. Their instant utility is given by:

$$u(c,l) = \frac{c^{1-\sigma}}{1-\sigma} - \psi_i \frac{l^{1+\eta}}{1+\eta}.$$

Workers can accumulate human capital through learning-by-doing, following Heckman et al. (2003) who suggest that evidence favors this type of model over on-the-job training. Human capital evolves according to the function $H(\cdot)$, which I will describe in detail in the next subsection. Individuals can be employed at one of J occupations, which are characterized by occupation specific human capital h_j, and the rate at which current labor supply (ℓ) affects future wages θ^j.

For ages $a_0 < a < T$, the problem of a worker with learning ability parameter ξ_i, leisure preference parameter ψ_i, employed in occupation j is given by

$$V(\boldsymbol{h}_i, j, a; \xi_i, \psi_i) = \max_{c_i, \ell_{i,j}} \left\{ u(c_i, \ell_{i,j}) + \beta \mathbb{E}\left[\max \left\{ V(\boldsymbol{h}_i', k, a+1; \xi_i, \psi_i) + \varepsilon^{j,k} \right\}_{k=1}^{J} \right] \right\}, \quad (3)$$

$$\text{s.t.} \quad c_i = (1-\tau_0)\left(A(h_{i,j}\ell_{i,j})^{\zeta}\right)^{1-\tau_1} + Tr,$$

$$\boldsymbol{h}_i' = H(\boldsymbol{h}_i, j, \xi_i, \ell_{i,j}, k) \quad \forall\, k \in \{1, 2, ..., J\},$$

where $\varepsilon^{j,k}$ is a taste shock that affects the utility of choosing occupation k while being employed at occupation j, and follows an extreme value distribution with scale parameter α and location parameter $\mu^{j,k}$. In this economy the labor market is frictionless and workers can switch occupations at no monetary cost. They get paid their marginal productivity, which is characterized by the parameters of the production function A and ζ, and only depends on occupation-specific human capital. τ_0 and τ_1 are the parameters of the tax system, where the former is the flat tax rate and the latter is the tax progressivity. This functional form for post-tax earnings follows Bénabou (2002) and Heathcote et al. (2010). Tr are the transfers that the government gives to workers.

Given the high dimensionality of the problem, I assume that agents cannot save. Although this is a strong assumption, I do not expect it to affect the main mechanisms of the model, as the effects of learning ability on labor supply and occupational choice are particularly relevant for younger individuals, that tend to have very low wealth. Differences in initial wealth across individuals and occupations will partially be accounted by differences in the initial distribution at age 23[7].

It is worth noting that in this model learning ability does not make a person immediately more productive. It only affects the rate at which they accumulate human capital by working in a particular occupation, and therefore only benefits them through future wages. However, workers will also be able to transfer their human capital to other occupations, as I will explain in the next subsection.

Given that the taste shocks $\varepsilon^{j,k}$ follow an extreme value distribution with scale parameter α and location parameter $\mu^{j,k}$, we have analytical expressions for the choice probabilities of each occupation. In particular, the probability of choosing occupation k when employed at occupation j is given by

$$q(\boldsymbol{h}_i, j, a, k; \xi_i, \psi_i) = \frac{exp\left(\alpha\left(V(\boldsymbol{h}_i, k, a; \xi_i, \psi_i) + \mu^{j,k}\right)\right)}{\sum_{\iota \neq j} exp\left(\alpha\left(V(\boldsymbol{h}_i, \iota, a; \xi_i, \psi_i) + \mu^{j,\iota}\right)\right) + \exp(\alpha(V(\boldsymbol{h}_i, j, a; \xi_i, \psi_i) + \mu^{j,j}))}. \quad (4)$$

[7] As I explain in detail in Section 4, to bring the model to the data, I assume the initial occupation choice to have already taken place, and then I estimate the mean and variance of the human capital in each occupation at age a_0. The differences in the mean and variance of initial human capital by occupation should partly capture differences in wealth or other factors occurring before age a_0 that might have impacted their initial choice and level of human capital.

3.2 Evolution of Human capital

As reported in Section 2 evidence suggests a positive correlation between occupational average wage growth, work week, and cognitive ability. This fact is built into the human capital accumulation function $H(\cdot)$, and will be the main mechanism driving the cross-sectional and between-occupation dispersion of hours worked in this model. More specifically, $H(\cdot)$ takes the following form

$$H(h_j, j, \xi_i, \ell_{i,j}, k) = \begin{cases} \left[\chi_{a,j}(h_{i,j})\left(\xi_i \ell_{i,j}\right)^{\theta_j} + (1-\delta_j)\right] h_{ij} & \text{if } j = k, \\ \Gamma_{j,k}\left[A_{j,k} h_{ij} + B_{j,k}\right] & \text{otherwise,} \end{cases} \tag{5}$$

where θ_j is the occupation-specific human capital accumulation technology, δ_j is the occupation specific depreciation rate of human capital. $\chi_{a,j}(\cdot)$ is a function that captures the fact people's cognitive ability declines with age, making their learning more productive when they are young, and also provides decreasing returns to learning as human capital increases, accounting for differences in average initial human capital by occupation. In that sense, $\chi_{t,j}(\cdot)$ is defined as

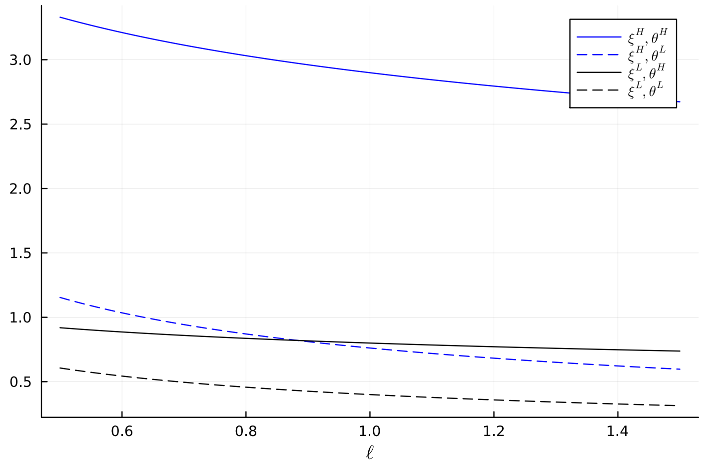
$$\chi_{a,j}(h_{i,j}) = \chi \left(\frac{a}{45 - a_0}\right)^{-\alpha_{age}} \left(\frac{h_{ij}}{C_j}\right)^{-\theta_j},$$

where C_j is a normalization constant to account for differences in initial wages (and human capital) across occupations. The term that depends on age is divided by $45 - a_0$ to reflect cognitive ability decline after age 45. The speed of the decline is governed by α_{age}.

Γ is a symmetric square matrix that denotes the level of transferability of human capital between occupations[8]. All the off-diagonal elements of Γ are smaller than 1, which means changing occupations has a cost in terms of human capital. $A_{j,k}$ and $B_{j,k}$, are two normalization coefficients that will be calibrated to preserve the relative standing of individuals in terms of human capital when switching occupations. That is, individuals that move from occupation j to occupation k will have their human capital transformed to be the same number of standard deviations above or below the mean of the human capital distribution at age 0 in the new occupation as it was in the old one. After this transformation, the matrix Γ will determine how much of the standardized human capital is lost

[8] Human capital transferability is computed by measuring the correlation of skill requirements at each occupation as defined in ONET. I provide more details about the computation of the elements of this matrix in Appendix D

Figure 6. *Derivative of $H(h_j, j, \xi, \ell, j)$ w.r.t. ℓ*



when transferring. The advantage of this functional form is that it greatly reduces the dimensionality of the problem, as it means that for each occupation I only need to keep track of one value for human capital, rather than the whole $J \times 1$ vector.

From the definition of $H(\cdot)$ we can see that the model features supermodularity between θ^j and ξ. That is

$$\frac{\partial H(h_j, j, \xi^H, \ell, j)}{\partial \theta^j} > \frac{\partial H(h_j, j, \xi^L, \ell, j)}{\partial \theta^j} \quad \forall\, \xi^H > \xi^L.$$

In other words $H(\cdot)$ is designed to capture the fact that individuals with higher learning ability will accumulate more human capital from the learning-by-doing mechanism in occupations that feature more significant learning opportunities (higher θ^j). This property can be seen more clearly in Figure 6, which shows the derivative of $H(\cdot)$ with respect to ℓ for two different values of $\xi = \{\xi^H, \xi^L\}$ and $\theta = \{\theta^H, \theta^L\}$. As we can see, for any level of labor, the gap between the derivative of $H(\cdot)$ with respect to ℓ for a high and low θ^j is larger for individuals with higher learning ability. That means that high cognitive ability workers lose more in terms of future wages by switching to an occupation with low θ than low cognitive ability workers.

From Equation 4, we can see that occupation choice probabilities will be driven by differences between the expected value of an individual for being employed at occupation j vs occupation k. The higher the difference, the higher the probability of picking occupation j over k. This, together with Equation 5 will guarantee that workers with higher learning ability, who learn more in high growth–high θ^j– occupations will be more likely to choose an occupation with high θ^j than workers with lower learning ability. This mechanism will be the main driver of the correlation between cognitive ability, hours worked, and wage growth by occupation in this model.

3.3 Initial and Terminal Conditions

At age $a = a_0$, workers choose their first occupation based on their preference and skill types and initial endowment of each occupation specific human capital. However, since I am not keeping track of the entire $J \times 1$ dimensional human capital vector, the only thing that matters for people's choices after age a_0 is their initial occupational choice. Therefore, I will assume an exogenous initial distribution for occupations at $a = a_0$, which I will estimate directly from the data to match the distribution of occupations for people aged 23.

Regarding the last period, at age $a = T$, dynamic incentives disappear, and therefore workers only maximize their instant utility, which simplifies the problem to:

$$V(\boldsymbol{h}_i, j, T; \xi_i, \psi_i) = \max_{c_i, \ell_{i,j}} \left\{ u(c_i, \ell_{i,j}) \right\},$$

$$\text{s.t.} \quad c_i = (1 - \tau_0) \left(A(h_{i,k} \ell_{i,k})^{\zeta} \right)^{1-\tau_1} + Tr.$$

3.4 Stationary Equilibrium and Model solution

The solution of the worker's problem is characterized by policy functions for labor $\ell(\boldsymbol{h}_i, j, a; \xi_i, \psi_i)$ and value functions $V(\boldsymbol{h}_i, j, a; \xi_i, \psi_i)$, satisfying the Bellman Equation defined in Equation 3. Value functions, in turn yield occupation choice probabilities $q(\boldsymbol{h}_i, j, a, k; \chi_i, \xi_i)$ defined in Equation 4.

I assume for simplicity that taxes collected by the government are used for unproductive expenditure from the point of view of workers, that is $Tr = 0$. Additionally, I assume that every period the mass of workers entering the labor market at age a_0 is equal to the mass of workers that die at age T. These assumptions, together with firms paying marginal productivity and making no profits, guarantee that solving the workers' problem will characterize the stationary equilibrium of the economy.

After solving for $\ell(\boldsymbol{h}_i, j, a, \xi_i, \psi_i)$ and $q(\boldsymbol{h}_i, j, a, k; \xi_i, \psi_i)$ I use Young (2010) non-stochastic simulation to build a sequence of transition matrices for each age $\{\Lambda_a\}_{a=a_0}^T$. With the sequence $\{\Lambda_a\}_{a=a_0}^T$ I can map any initial distribution of workers $x(\boldsymbol{h}_i, 0; \xi_i, \psi_i) = \{x(\boldsymbol{h}_i, j, 0; \xi_i, \psi_i)\}_{j=1}^J$ to a distribution at the end of their life cycle by iterating the following equation

$$x(\boldsymbol{h}_i, a; \chi_i, \xi_i) = \Lambda_a x(\boldsymbol{h}_i, a-1; \xi_i, \psi_i).$$

With the sequence of distributions $\left\{\{x(\boldsymbol{h}_i, j, a; \xi_i, \psi_i)\}_{j=1}^J\right\}_{a=a_0}^T$ I compute aggregate labor, wages, earnings, and welfare for each occupation for a cohort of individuals at every age in the stationary equilibrium.

4 Calibration and parameter estimation

In this section, I describe the process used to calibrate the model and present how well the model fits some targeted and untargeted aggregate moments. Given an exogenous initial distribution of people across occupations, I assume for every occupation j a jointly log-normal initial distribution for ξ, ψ, and h_j, and I estimate the parameters driving the main mechanisms of the model by using Generalized Method of Moments to match key empirical facts. To achieve identification of the role of preference and learning ability heterogeneity, I argue that the dispersion in hours at age 55 is attributable solely to differences in leisure preferences. I use this insight to disentangle and separately estimate the effects of learning ability and preference heterogeneity in driving hours dispersion among workers through the life cycle.

4.1 Defining the initial distribution

As stated in Section 3, I simplified the model by assuming the initial distribution $\{x(\boldsymbol{h}_i, j, 0, \chi_i, \xi_i)\}_{j=1}^J$ to be exogenous. In that sense, I estimate from the data a distribution over states (h_j, j, ξ, ψ) at age a_0. I start by defining

$$e^j = \int_{h_{ij}, \xi, \psi} x(\boldsymbol{h}_i, j, 0, \xi, \psi) dF(h_{ij}, \xi, \psi),$$

as the initial mass of workers in each occupation. This has a very clear empirical counterpart which is the share of young people in each occupation. Then, I assume that, conditional on being on occupation j, the distribution of human capital, learning ability, and disutility of labor is jointly log-normal, such that:

$$\begin{bmatrix} \log(h_{ij}) \\ \log(\xi_i) \\ \log(\psi_i) \end{bmatrix} \sim N \left(\begin{bmatrix} \tilde{h}_j \\ \tilde{\xi} + \tilde{\xi}_j \\ \tilde{\psi} \end{bmatrix}, \begin{bmatrix} \sigma^2_{\tilde{h}_j} & \sigma_\xi \sigma_{\tilde{h}_j \xi} & 0 \\ \sigma_\xi \sigma_{\tilde{h}_j \xi} & \sigma^2_\xi \sigma^2_{\xi,j} & 0 \\ 0 & 0 & \sigma^2_\psi \end{bmatrix} \right). \tag{6}$$

Individual disutility of labor (ψ) is an unobservable characteristic. However, the NLSY97 data set has information of cognitive test scores, which I argue will be informative of differences in the mean and dispersion of learning ability (ξ) by occupation. In that sense, although the cardinality of the test scores does not directly translate into the level of learning ability in my model, the mean and dispersion of log cognitive test scores by occupation should provide a good approximation of the differences in the mean and variance of log learning ability across them. As previously stated, since the cardinality of ξ is not directly observable, I will set the mean of log cognitive ability $\tilde{\xi}$ at an arbitrary number for normalization, and I will then estimate the parameter χ of the human capital accumulation function ($H()$) to fit the average wage growth observed in the data.

Equation 6 also contains a very strong assumption, which is that $\log(\psi)$ is orthogonal to both initial human capital and learning ability. In the base calibration of this model, I opt for this assumption given the difficulties associated with measuring labor disutility. However, in future versions of this document I will explore the possibility of relaxing this assumption, and use a relevant moment in the data to estimate the covariances between initial human capital, learning ability, and disutility of labor.

4.2 Parameter estimation

Some of the parameters in the initial distribution affect moments that do not depend on choices made after age a_0. Therefore, these parameters can be estimated without solving the full model. Let us define the set of such parameters as Θ_0:

$$\Theta_0 = \left\{ e^j, \tilde{h}_j, \tilde{\xi}_j, \sigma_{\tilde{h}_j}, \sigma_{\xi_j}, \sigma_{\tilde{h}_j,\xi} \right\}_{j=1}^{J},$$

where e^j is the initial mass of workers in each occupation, $\tilde{h}_j$ is the mean initial log human capital in each occupation, $\tilde{\xi}_j$ is the mean log learning ability by occupation, $\sigma_{\tilde{h}_j}$ is the standard deviation of initial log human capital, σ_{ξ_j} is the standard deviation of log learning ability, and $\sigma_{\tilde{h}_j,\xi}$ is the covariance between initial log human capital and log learning ability.

In fact, the result of minimizing the moment condition will in these cases yield analytical solutions for the elements of Θ_0. On the other hand, other parameters will require solving for the full model given their dependence on choices after age a_0. These parameters are defined as Θ_{full};

$$\Theta_{full} = \{\theta_1, \ldots, \theta_J, \sigma_\xi, \sigma_\psi, \chi\},$$

where θ^j is the human capital accumulation technology for each occupation, σ_ξ is the scale parameter for the standard deviation of learning ability, σ_ψ is the standard deviation of labor disutility, and χ is a parameter in the human capital accumulation function that denotes a general learning technology.

I will refer collectively to the elements of these two sets as $\Theta = \{\Theta_0, \Theta_{full}\}$. To estimate elements in Θ, I will use the Generalized Method of Moments (GMM). To do so, I define the set $\mathcal{G}(\Theta)$, which consists of the aggregate moments of the economy computed from solving the model given the parameter values Θ. As with the set Θ, I will separate the elements of $\mathcal{G}(\Theta)$ into two groups, corresponding to the parameters that only depend on the initial distribution and those that depend on the full model solution. I call these two sets $\mathcal{G}_0(\Theta_0)$ and $\mathcal{G}_{full}(\Theta_0, \Theta_{full})$, respectively. Let us also define $\bar{\mathcal{G}}_0$ and $\bar{\mathcal{G}}_{full}$ as the empirical counterparts of these moments in the data.

Then, I can set up a two-stage process to estimate each set of parameters. I start by finding the $\hat{\Theta}_0$ that solves the following problem

$$\hat{\Theta}_0 = \arg\min_{\Theta_0} \left\{ \left[\mathcal{G}_0(\Theta_0) - \bar{\mathcal{G}}_0\right]' W_0 \left[\mathcal{G}_0(\Theta_0) - \bar{\mathcal{G}}_0\right] \right\},$$

where W_0 is a weighting matrix. With the solution $\hat{\Theta}_0$ I can then obtain the estimate $\hat{\Theta}_{full}$ by solving

$$\hat{\Theta}_{full} = \arg\min_{\Theta_{full}} \left\{ \left[\mathcal{G}_{full}(\hat{\Theta}_0, \Theta_{full}) - \bar{\mathcal{G}}_{full} \right]' W_{full} \left[\mathcal{G}_{full}(\hat{\Theta}_0, \Theta_{full}) - \bar{\mathcal{G}}_{full} \right] \right\},$$

where W_{full} is a different weighting matrix. It is worth noting that under the assumption that $\zeta = 1$, which is true for the base calibration in the model, the expression $\left[\mathcal{G}_0(\Theta_0) - \bar{\mathcal{G}}_0\right]$ will yield exact analytical solutions for Θ_0 given that each moment in the set $\mathcal{G}_0(\Theta_0)$ depends on a single element of Θ_0. This means I just need to solve for each element in Θ_0 such that each of the relevant moment conditions are satisfied. On the other hand, the moment conditions for the parameters in Θ_{full} will not yield exact analytical solutions.

The choice for the moments $\bar{\mathcal{G}}_0$ is straightforward, since these moments do not depend on endogenous choices, and therefore I can pin down exactly which element in $\bar{\mathcal{G}}_0$ will be affected by which element in Θ_0. Moreover, these moments will have exact empirical counterparts, which are defined in table 6. However, the choice for the moments $\bar{\mathcal{G}}_{full}$ is not as straightforward, given that due to the endogenous choices made after age a_0, all parameters affect all moments. However, some empirical moments should be more informative about some parameters. In fact, I will argue that the estimated average wage growth by each occupation j should mostly be informative about the value of each θ^j, whereas the variance of hours worked by age should be informative about σ_ξ and σ_ψ.

As mentioned earlier, since the mapping between $\mathcal{G}_{full}(\hat{\Theta}_0, \Theta_{full})$ and $\bar{\mathcal{G}}_{full}$ is much more convoluted than the mapping between $\mathcal{G}_0(\Theta_0)$ and $\bar{\mathcal{G}}_0$, I rely on a global optimizer to find the values of Θ_{full} that minimize the distance between the model moments and the empirical moments. For a detailed account of which moments are matched to the set $\mathcal{G}(\Theta_0)$, and which moments are considered to be more informative about each parameter in Θ_{full}, see Table 6. Additional details of the estimation process are provided in Appendix E.

Table 6. *Estimated Parameters*

Estimated Parameters (Θ_0)			
Parameter	**Description**	**Moment**	**Source**
e^j	Initial mass at each occupation	Occupation shares at age 23	ACS
$\tilde{h}_j$	Occ. specific mean initial human capital	Mean wage across occupations, age 23	ACS
$\sigma_{\tilde{h},j}$	Std. dev. of initial human capital	Std. dev. of log(wage), age 23	ACS
$\tilde{\xi}_j$	Mean Learning ability by occ.	Mean log cog. test score by occupation	NLSY
$\sigma_{\xi,j}$	Variance of learning ability	V(log cog. test score) by occupation	NLSY
$\sigma_{h_j,\xi}$	Cov(human capital, learning ability)	Cov(wages, test scores), age 23	NLSY

Estimated Parameters (Θ_{full})			
Parameter	**Description**	**Moment**	**Source**
θ_j	Human capital accumulation technology	Occupation-specific avg. wage growth	ACS
σ_ξ	Scale for std. dev. of learning ability	Variance of hours, age 23	ACS
σ_ψ	Std. dev. of labor disutility	Variance of hours, age 55	ACS
χ	General learning technology	Full sample avg. wage growth	ACS

The remaining parameters in the model are calibrated externally, and their values and sources are provided in Table 7. I end this subsection by briefly discussing these choices. The parameters that are most important in determining the behavior labor supply over the life cycle, such as the discount factor, and the Frisch elasticity, and the tax rates are set to values that are standard in the literature (in particular, I follow Heathcote et al., 2017, 2020; Bick et al., 2024). To pick the depreciation rate of occupation-specific human capital I started by following Huggett et al. (2011), who use a value of 0.02. However, I believe that using the same depreciation rate for all occupations is not a realistic assumption given that the task requirements are very different. Having this in mind, I allow for the depreciation rate to be different in manual and cognitive occupations, and in particular I assume the former have a depreciation rate of 0.005, and the latter of 0.025. The relative risk aversion was set to 1 (log utility) for the model to be consistent with balanced growth path, as is standard in the literature. $\tilde{\psi}$ was picked to make the mean life-cycle labor to be equal to one, so that log labor can be interpreted as deviations from the mean. $\tilde{\xi}$ were set to an arbitrary number, given that the parameter χ in the human capital accumulation function, which I estimate to match general wage growth, has a similar effect.

Matrix Γ was computed by looking into the correlation between skill requirement scores by occupation in ONET. This dataset provides scores from 1 to 100 to 25 different skills that are required in each occupation. Using these scores, I computed a correlation matrix that I use as a proxy of how similar occupation-specific human capital should be for each pair of occupations. I provide more details about the computation of the elements of this matrix in Appendix D. The coefficients in the transformation of human capital when switching occupations $A_{j,k}$ and $B_{j,k}$ are set to transform human capital in an occupation into standard deviations from the mean. The normalization coefficient in the concavity of human capital accumulation function C_j is set to the mean initial human capital in each occupation.

Technical parameters associated to occupation switching such as $\mu^{j,j}$, $\mu^{j,k}$ and α are set to achieve high persistence in occupations, and preliminary sensitivity analysis indicate that small changes on them have no impact in the results of the model. In the base calibration, $\mu^{j,j}$, and $\mu^{j,k}$ are set to 0, and all the occupation persistence is achieved through α. Finally, when picking the parameters of the production function, ζ was set to be 1 so that production is linear and all the convexity in the compensation to labor comes from future wage growth, and TFP was arbitrarily set to 4, since it has no effect on wages given that I estimate the value of initial human capital to match initial wages. It is worth noting that making ζ greater than 1, and allowing it to vary by occupation would make this model closer to Erosa et al. (2022), and would introduce an additional mechanism of occupational choice to the model, based entirely on leisure preferences. However, in this document I want to focus on the role of learning ability in occupational choice, which is why I will not allow for this in the base calibration.

Finally, the parameter α_{age} was picked to match the slope of the variance of hours between ages 23 and 55. Although I am targeting the initial and endpoint of this variance, changing α_{age} allows me to better fit the entire life cycle profile. In future versions of this document, I will estimate this parameter to match the entire life cycle profile of the variance of hours.

Table 7. *Externally calibrated Parameters*

Parameter	Description	Moment / Value	Source
Γ	Occupation human capital transferability	Skill requirements by occupation	ONET
σ	Relative Risk aversion	1.0	BGP
$\tilde{\psi}$	Mean log disutility of labor	0.47	Normalization
$\tilde{\xi}$	Mean log learning ability	1.5	Normalization
$1/\eta$	Frisch Elasticity of labor supply	0.3	Bick et al. (2024)
β	Discount factor	0.98	Huggett et al. (2011)
ζ	Returns to scale in production	1.0	-
A	TFP	4.0	Normalization
δ_c	Depreciation rate of human capital cognitive occs.	0.025	-
δ_m	Depreciation rate of human capital manual occs.	0.005	-
τ_0	Flat tax rate	0.19	Heathcote et al. (2017)
τ_1	Tax progressivity	0.181	Heathcote et al. (2017)
α_{age}	Age decay of learning ability	0.2	-
$\mu^{j,j}$	Location par. for taste shocks, same occ.	0.0	-
$\mu^{j,k}$	Location par. for taste shocks, diff. occ.	0.0	-
α	Scale par. for taste shocks	1.3	-
$A_{j,k}$	Slope transformation of human capital	σ_k/σ_j	ACS
$B_{j,k}$	Intercept transformation of human capital	$(\sigma_j\bar{h}_k - \sigma_k\bar{h}_j)/\sigma_j$	ACS
C_j	Normalization, concavity of $H(\cdot)$	$\bar{h}_j$	ACS
a_0	Initial period	23	-
T	Terminal period	60	-

4.3 Identification strategy for variance of learning ability (σ_ξ) and leisure preferences (σ_ψ)

One of the main goals of this project is to quantify how much of the dispersion in hours worked over the life cycle is due to either dispersion in learning ability (σ_ξ) or in preferences for leisure (σ_ψ). Doing so requires disentangling in a convincing way the effect of each on the lifetime dispersion in hours. To do so, I will use the fact that in the model, by construction, σ_ξ does not affect the variance of hours worked at age T, which is caused by the absence of learning. The assumption that learning incentives disappear later in life is fairly standard in numerous theories of life-cycle wage growth, and there is a long history of using this assumption to identify structural parameters in models of human capital accumulation (see for example Heckman et al., 1998; Huggett et al., 2011; Bowlus and Liu, 2013; Schulhofer-Wohl, 2018; Lagakos et al., 2018).

Thanks to this assumption, I know that the variance in the learning ability parameter will have no effect on the variance of hours at age T, whereas the variance of hours at every age $a < T$ will be affected by both learning ability and leisure preferences. Therefore, I can separately identify σ_ξ and σ_ψ by matching the variance of hours in the data at age 23 and at age 55. Although both parameters are estimating jointly, the fact that σ_ξ will not affect the variance of hours worked in the last period means this moment will have to be matched entirely by σ_ψ. Then, σ_ξ will have to be such that the variance of hours at age $a = a_0$ will match the higher variance observed in the data for young workers. In other words, this identification strategy implies that if the hours profile was flat, I would load everything on leisure preferences, whereas if it was steep and going to zero, I would load everything on learning ability.

4.4 Evaluating Model Fit

In this subsection I report the fit of the model's aggregate moments to the data. In Figure 7 we can observe the model fit for the between-occupation and cross-sectional variance of hours that were reported in Figure 1. More specifically, Figure 7**(a)** shows the variance of log hours by age. Even though I am only targeting the variance of hours at ages 23 and 55, the model does a good job at matching the shape of the variance of hours profile for all ages.

In Figure 7**(b)** I show the between-occupation variance of hours worked, which is an untargeted moment. The model is able to match the decreasing profile of the between-occupation variance of hours over age, however it is not able to match the level. The fact that the between-occupation variance of hours is 0 at age 55 is due to a combination of the absence of learning in the last period and having log utility. These two factors make it so that the incentives to work in the last period are the same for all occupations. A possible way to make the between-occupation variance of hours higher at age 55 would be to introduce some covariance in the initial distribution between learning ability and leisure preferences, which I will explore in future versions of this document [9].

[9] Preliminary analysis shows that allowing for a positive covariance between ξ and ψ, keeping all else equal, increases the between occupation variance at the end of the life cycle, but reduces the cross-sectional variance of hours at every age, and flattens its life cycle profile. It also makes the between occupation variance of hours to be more L-shaped than in the model in which ψ is orthogonal. Introducing a negative covariance makes both the between occupation and cross-sectional variance of hours steeper, while also increasing the between occupation variance of hours at 55.

Figure 7. *Model Fit: Cross-sectional and between-occupation variance of hours*

(a) Variance of Log hours by age (Targeted) | **(b)** Between Occupation variance of log hours (Untargeted)

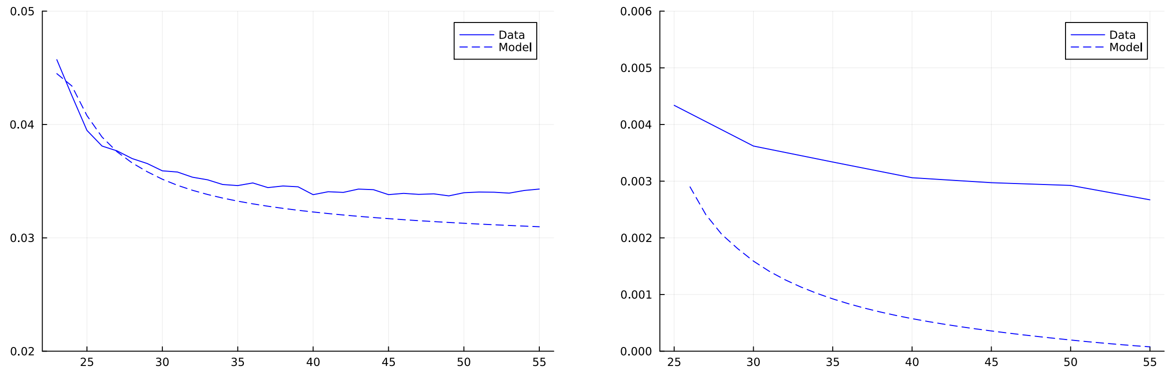
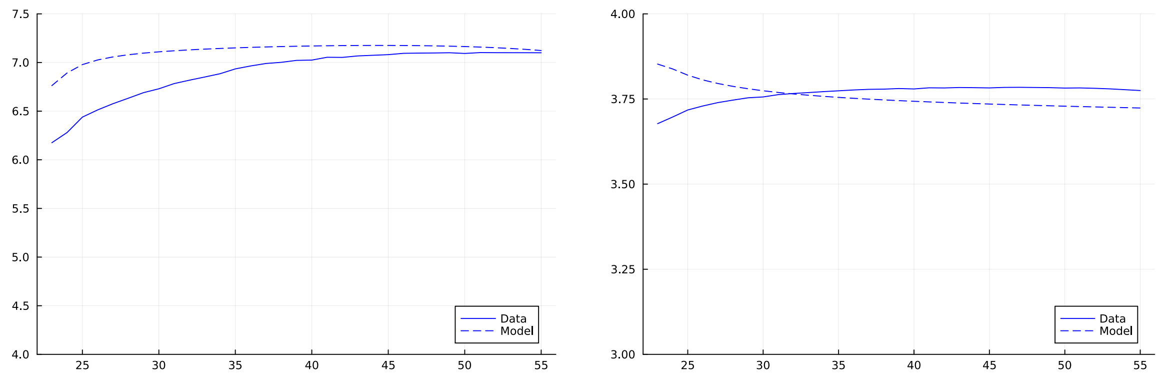


Figure 8(**a**) and 8(**b**) shows the fit of the model to the level of log earnings and log labor by age. In the current calibration the model can match the general shape of log earnings, but predicts higher earnings at young ages than the data. This is due to the fact that in the model the level of log labor at younger ages is higher than it is in the data. The reason for this is that the model underestimates the mass of people that stay in the low growth occupations when they are young with respect to the data. In a future version of this document, I will include an alternative version of the matrix Γ, that is computed from observing occupation switching patterns in the data between the 25 major census occupations, to achieve a more realistic composition of occupations as people age.

Figure 8. *Model Fit: Log Earnings and Log labor by age*

(a) Log Earnings by age (Untargeted) | **(b)** Log Labor by age (Untargeted)



In Table 8, I present the correlations between average workweek, wage growth and cognitive skill by occupation, all of which are untargeted. The model is able to match the sign of both correlations. However, although correlation between average cognitive ability and wage growth is relatively close to what is observed in the data, the model overstates the correlation between wage growth and workweek. A possible way to reduce this correlation is to introduce some covariance between learning ability and leisure preferences in the initial distribution, which I will explore in future versions of this document.

Table 8. *Correlation between average hours, wage growth and cognitive skill by occupation*

Statistic	**Model**	**Data**
Correlation Wage Growth and Workweek	0.7343	0.5510
Correlation Wage Growth and LA	0.4031	0.4786

5 Decomposing hours and Earnings Inequality

In this Section I use the calibrated model to conduct counterfactual experiments aimed at decomposing the variance in hours worked and earnings into components attributable to differences in learning ability, preferences for leisure, and initial human capital. My findings indicate that in this model with occupational choice, differences in learning ability account for 24% of the variance in log hours at age 23 and 30% of the variance in log earnings at age 55. In contrast, differences in leisure preferences explain 73% of the variance in log hours at age 23, 10% of the variance in earnings at age 55.

Table 9. *Decomposing the variance of hours worked and earnings*

	Statistic	Benchmark	$\sigma_\psi = 0$ No leisure het.	$\sigma_\xi = 0$ No Learning het.	$\sigma_\xi, \sigma_\psi = 0$ No leisure, No Learning het.
age = 23	V(Hours)	0.043	0.01	0.032	0.0
	% Benchmark	100	24	73	1
	V(Earnings)	0.551	0.461	0.457	0.351
	% Benchmark	100	84	83	64
age = 55	V(Hours)	0.031	0.0	0.031	0.0
	% Benchmark	100	0	99	0
	V(Earnings)	1.062	0.672	0.445	0.343
	% Benchmark	100	63	42	32

If we look at table 9, we can see that in the absence of both learning ability and leisure preference heterogeneity, the variance of hours worked and earnings is only 1% of the benchmark model at age 23 and 0% at age 55. This confirms that virtually all the variance of hours in the model is coming from differences in learning ability and leisure preferences, as expected given the main mechanism in the model[10]. The variance of earnings at both 23 and 55 when there is no learning or leisure heterogeneity is still significant, and this is because it is driven mechanically by the initial variance of wages estimated from the data. However, it is worth noting that in this counterfactual experiment, the variance of earnings stays relatively flat over the life cycle, and in fact it slightly decreases when comparing age 23 to age 55.

If we allow for learning ability dispersion, the variance of hours at 23 increases to about 24% of the benchmark. However, if we look at earnings at age 55, we can see it increases to 63% of the benchmark. This suggests that even if learning ability explains only 1/4 of the variance of hours at age 23, that additional dispersion combined with the learning ability heterogeneity explains an additional 31 percentage points of the variance of earnings at age 55 with respect to the benchmark

[10]The remaining 23 percentage points are explained by the extreme value shocks and people starting at different occupations.

model.

On the other hand, if we allow for leisure preference dispersion, the variance of hours at 23 increases to 73% of the benchmark. The variance of earnings at age 55, on the other hand, increases to 42% of the benchmark. That is, they only explain an additional 10 percentage points of the variance of earnings at age 55 with respect to the benchmark model. Interestingly, these results suggest that even if learning ability only accounts for a small fraction of the variance of hours when people are young, given how it interacts with the human capital accumulation technology, it has a much larger impact on earnings dispersion over the life cycle, which is 3 times larger than the role of leisure preference heterogeneity.

The remaning 27 percentage points of the variance of earnings at age 55 can be mostly explained by differences between occupations and at a smaller extent by the variance of the taste shocks.

6 Policy Applications

In this Section, I present the results of a policy experiment aimed at understanding the role of the progressivity of the tax system in shaping lifetime earnings and consumption inequality for one cohort of individuals. I use the calibrated model to assess the impact of increasing tax progressivity, while maintaining total government revenue constant, on lifetime earnings inequality, as well as lifetime output, welfare, and wage growth. This analysis offers novel insights into how these policies influence inequality by distorting human capital accumulation incentives.

To conduct this policy experiment, I first recalibrated the model shutting down differences in learning ability to match the lifetime variance in hours (as opposed to matching variance in hours at age 23 and 55). That is, I look at the world through two different lenses: all the variance in hours worked is caused leisure preferences, or part of it is caused by learning ability. In other words, we could think of the model with no learning ability heterogeneity as a restricted version of the base calibration. I then compare what would happen to a cohort's lifetime earnings inequality, output, welfare and wage growth if they had been born to a more progressive tax system in both the base calibration and the new calibration with no skill heterogeneity. To achieve revenue equivalence of the tax systems I increase progressivity parameter τ_1 and then adjust τ_0 such that the tax revenue is the same as in the baseline. It is worth noting that since in the model agents have log utility and cannot save, τ_1 is the only parameter of the tax system that affects labor supply. For the results for the base calibration, see Table 10. For the results in the calibration with no skill heterogeneity, see Table 11.

In the model with skill heterogeneity, the revenue-maximizing level of τ_1 is higher. This is evident from the fact that, as τ_1 increases, τ_0 initially decreases to maintain revenue equivalence, but eventually begins to rise again. This non-monotonic behavior arises from the interaction of two opposing forces. On one hand, increasing τ_1 allows the government to collect a larger share of earnings from high-income individuals, boosting revenue. On the other hand, higher τ_1 reduces the incentive to work, leading to lower output and, consequently, reduced tax revenue. Initially, the revenue-boosting effect outweighs the output-reducing effect, but as τ_1 continues to rise, the latter becomes dominant. In contrast, in the model without skill heterogeneity, τ_0 consistently decreases as τ_1 rises, suggesting that the output-reducing effect is always stronger in this scenario. Consequently, the revenue-maximizing level of τ_1 is lower in this calibration.

In both calibrations, increasing progressivity reduces incentives to work, particularly when people are older and have higher earnings. However, labor is more sensitive to changes in progressivity in the calibration with learning ability heterogeneity, and consequently, output and average wage growth respond more than in the model without learning ability heterogeneity.

Table 10. *Revenue-neutral increase in tax progressivity*

τ_1	τ_0	**Earnings Ineq.**	**Ouput**	**Welfare**	**Wage Growth**
0.181	0.185	100.0	100.0	100.0	100.0
0.191	0.185	93.5	99.0	97.3	99.2
0.211	0.181	79.9	97.0	91.4	97.5
0.231	0.185	65.3	94.8	84.6	95.6

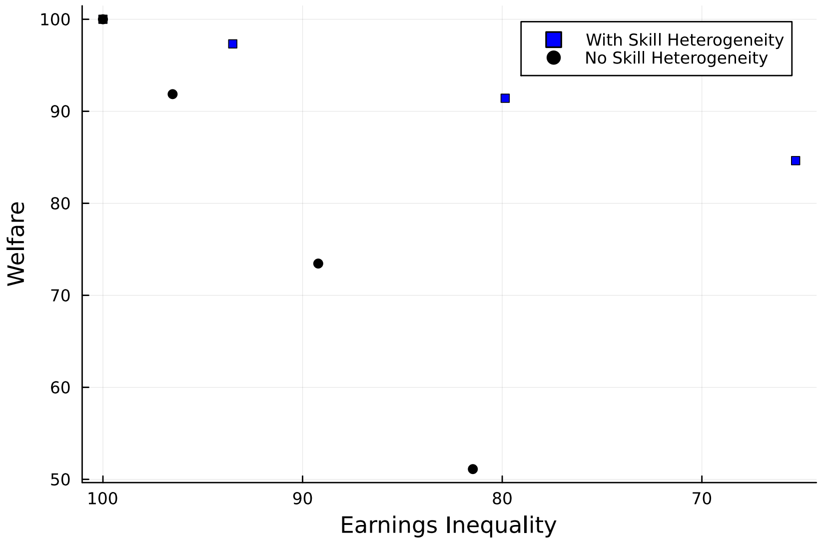
Table 11. *Revenue-neutral increase in tax progressivity, no skill heterogeneity*

τ_1	τ_0	**Earnings Ineq.**	**Ouput**	**Welfare**	**Wage Growth**
0.181	0.185	100.0	100.0	100.0	100.0
0.191	0.216	96.5	99.4	91.9	99.8
0.211	0.281	89.2	98.3	73.5	99.3
0.231	0.366	81.5	97.1	51.1	98.8

In Figure 9 I illustrate the differences in terms of the lifetime welfare cost of reducing inequality. As we can see, increasing tax progressivity in the model with skill heterogeneity is more effective at reducing lifetime earnings inequality than in the model without skill heterogeneity. This is because high ability people are more sensitive to changes in progressivity, which means they adjust their labor supply more in response to the increase in tax progressivity, therefore having a bigger impact in the increase of earnings inequality over the life cycle. We can also see that in the model with no skill heterogeneity, the same change in progressivity causes a larger reduction of lifetime welfare. There are two reasons behind this result. First, when all hours dispersion is due to leisure preferences inequality is a reflection of people's tastes, and therefore the gap in welfare between the low and high earners is smaller than in the model with learning ability heterogeneity. Consequently, bringing low and high earners closer in terms of earnings results in a smaller welfare gain from the point of view of the low earners. On the other hand, when learning ability is also responsible for some of the dispersion in hours, progressivity serves as a sort of insurance for being born with low learning ability.

The second reason for the bigger reduction in welfare is that in the model with no skill heterogeneity, to maintain revenue equivalence we need higher increases in the tax level parameter τ_0, which means that low earners are paying a relatively higher tax rate. The fact that people earn less when they are young means that under this more progressive scheme, young people are paying more taxes in the model with no skill heterogeneity relative to the model with skill heterogeneity. This loading of taxes on young people combined with time discounting, contributes to the welfare loss being higher in the model with no skill heterogeneity.

Figure 9. *Equality-Efficiency Trade-off*



Notes: Each point in the graph corresponds to a different level of progressivity τ_1 for each model. The values of progressivity tested in both models are: 0.181, 0.201, 0.221, 0.241. The scenario with $\tau_1 = 0.181$ corresponds to the base calibration, and both welfare and earnings inequality are normalized to 100 in this scenario for both models.

7 Conclusions

This paper provides new insights into the role of hours worked in shaping lifetime earnings inequality, emphasizing the importance of occupational choice, learning ability, and leisure preferences in driving the cross-sectional and between-occupation variance in hours over the life cycle.

I contribute to the empirical research on hours dispersion by documenting novel correlations between cognitive ability, hours worked, and occupational wage growth. Previous studies have highlighted the correlation between average wages and workweek length, but this paper is the first to examine the dynamic relationship between hours worked and wage growth across occupations. I find that the correlation between hours worked and wage growth is stronger for younger individuals, shedding light on how early work decisions influence long-term wage outcomes.

On the structural side, I present a life-cycle model of labor supply and occupational choice, with complementarity between learning ability and occupation-specific human capital accumulation at its core. This model successfully disentangles the contributions of learning ability and leisure preferences to lifetime hours worked and earnings inequality. Crucially, the model matches the observed decrease in both cross-sectional and between-occupation variance in hours over the life cycle, providing a robust framework for understanding these dynamics.

Through the calibration and estimation of the model, I demonstrate that differences in learning ability and leisure preferences explain a substantial portion of the variance in hours worked and earnings across occupations and over the life cycle. Specifically, learning ability accounts for a significant share of the variance in hours worked at younger ages and a considerable portion of earnings variance in later life stages. These findings imply that policy measures addressing earnings inequality must account for the role of hours worked and occupational sorting based on individual ability and preferences.

The policy experiments conducted in this paper provide important insights into how progressive taxation impacts inequality and other economic aggregates. In particular I show that changing progressivity in a world in which all labor decisions are based on leisure preferences has a different impact on labor, wage growth, output and welfare than in a world in which learning ability is an important source of hours dispersion. This suggests that the effectiveness of tax policies aimed at reducing inequality depends on the underlying sources of hours dispersion.

Future research could extend this framework by revisiting the question posed by Huggett et al. (2011)—how much of lifetime earnings inequality is due to initial conditions versus shocks—within a model that incorporates learning ability as a key driver of hours dispersion in early life. A similar extension could be applied to the study of hours worked by women, taking into account fertility decisions and see how they interact with occupational choice and learning ability heterogeneity.

In addition, the theory presented in this paper has predictions of what happens to labor, output, and other economic aggregates as a result of changes in the progressivity of the tax system. These predictions can be tested empirically to see if the model is able to replicate the effects of tax changes on the economy, which opens another avenue of future research.

By integrating learning ability and occupational choice into the analysis of hours dispersion, this paper not only advances the literature on earnings inequality but also opens the door to new avenues of research on how human capital accumulation and labor market policies shape lifetime inequality.

References

Badel, A., Huggett, M., and Luo, W. (2020). Taxing Top Earners: a Human Capital Perspective. *The Economic Journal*, 130(629):1200–1225.

Bénabou, R. (2002). Tax and Education Policy in a Heterogeneous-Agent Economy: What Levels of Redistribution Maximize Growth and Efficiency? *Econometrica*, 70(2):481–517.

Bick, A., Blandin, A., and Rogerson, R. (2022). Hours and Wages. *Quarterly Journal of Economics*, 137(4069859).

Bick, A., Blandin, A., and Rogerson, R. (2024). Hours Worked and Lifetime Earnings Inequality. (32997).

Bowlus, A. J. and Liu, H. (2013). The contributions of search and human capital to earnings growth over the life cycle. *European Economic Review*, 64:305–331.

Carrillo-Tudela, C. and Visschers, L. (2023). Unemployment and endogenous reallocation over the business cycle. *Econometrica*, 91(3):1119–1153.

Erosa, A., Fuster, L., Kambourov, G., and Rogerson, R. (2022). Labor Supply and Occupational Choice. (30492).

Guvenen, F., Kaplan, G., Song, J., and Weidner, J. (2022). Lifetime Earnings in the United States over Six Decades. *American Economic Journal: Applied Economics*, 14(4):446–479.

Guvenen, F., Kuruscu, B., and Ozkan, S. (2014). Taxation of Human Capital and Wage Inequality: A Cross-Country Analysis. *The Review of Economic Studies*, 81(2):818–850.

Heathcote, J., Storesletten, K., and Violante, G. L. (2010). The Macroeconomic Implications of Rising Wage Inequality in the United States. *Journal of Political Economy*, 118(4):681–722.

Heathcote, J., Storesletten, K., and Violante, G. L. (2017). Optimal Tax Progressivity: An Analytical Framework*. *The Quarterly Journal of Economics*, 132(4):1693–1754.

Heathcote, J., Storesletten, K., and Violante, G. L. (2020). Optimal progressivity with age-dependent taxation. *Journal of Public Economics*, 189:104074.

Heckman, J. J., Lochner, L., and Cossa, R. (2003). Learning-by-doing versus on-the-job training: using variation induced by the EITC to distinguish between models of skill formation. pages 74–130. Cambridge University Press, Cambridge.

Heckman, J. J., Lochner, L., and Taber, C. (1998). Explaining Rising Wage Inequality: Explorations with a Dynamic General Equilibrium Model of Labor Earnings with Heterogeneous Agents. *Review of Economic Dynamics*, 1(1):1–58.

Huggett, M., Ventura, G., and Yaron, A. (2006). Human capital and earnings distribution dynamics. *Journal of Monetary Economics*, 53(2):265–290.

Huggett, M., Yaron, A., and Ventura, G. (2011). Sources of Lifetime Inequality. *American Economic Review*, 101(7):2923–2954.

Jovanovic, B. (1979). Job Matching and the Theory of Turnover. *Journal of Political Economy*, 87(5):972–990.

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

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

Kambourov, G., Manovskii, I., and Plesca, M. (2020). Occupational mobility and the returns to training. *Canadian Journal of Economics/Revue canadienne d'Ã©conomique*, 53(1):174–211.

Lagakos, D., Moll, B., Porzio, T., Qian, N., and Schoellman, T. (2018). Life Cycle Wage Growth across Countries. *Journal of Political Economy*, 126(2):797–849.

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

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.

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

Schulhofer-Wohl, S. (2018). The age-time-cohort problem and the identification of structural parameters in life-cycle models. *Quantitative Economics*, 9(2):643–658.

Young, E. R. (2010). Solving the incomplete markets model with aggregate uncertainty using the Krusell-Smith algorithm and non-stochastic simulations. *Journal of Economic Dynamics and Control*, 34(1):36–41.

A Appendix: ACS descriptive statistics

Table 12. *ACS descriptive statistics and EWP estimates*

	occ_name	meanWage	meanHours	stdHours	EWP
1	Architecture and Engineering Occupations	3.58	43.71	7.02	1.78
2	Arts, Design, Entretainment, Sports, and Media Occupations	3.2	43.08	9.3	2.07
3	Building and Grounds Cleaning and Maintenance Occupations	2.59	40.48	7.41	1.62
4	Business Operations Specialists	3.45	44.38	8.31	1.92
5	Computer and Mathematical Occupations	3.61	42.81	6.53	1.85
6	Construction and Extraction Occupations	2.92	42.19	7.76	1.75
7	Counselors, Social and Religious Workers	3.01	43.7	8.97	1.6
8	Education, Training, and Library Occupations	3.24	43.11	9.17	2.05
9	Farming, Fishing, and Forestry Occupations	2.44	45.2	10.6	1.43
10	Food preparation and Serving Occupations	2.39	38.47	9.39	1.51
11	Healthcare Practitioners and Technical Occupations	3.67	45.25	11.22	2.51
12	Healthcare Support Occupations	2.7	39.98	8.6	1.59
13	Installation, Maintenance, and Repair Workers	3.02	43.3	7.64	1.81
14	Legal Occupations	3.9	47.49	9.59	2.63
15	Life, Physical and Social Science Occupations	3.44	43.55	7.8	2.35
16	Management, Business, Science and Arts Occupations	3.51	46.74	9.18	2.21
17	Office and Administrative Support Occupations	2.87	40.89	7.6	1.94
18	Personal Care and Service Occupations	2.57	40.2	9.76	1.72
19	Production Occupations	2.91	42.91	7.27	1.75
20	Protective Services Occupations	3.1	43.57	8.96	2.17
21	Sales and Related Occupations	3.08	43.9	9.65	2.24
22	Transportation and Material Moving Occupations	2.77	43.45	10.02	1.81

B Appendix: Checking robustness of correlation between EWP and hours worked

Figure 10. *Employment outcomes over the life cycle by Occupation type*

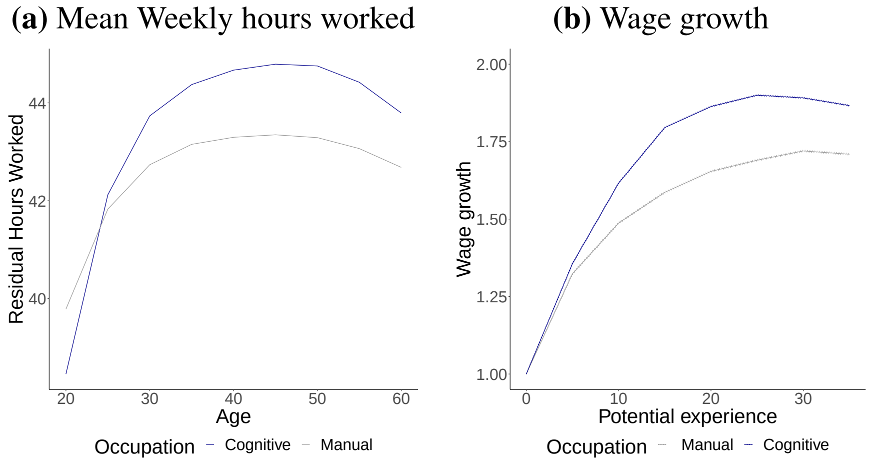
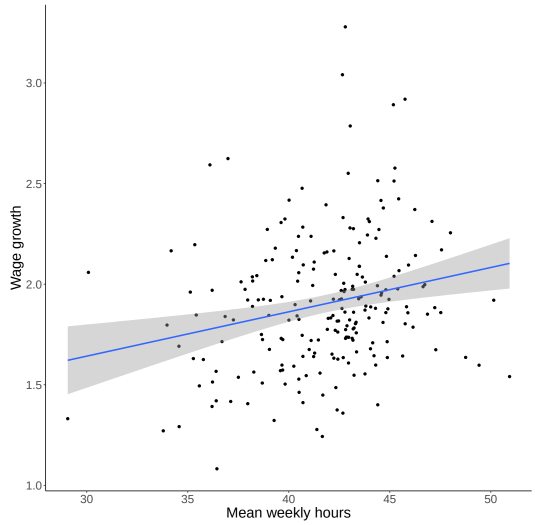


Figure 11. *Correlation EWP and hours worked by Occupation-Industry pairs*



C Appendix: Details of model solution

Since this is a life cycle model, I start by solving for the labor supply and value function workers at the last period:

$$V(\boldsymbol{h}_i, j, T; \xi_i, \psi_i) = \max_{c_i, \ell_{i,j}} \{u(c_i, \ell_{i,j})\},$$

$$\text{s.t.} \quad c_i = (1-\tau_0)\left(A(h_{i,k}\ell_{i,k})^{\zeta}\right)^{1-\tau_1}.$$

Since this is a static problem, we can find the optimal labor supply by solving the following first-order condition:

$$u_c(c_i,\ell_{ij})(1-\tau_0)(1-\tau_1)\zeta\left(A(h_{i,j}\ell_{i,j})^{\zeta}\right)^{-\tau_1}A(h_{i,j}\ell_{i,j})^{\zeta-1}h_{i,j}+u_{\ell}(c_i,,l_{ij})=0,$$

solving this equation we can find $\ell_T^j(\boldsymbol{h}_i,\chi_i,\xi_i)$, which we can plug into the utility function to find the value function $V_T^j(\boldsymbol{h}_i,\chi_i,\xi_i)$. To find $\ell_t^j(\boldsymbol{h}_i,\chi_i,\xi_i)$ and $V_t^j(\boldsymbol{h}_i,\chi_i,\xi_i)\ \forall \quad t=0\ldots T-1$, we cannot rely on an analytical solution, as labor in the current period affects future capital in all possible occupations in the future. For this reason, we need to approximate the solution by defining a grid for human capital, learning ability and leisure preferences, and solve the model by iterating backwards and finding the optimal labor for each combination of points in the grids.

It is worth noting that since taste shocks to the utility of picking a certain occupation are assumed to follow a Gumbel distribution with location parameter $\mu^{j,k}$, and scale parameter α, the worker's problem at $0<t<T-1$ can be rewritten as

$$V(\boldsymbol{h}_i,j,a;\xi_i,\psi_i)=\max_{c_i,\ell_{i,j}}\left\{u(c_i,\ell_{i,j})+\beta\sum_{k=1}^{J}q(\boldsymbol{h}_i',k,a+1;\xi_i,\psi_i)V(\boldsymbol{h}_i',k,a+1;\xi_i,\psi_i),\right\}$$

$$\text{s.t.} \quad c=(1-\tau_0)\left(A(h\ell)^{\zeta}\right)^{1-\tau_1},$$

$$\boldsymbol{h}_i'=H(\boldsymbol{h},j,\xi_i,\ell_i,k),$$

where

$$q(\boldsymbol{h}_i,j,a,k;\xi_i,\psi_i)=\frac{exp\left(\alpha\left(V(\boldsymbol{h}_i,k,a;\xi_i,\psi_i)+\mu^{j,k}\right)\right)}{\sum_{\iota\neq j}exp\left(\alpha\left(V(\boldsymbol{h}_i,\iota,a;\xi_i,\psi_i)+\mu^{j,\iota}\right)\right)+\exp(\alpha(V(\boldsymbol{h}_i,j,a;\xi_i,\psi_i)+\mu^{j,j}))}.$$

If we take derivatives of what is inside the $\max$ operator, we get

$$\begin{aligned}
&u_c(c,\ell)(1-\tau_0)(1-\tau_1)\zeta\left(A(h\ell)^{\zeta}\right)^{1-\tau_1}\ell^{-1}+u_\ell(c,\ell)\\
&\quad+\beta\left[q_{h'}(\boldsymbol{h}_i',j,a+1,j;\xi_i,\psi_i)H_\ell(\boldsymbol{h},j,\xi_i,\ell_i,k)V(\boldsymbol{h}_i',j,a+1;\xi_i,\psi_i)\right.\\
&\quad\left.+V_{h'}(\boldsymbol{h}_i',j,a+1;\xi_i,\psi_i)H_\ell(\boldsymbol{h},j,\xi_i,\ell_i,k)q(\boldsymbol{h}_i',j,a+1,j;\xi_i,\psi_i)\right]=0.
\end{aligned}$$

Taking derivatives of the value function with respect to h, H with respect to ℓ, and $q()$ with respect to ℓ we get that

$$V_h(\boldsymbol{h}_i,j,a;\xi_i,\psi_i)=u_c(c,\ell)(1-\tau_0)(1-\tau_1)\zeta\left(A(h\ell)^{\zeta}\right)^{1-\tau_1}h^{-1},$$

$$H_\ell(\boldsymbol{h},j,\xi_i,\ell_i,k)=\chi_{a,j}\theta^j\left(\xi\ell\right)^{\theta j-1}\xi,$$

$$q_h(\boldsymbol{h},j,a,k;\xi_i,\ell_i)=q(\boldsymbol{h},j,a,k;\xi_i,\ell_i)\left(1-q(\boldsymbol{h},j,a,k;\xi_i,\ell_i)\right)V_h(\boldsymbol{h}_i,j,a;\xi_i,\psi_i).$$

If we plug the values of the base calibration into these equations, that is $\sigma=1$, $\zeta=1$ into the previous expressions we get

$$V_h(\boldsymbol{h}_i,j,a;\xi_i,\psi_i)=(1-\tau_1)h^{-1},$$

$$H_\ell(\boldsymbol{h},j,\xi_i,\ell_i,k)=\chi_{a,j}\theta^j\left(\xi\ell\right)^{\theta j-1}\xi,$$

$$q_h(\boldsymbol{h},j,a,k;\xi_i,\ell_i)=q(\boldsymbol{h},j,a,k;\xi_i,\ell_i)\left(1-q(\boldsymbol{h},j,a,k;\xi_i,\ell_i)\right)(1-\tau_1)h^{-1}.$$

Then, the first order condition for labor becomes

$$\begin{aligned}
&(1-\tau_1)\ell^{-1}-\psi\ell^{\eta}+\beta\left[q_{h'}(\boldsymbol{h}_i',j,a+1,j;\xi_i,\psi_i)H_\ell(\boldsymbol{h},j,\xi_i,\ell_i,k)V(\boldsymbol{h}_i',j,a+1;\xi_i,\psi_i)\right.\\
&\quad\left.+V_{h'}(\boldsymbol{h}_i',j,a+1;\xi_i,\psi_i)H_\ell(\boldsymbol{h},j,\xi_i,\ell_i,k)q(\boldsymbol{h}_i',j,a+1,j;\xi_i,\psi_i)\right]=0.
\end{aligned}$$

D Appendix: Computing transferability of human capital matrix Γ

To compute a sensible measure of the correlation between the initial levels of occupation-specific human capital, I use the information available in the ONet portal. ONet gives a score in each of the 25 skills presented in Table 13 for each occupation. I use this information to compute the correlation between the initial levels of human capital in each occupation by using the following procedure.

Table 13. *List of ONet Skills*

	Skill.name
1	Complex Problem Solving
2	Management of Financial Resources
3	Management of Material Resources
4	Management of Personnel Resources
5	Time Management
6	Coordination
7	Instructing
8	Negotiation
9	Persuasion
10	Service Orientation
11	Social Perceptiveness
12	Judgment and Decision Making
13	Systems Analysis
14	Systems Evaluation
15	Equipment Maintenance
16	Equipment Selection
17	Installation
18	Operation and Control
19	Operations Analysis
20	Operations Monitoring
21	Programming
22	Quality Control Analysis
23	Repairing
24	Technology Design
25	Troubleshooting

I assume that all ONet skills, which I call x_k are iid and normally distributed with mean 1 and variance σ^2. I also assume that human capital in occupation j is a linear combination of these skills, where the weights are given by the importance score of each skill in occupation j described in ONet. I denote the importance score of skill k in occupation j as α_{jk}. Then, the human capital in occupation j is then given by:

$$h_j = \mu_j + \sum_{k=1}^{25} \alpha_{jk} x_k.$$

Then, the Variance-covariance matrix occupation-specific human capital is given by

$$\Sigma_{ij} = \sum_{k=1}^{25} \alpha_{ik}\alpha_{jk}\sigma^2.$$

We can then compute the correlation matrix between occupation-specific human capital as

$$\rho_{ij} = \frac{\sum_{k=1}^{25} \alpha_{ik}\alpha_{jk}}{\sqrt{\left(\sum_{k=1}^{25} \alpha_{ik}^2\right)\left(\sum_{k=1}^{25} \alpha_{jk}^2\right)}}.$$

This correlation matrix can be used to compute the transferability of human capital from one occupation to another.

E Appendix: Details of the estimation procedure

Given the production function, the logarithm of wages can be expressed as

$$\log(w_{ij}) = \log(A) + \zeta \log(h_{ij}) + (\zeta - 1)\log(\ell_{ij}),$$

which means we can obtain analytical expressions for the expectation and variance of log wages as a function of log human capital:

$$\mathbb{E}[\log(w_{ij})] = \log(A) + \zeta\mathbb{E}[\log(h_{ij})] + (\zeta - 1)\mathbb{E}[\log(\ell_{ij})],$$

$$\mathbb{V}[\log(w_{ij})] = \zeta^2\mathbb{V}[\log(h_{ij})] + (\zeta-1)^2\mathbb{V}[\log(\ell_{ij})] + (\zeta-1)\mathbb{C}\left(\log(h_{ij}), \log(\ell_{ij})\right).$$

In the base calibration of this model, I assume $\zeta = 1$, which means that wages are independent of labor, and therefore there is an exact mapping from the moments of wages and human capital. In fact, I now have that

$$\mathbb{E}[\log(w_{ij})] = \log(A) + \mathbb{E}[\log(h_{ij})],$$

$$\mathbb{V}[\log(w_{ij})] = \mathbb{V}[\log(h_{ij})].$$

Moreover, under this production function, the covariance between log wages and log skill is also equal to the correlation between log skill and log human capital.

$$\mathbb{C}[\log(w_{ij}), \log(testScore_i)] = \mathbb{C}[\log(h_{ij}), \log(\xi_i)].$$

Then, for every occupation j, the parameters of the initial distribution can be mapped to the following moments at age 23:

$$\tilde{h}_j = \mathbb{E}[\log(w_{ij})],$$

$$\sigma_{\tilde{h}_j} = \mathbb{V}[\log(w_{ij})],$$

$$\sigma_{\tilde{h}_j,\xi} = \mathbb{C}[\log(w_{ij}), \log(testScore_i)],$$

$$\tilde{\xi}_j = \mathbb{E}[\log(testScores_{ij})],$$

$$\sigma_{\xi_j} = \mathbb{V}[\log(testScores_{ij})].$$

F Appendix: Estimation Results

Table 14. *Standard deviation of initial log wages*

holahola	**Model**	**Data**
Architecture and Engineering Occupations	0.64	0.64
Arts, Design, Entretainment, Sports, and Media Occupations	0.76	0.76
Building and Grounds Cleaning and Maintenance Occupations	0.59	0.61
Business Operations Specialists	0.62	0.62
Computer and Mathematical Occupations	0.68	0.67
Construction and Extraction Occupations	0.62	0.61
Counselors, Social and Religious Workers	0.61	0.60
Education, Training, and Library Occupations	0.65	0.65
Farming, Fishing, and Forestry Occupations	0.61	0.60
Food preparation and Serving Occupations	0.58	0.58
Healthcare Practitioners and Technical Occupations	0.59	0.59
Healthcare Support Occupations	0.56	0.57
Installation, Maintenance, and Repair Workers	0.56	0.56
Legal Occupations	0.78	0.77
Life, Physical and Social Science Occupations	0.64	0.64
Management, Business, Science and Arts Occupations	0.63	0.64
Office and Administrative Support Occupations	0.57	0.58
Personal Care and Service Occupations	0.68	0.69
Production Occupations	0.57	0.58
Protective Services Occupations	0.60	0.60
Sales and Related Occupations	0.62	0.62
Transportation and Material Moving Occupations	0.59	0.60

Table 15. *Average log hours by occupation and Variance of log Hours worked*

Occupation	Model	Data	Diff.
Architecture and Engineering Occupations	-0.02	0.09	-0.11
Arts, Design, Entretainment, Sports, and Media Occupations	0.02	0.08	-0.06
Building and Grounds Cleaning and Maintenance Occupations	-0.01	0.02	-0.03
Business Operations Specialists	-0.02	0.10	-0.12
Computer and Mathematical Occupations	-0.02	0.07	-0.09
Construction and Extraction Occupations	-0.03	0.04	-0.07
Counselors, Social and Religious Workers	-0.03	0.08	-0.11
Education, Training, and Library Occupations	0.06	0.07	-0.02
Farming, Fishing, and Forestry Occupations	-0.06	0.11	-0.17
Food preparation and Serving Occupations	-0.03	-0.00	-0.03
Healthcare Practitioners and Technical Occupations	-0.09	0.11	-0.21
Healthcare Support Occupations	-0.07	0.00	-0.07
Installation, Maintenance, and Repair Workers	-0.03	0.08	-0.11
Legal Occupations	0.03	0.18	-0.14
Life, Physical and Social Science Occupations	-0.09	0.09	-0.18
Management, Business, Science and Arts Occupations	-0.05	0.15	-0.20
Office and Administrative Support Occupations	-0.08	0.04	-0.12
Personal Care and Service Occupations	-0.05	0.03	-0.07
Production Occupations	-0.04	0.07	-0.11
Protective Services Occupations	-0.06	0.10	-0.16
Sales and Related Occupations	0.01	0.11	-0.10
Transportation and Material Moving Occupations	-0.04	0.09	-0.12
Variance of hours at 55	0.03	0.03	-0.00
Variance of hours at 23	0.04	0.05	-0.00

Table 16. *Average wage growth by occupation*

Occupation	Model	Data	Diff.
Architecture and Engineering Occupations	1.07	1.78	-0.71
Arts, Design, Entretainment, Sports, and Media Occupations	1.14	2.07	-0.93
Building and Grounds Cleaning and Maintenance Occupations	1.16	1.62	-0.46
Business Operations Specialists	1.03	1.92	-0.89
Computer and Mathematical Occupations	1.04	1.85	-0.81
Construction and Extraction Occupations	1.19	1.75	-0.56
Counselors, Social and Religious Workers	1.09	1.60	-0.51
Education, Training, and Library Occupations	1.24	2.05	-0.81
Farming, Fishing, and Forestry Occupations	1.16	1.43	-0.27
Food preparation and Serving Occupations	1.06	1.51	-0.45
Healthcare Practitioners and Technical Occupations	0.95	2.51	-1.56
Healthcare Support Occupations	1.12	1.59	-0.47
Installation, Maintenance, and Repair Workers	1.16	1.81	-0.65
Legal Occupations	1.35	2.63	-1.28
Life, Physical and Social Science Occupations	0.97	2.35	-1.38
Management, Business, Science and Arts Occupations	0.85	2.21	-1.36
Office and Administrative Support Occupations	0.76	1.94	-1.18
Personal Care and Service Occupations	1.16	1.72	-0.56
Production Occupations	1.16	1.75	-0.59
Protective Services Occupations	1.21	2.17	-0.96
Sales and Related Occupations	0.81	2.24	-1.43
Transportation and Material Moving Occupations	1.16	1.81	-0.65