PERUVIAN ECONOMIC ASSOCIATION

Lifetime Hours Inequality and Occupational Choice

César Urquizo

Working Paper No. 208, April 2025

Lifetime Hours Inequality and Occupational Choice

César Urquizo *
Vanderbilt University

August 2026
Most recent version available here

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 heterogeneous 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 only 8% of the variance in log hours at age 23, while leisure preferences account for 80%. Despite its seemingly small contribution to hours dispersion, cognitive ability accounts for roughly twice as much of the variance of earnings at age 55 (54%) compared to leisure preferences (28%). 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	:	H21, J22, J24, J31, J62
Keywords	:	Labor Supply, Occupational Choice, Occupational Mobility, Life Cycle, Tax Progressivity, Inequality

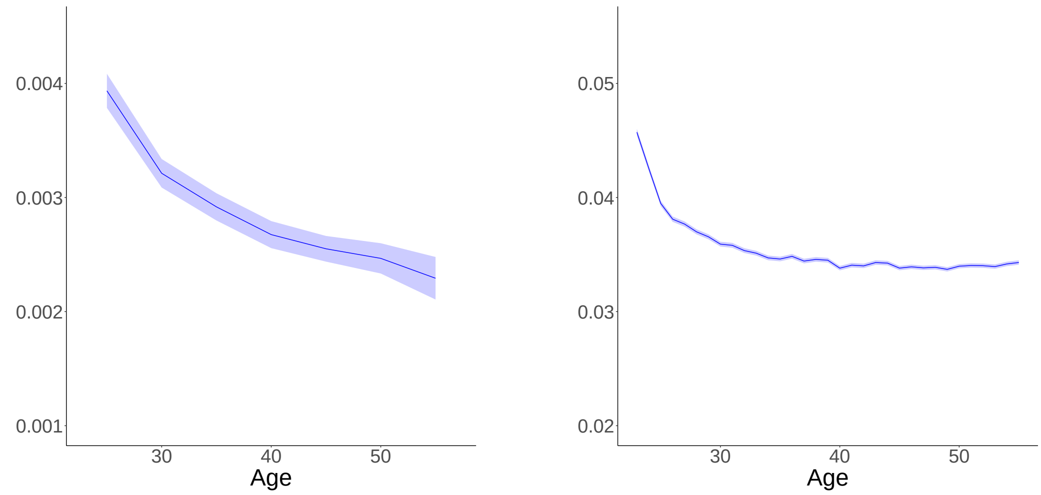
*I am grateful to my committee: Víctor Rios-Rull, Dirk Krueger, Iourii Manovskii, and Joachim Hubmer for their feedback and support on this project. I also thank Antonella Rivera, Daniel Jaar, Alberto Ramírez de Aguilar, Jordan Peeples, and participants at the Penn Macro Seminar, Vanderbilt Macro Seminar,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

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. Yet, if hours worked affect the accumulation of human capital and therefore future wages, differences in hours may have important consequences for earnings inequality over the life cycle beyond their direct effect on earnings. 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] The declining age profile points toward a role for human capital accumulation: incentives to supply hours for learning purposes are strongest early in life and naturally weaken as workers age. At the same time, existing theories that explain life-cycle and occupational differences in hours primarily through heterogeneity in leisure preferences (Bick et al., 2025; Erosa et al., 2022) do not naturally generate this pattern.

Figure 1. *Variance of Log Hours Worked Over the Life Cycle*

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



Notes: Author's calculations from the American Community Survey (ACS), 2003-2023. The sample includes males aged 23-55 working at least 20 hours per week. Shaded areas represent 95% confidence intervals. Between-occupation variance is computed as the variance of average log hours worked across occupation-age groups, where age groups are defined in five-year intervals. [2]

[1]This pattern also holds for females. However, a substantial share of women temporarily exit the labor force during their 20s, which complicates some of the assumptions underlying the empirical analysis. In addition, given the unequal distribution of childcare responsibilities within households, female labor supply is more responsive to family composition, which is beyond the scope of this document. For these reasons, the analysis focuses on males.

This raises an important question: What explains differences in hours worked across individuals and occupations, and how does the source of these differences shape earnings inequality over the life cycle? The literature has mainly relied on leisure preference heterogeneity to explain differences in hours worked over the life cycle and across occupations (Bick et al., 2025; Erosa et al., 2022). I relax this restriction by allowing hours dispersion to also arise from heterogeneity in learning ability. When occupations differ in how strongly current hours are rewarded through future wage growth, workers with different learning abilities face different incentives to supply hours. I use the data to discipline the relative importance of learning ability and leisure preferences in generating observed hours dispersion. This theory can naturally account for the declining age profiles presented in Figure 1, as well as the positive relationship between average wage growth, hours worked, and cognitive test scores across occupations documented later in the paper. It also generates a decline in occupational switching with age without relying on age-specific parameters. Using a life-cycle model of labor supply and occupational choice, I disentangle the roles of leisure preferences and learning ability in generating the dispersion in working hours and quantify their contributions to life-cycle hours and earnings inequality. Distinguishing between these two sources of hours dispersion has important normative implications: when hours dispersion is driven entirely by leisure preferences, progressive taxation is less effective at reducing lifetime earnings inequality and entails a slightly larger welfare cost than when learning ability also contributes to hours dispersion.

I begin by presenting several empirical facts on hours worked, wage growth, and cognitive ability over the life cycle, using data from the American Community Survey (ACS) 2003–2023 and the National Longitudinal Survey of Youth 1997 (NLSY97). Using the ACS, I estimate average life-cycle wage growth for 22 major Census occupations following Lagakos et al. (2018) to account for time and cohort effects in repeated cross-sectional data. I document a positive relationship between average hours worked and wage growth across occupations, which remains after controlling for current wages. I also show that occupations with higher average cognitive test scores exhibit higher wage growth. Importantly, the relationship between cognitive ability and hours worked varies systematically with occupational wage growth: cognitive scores are essentially uncorrelated with hours in occupations with low wage growth, but positively correlated with hours in occupations with high wage growth. This pattern is consistent with workers with greater learning ability supplying more hours precisely in occupations where hours are more strongly rewarded through future wage growth. The relationships between hours, cognitive ability, and wage growth are robust to alternative measures of wage growth constructed using panel data and to controlling for selection on observable and unobservable fixed characteristics.

Informed by the empirical evidence, I develop a structural life-cycle 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 is that workers with greater learning ability have stronger incentives to supply hours in occupations where working more generates greater future wage growth. In the model, this arises from a complementarity between learning ability and occupation-specific human capital accumulation technologies.

I estimate the parameters governing the model's key mechanisms using Generalized Method of Moments. The identification of learning ability and leisure preference heterogeneity exploits their different implications for hours dispersion over the life cycle. Drawing on a framework similar to Heckman et al. (1998), I assume that by age 55 incentives to work for human capital accumulation are negligible, so that the remaining dispersion in hours is attributable to heterogeneity in leisure preferences. I therefore use hours dispersion at age 55 to discipline the variance of leisure preferences, while the additional dispersion in hours at younger ages identifies the variance of learning ability. I use the between-occupation variance of hours at age 55 to discipline the correlation between learning ability and leisure preferences,[3] and data on cognitive test scores to discipline differences in the mean and variance of learning ability across occupations. Finally, wage data discipline the mean and variance of initial human capital by occupation and its correlation with learning ability. The estimated model successfully replicates the declining between-occupation and cross-sectional variance of hours documented in Figure 1, as well as the decline in occupational switching with age, without relying on age-varying preference or technology parameters.

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. The decomposition reveals a sharp distinction between the importance of learning ability for hours dispersion early in life and for earnings inequality later in life. Using

[3] In the absence of correlation between learning ability and leisure preferences, the between-occupation variance of hours converges toward zero as workers age. Its level at age 55 therefore provides information about the correlation between these two sources of heterogeneity.

Shapley values[4] to account for interaction effects, I find that differences in learning ability account for only 8% of the variance of log hours at age 23, compared with 80% attributable to leisure preferences. Yet these early differences are amplified through human capital accumulation: by age 55, learning ability accounts for 54% of the variance of log earnings, while leisure preferences account for only 28%. Thus, a source of heterogeneity that explains relatively little of hours dispersion early in life becomes a major determinant of earnings inequality later in life. More broadly, despite the modest direct contribution of hours to contemporaneous earnings inequality, I find that hours dispersion accounts for 57% of lifetime earnings inequality once its effects on human capital accumulation and future wages are taken into account.

Finally, I show that identifying the source of hours dispersion matters for the evaluation of redistributive policy. I compare the estimated model with a recalibrated restricted model in which learning ability is homogeneous and all hours dispersion is attributed to leisure preferences. Increasing tax progressivity is more effective at reducing lifetime earnings inequality in the model with learning heterogeneity and entails a slightly smaller welfare loss. Consequently, the welfare cost of reducing lifetime earnings inequality by one percentage point is approximately 4.0% of lifetime consumption in the estimated model, compared with 6.5% in the preference-heterogeneity-only model.[5]

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.

[4]Because this section focuses exclusively on the roles of learning ability and leisure preferences, the Shapley decompositions reported here do not sum to 100%. The remainder is attributable to other sources of heterogeneity in the model, including initial human capital, differences across occupations, and taste shocks.

[5]Two forces drive this difference. First, high earners tend to have high learning ability and are more responsive to changes in progressivity, making progressivity more effective at reducing earnings inequality. Second, when hours dispersion is driven entirely by leisure preferences, earnings differences partly reflect differences in tastes rather than differences in welfare. Progressive taxation therefore provides less valuable insurance than when earnings differences also reflect heterogeneity in learning ability. See Section 6 for a detailed discussion.

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. Checchi et al. (2022) show that hours dispersion and its covariance with wages are an important, and growing, driver of earnings inequality in the UK and Germany, though not in the US and France, over the period 1991-2016. Bick et al. (2025) extends this work, emphasizing how differences in leisure preferences –both permanent and transitory[6]– can drive hours and earnings dispersion over the life cycle. [7]

I build on the emerging literature about hours inequality (Bick et al., 2025) 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, labor supply, and earnings inequality. Using a new database covering hours worked in 160 countries, Gethin and Saez (2026) document a strong negative association between labor income taxes and hours worked across countries, although this relationship weakens once social spending is accounted for and disappears after also controlling for working-hours regulations. Closer to my policy experiment, Badel et al. (2020) examines the taxation of top earners in a model with human capital accumulation, while Guvenen et al. (2014) studies how progressive taxation affects wage inequality. I contribute to this literature by evaluating the effectiveness of tax progressivity in reducing inequality and its associated welfare cost. Specifically, I compare outcomes in a model where all hours dispersion stems from preference heterogeneity with those in a model where part of the dispersion is driven by heterogeneity in learning ability. I find that, in the presence of learning heterogeneity, increasing progressivity is

[6]Bick et al. (2025) 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.

[7]While Bick et al. (2025) 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.

more effective at reducing lifetime earnings inequality and entails a slightly lower welfare cost 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. This dynamic mechanism can account for the declining between-occupation dispersion of hours worked over the life cycle and is consistent with the empirical finding that, conditional on current wages, individuals work longer hours in occupations with higher average wage growth.

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 three novel empirical facts that motivate the main mechanism in my model: a positive, statistically significant, correlation between occupation-specific average wage growth and workweek, a positive, statistically significant, correlation between occupation-specific average wage growth and cognitive test scores, and a correlation between hours worked and cognitive scores that is stronger for occupations with high average wage growth.

2.1 Data

The empirical facts presented in this section come mainly from two data sources: American Community Survey (ACS) 2003-2023, and National Longitudinal Survey of Youth 1997 (NLSY97). Throughout this section I focus mainly on highly attached males aged 23-55, defined as individuals that responded that their usual week throughout the past year was 20 hours or more. This restriction helps exclude workers with weak labor force attachment and part-time workers who may hold multiple jobs/occupations[8]. In line with other papers in this literature, I exclude women from the sample because female labor supply at younger ages is more closely linked to fertility decisions and women experience more frequent spells outside the labor force during their 20s. I do not attempt to model these margins in this paper.

I start the sample at age 23, as the average age of college graduation in the U.S. is between 23 and 24. A potential concern is that the decline in the cross-sectional variance of hours is driven by young individuals who work while still in school. Starting the sample at age 25 reduces the magnitude of the decline: the variance of hours at age 23 is approximately 33% higher than at age 55, whereas at age 25 it is approximately 15% higher. Although the decline is smaller, it remains substantial, suggesting that simultaneous work and schooling cannot fully account for the life-cycle pattern in hours dispersion. Moreover, it is not clear that work while in school can explain the decline in the between-occupation variance of hours documented above. I end the sample at age 55 because, beyond this age, the fraction of individuals who partially or fully retire begins to increase

[8] Within this sample, 88% of workers reported working 52 weeks over the past year, and 93% reported working more than 48 weeks.

substantially, introducing an additional margin of labor supply adjustment that is outside the scope of this paper.

A final note on sample selection is that I exclude three of the 25 Census occupations: Financial Specialists, Extraction Workers, and Military Specific Occupations. These occupations are excluded due to small sample sizes or the unavailability of the data required for model calibration in one or more of the datasets used in the analysis.

From ACS 2003-2023 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 15 in Appendix B.

Table 1. *ACS and NLSY97 sample descriptions*

Statistic	**ACS**	**NLSY97**
Observations	14,172,908	34,605
Years	2003-2023	1998-2019
Frequency	Annual	Annual
Workweek Range	21-80	21-80
Age Range	23-55	23-39
# 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.

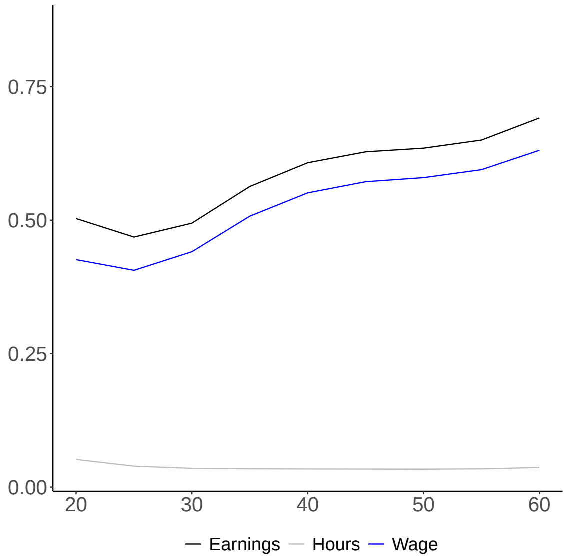
As mentioned earlier, a common concern when working with cross-sectional data to infer about life cycle behavior is that we might be confusing time, cohort and age effects. I deal with this potential problem in multiple ways. Firstly, I try computing the same statistics with ACS by restricting the sample to 2003-2013 and compare it to the 2013-2023 sample as well as the full sample, finding no significant changes in the observed patterns. Secondly, I compare 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 2003-2023. It has been widely documented that the dispersion of wages and earnings increases with age, but there has 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 0.0039 to 0.0023, which means there is a nearly 40% drop in the between-occupation variance of hours over the life cycle.[9] Figure 1(**b**) on the other hand, shows the cross-sectional variance of hours worked drops from 0.047 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 does not remain constant over time.

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



Notes: Author's calculations from the American Community Survey (ACS), 2003-2023.

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

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

[9]This decline is computed using the same five-year age bins plotted in Figure 1(**a**). As a robustness check, computing the between-occupation variance at the exact single ages of 23 and 55, rather than pooling within five-year bins, yields an even larger decline, from 0.0051 to 0.0021, a 58% drop (restricting the sample to 2005 onward, as in Figure 1(**b**), yields a decline from 0.0055 to 0.0025, a 54% drop).

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 specificly, 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 2 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 15 in Appendix B.

Table 2. *Slope of life cycle wage profile by occupation*

Statistic	**Wage Growth**
Full sample average	1.38
Maximum	1.89
Minimum	1.19
Standard deviation	0.21
Range	0.70
Range / Sample mean	0.51

Notes: Author's calculations from the American Community Survey 2003-2023.

As shown in Table 2 there is significant dispersion in wage growth across occupations. In the full sample average, a person at age 50 earns about 1.38 times their salary at age 20. When we look at this figure by occupation, we can see that the occupation with the highest EWP has a slope of 1.89, whereas the one with the lowest has a slope of 1.19. This means that the difference in the wage growth of the highest and lowest wage growth occupations is 70 percentage points. The standard deviation across occupations is also sizable, at 0.21. This means that the dispersion in wage growth across occupations, measured by the range, is about 51% 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

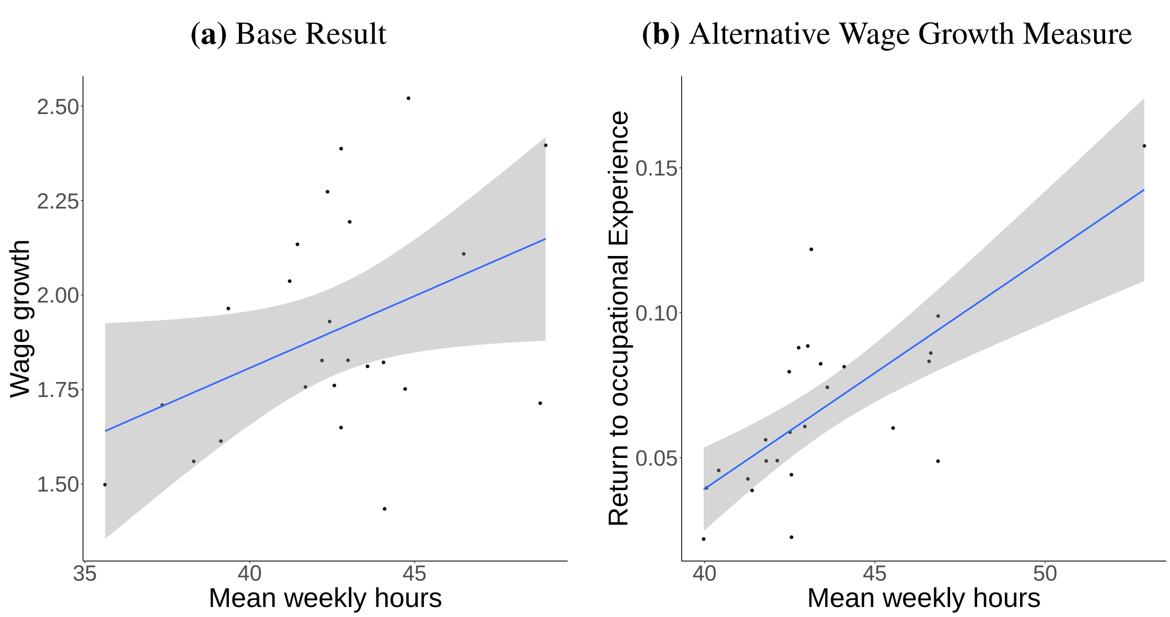
$$ln(w_{it}) = \alpha_i + \theta\ schooling_{it} + \sum_{j=1}^{J} \delta^j exp_{it}^j \times D_{it}^j + \gamma otherExp_{it} + \varepsilon_{it},$$

where i denotes an individual, t denotes time, α_i is an individual fixed effect, $schooling_{it}$ are the years of schooling, exp^j_{it} 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 3**(a)**. As we can see, there is a positive, statistically significant correlation (0.55) 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 3**(b)**. This correlation is also present when I restrict the samples to the periods 2003-2013 and 2013-2023.

Figure 3. *Correlation between wage growth and workweek*



Notes: Author's calculations from the American Community Survey 2003-2023, and the 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 C.

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

$$\log(\ell_{ij}) = \beta_0+\beta_1 EWP_j+\beta_2 \log(w_{ij})+\beta_3 age_{ij}+\beta_4 age_{ij}\times EWP_j+\beta_5 child_{ij}+\beta_6 marr_{ij}+\varepsilon_{ij}, \quad (2)$$

where EWP_j is the slope of the Experience Wage Profile 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, $child_{ij}$ denotes the number of children of individual i in occupation j, and $marr_{ij}$ is a binary variable that takes the value of 1 if individual i in occupation j has ever been married. The results of this regression are shown in Table 3. As we can see in all specifications, the coefficient of the Experience Wage Profile 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 [10].

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.

[10]The coefficient on the married indicator appears to run counter to the results in Blandin et al. (2025), who find that married men work more. This difference may reflect the fact that some effects of marriage are captured by the number of children variable. In addition, occupation choice may account for part of the labor shocks emphasized by the authors, which could help explain why the estimated coefficient is negative and very small in absolute value.

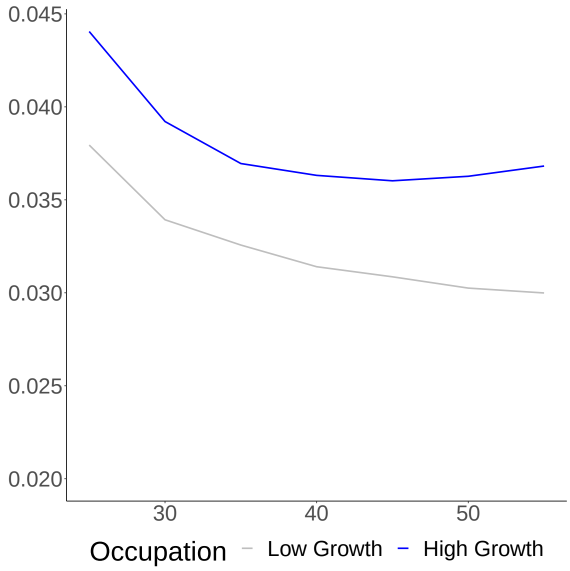
Table 3. *Regressions of log weekly hours worked on occupation average wage growth*

	Dependent variable:					
	$\log(hours)$					
	Full Sample	Full Sample	Ages 23-30	Ages 50-60	Full Sample	Full Sample
	(1)	(2)	(3)	(4)	(5)	(6)
Wage Growth	0.171***	0.145***	0.173***	0.135***	0.265***	0.224***
	(0.0004)	(0.0004)	(0.001)	(0.001)	(0.002)	(0.002)
Age		0.0002***			0.004***	0.002***
		(0.00001)			(0.00005)	(0.0001)
$\log(wage)$		0.016***	0.023***	0.009***	0.017***	0.009***
		(0.0001)	(0.0002)	(0.0002)	(0.0001)	(0.0001)
# Children						0.009***
						(0.0001)
Married						−0.005***
						(0.0001)
Wage Growth × Age					−0.003***	−0.002***
					(0.00004)	(0.00004)
Constant	3.518***	3.493***	3.420***	3.549***	3.324***	3.434***
	(0.001)	(0.001)	(0.001)	(0.001)	(0.002)	(0.003)
Observations	9,867,523	9,867,523	1,751,347	2,594,844	9,867,523	8,137,445
Adjusted R^2	0.016	0.020	0.025	0.013	0.021	0.018
Note:						*p<0.1; **p<0.05; ***p<0.01

Notes: Author's calculations from the American Community Survey 2003-2023.

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*

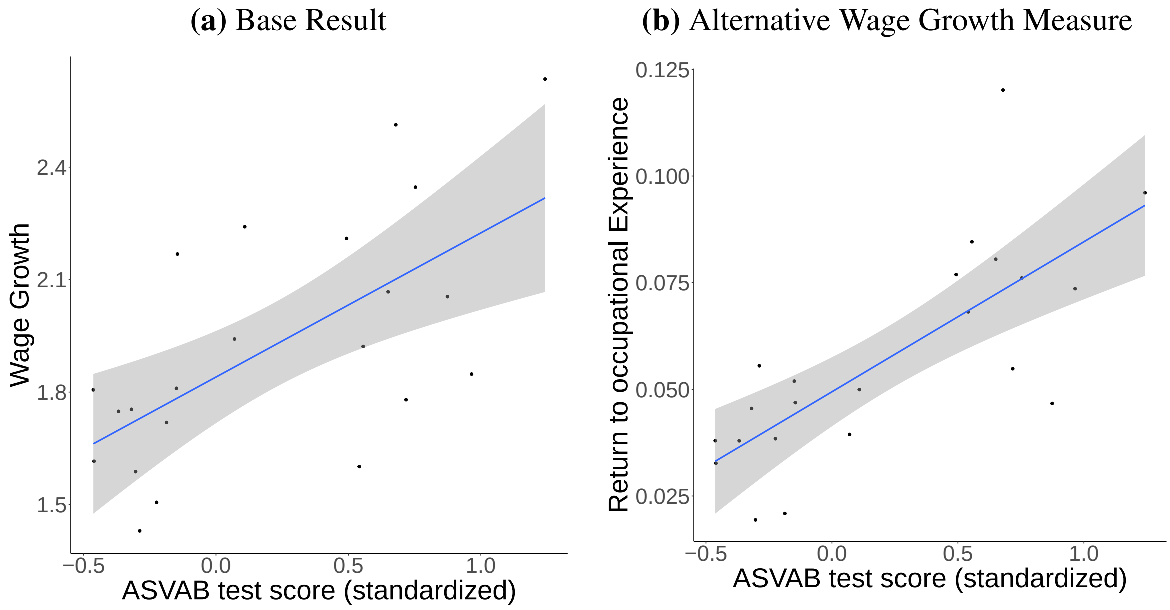


Notes: Author's calculations from the American Community Survey 2003-2023.

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 0.48 and 0.75, respectively, and both are statistically significant.

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



Notes: Author's calculations from the American Community Survey 2003-2023, and the 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 1.7% more hours per week. Although this number might seem small, a 1.7% longer workweek 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 4 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 4. *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		

Notes: Author's calculations from the 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 by featuring supermodularity between learning ability and occupation-specific human capital accumulation technology. The model is designed to capture how dispersion in hours, wages and earnings evolve through the life cycle in the presence of occupations with different learning-by-doing technologies.

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 is 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:

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

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 $a_0 < a < T_{ret}$, the problem of a worker with learning ability parameter ξ_i, leisure preference parameter ψ_i, employed in occupation j is given by

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

$$\text{s.t.} \quad c_i = \lambda \left(A (h_{i,j} \ell_{i,j})^{\varsigma} \right)^{1-\tau} + Tr, \tag{3}$$

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

where ε^{st} is a taste shock that affects the utility of staying in the current occupation and ε^{sw} is a shock to the utility of switching, and they both follow an extreme value distribution with scale parameter α_s and location parameter μ^{st} and μ^{sw}, respectively. 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. λ is a scale parameter governing the overall level of after-tax income, while τ captures the degree of tax progressivity, with higher values of τ implying a more progressive tax schedule. 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. $\hat{V}^{sw}(\cdot)$ is the value of switching to another occupation, which is defined as

$$\hat{V}^{sw}(\boldsymbol{h}_i, j, a; \xi_i, \psi_i) = \mathbb{E}\left[\max\left\{V(\boldsymbol{h}_i', k, a+1; \xi_i, \psi_i) + \varepsilon^{occ}\right\}_{k=1,\ k\neq j}^{J}\right],$$

where ε^{occ} is a taste shock that affects the utility of switching to a particular occupation $k \neq j$, and follows an extreme value distribution with scale parameter α_{occ} and location parameter μ^{occ}.

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 liquid wealth. Differences in initial wealth across individuals and occupations will partially be accounted by differences in the initial distribution at age 23[11].

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.

Since ε^{sw}, and ε^{st} follow an extreme value distribution, I can derive an analytical expression for the probability of staying in the current occupation, which is given by

[11] 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.

$$p^{st}(\boldsymbol{h}_i, j, a; \xi_i, \psi_i) = \frac{exp\left(\alpha_s\left(V(\boldsymbol{h}_i, j, a; \xi_i, \psi_i) + \mu^{st}\right)\right)}{exp\left(\alpha_s\left(V(\boldsymbol{h}_i, j, a; \xi_i, \psi_i) + \mu^{st}\right)\right) + \exp(\alpha_s(\hat{V}^{sw}(\boldsymbol{h}_i, j, a; \xi_i, \psi_i) + \mu^{sw}))}. \quad (4)$$

Moreover, given that the taste shocks ε^{occ} follow an extreme value distribution with scale parameter α_{occ} and location parameter μ^{occ}, I can also derive analytical expressions for the choice probabilities of each occupation, as well as the expected value of switching, which will be given by

$$\hat{V}^{sw}(\boldsymbol{h}_i, j, a; \xi_i, \psi_i) = \frac{1}{\alpha_{occ}} \log\left(\sum_{k=1,\ k\neq j}^{J} \exp\left(\alpha_{occ}\left[V(\boldsymbol{h}_i', k, a+1; \xi_i, \psi_i) + \mu^{occ}\right]\right)\right) + \frac{\gamma}{\alpha_{occ}}, \quad (5)$$

where γ is the Euler-Mascheroni constant. The analytical expression for the choice probabilities of each occupation k when currently in occupation j will be given by

$$p^{occ}(\boldsymbol{h}_i, j, a, k; \xi_i, \psi_i) = \frac{\exp\left(\alpha_{occ}\left(V(\boldsymbol{h}_i, k, a; \xi_i, \psi_i) + \mu^{occ}\right)\right)}{\sum_{\iota=1,\iota\neq j}^{J} \exp\left(\alpha_{occ}\left(V(\boldsymbol{h}_i, \iota, a; \xi_i, \psi_i) + \mu^{occ}\right)\right)}. \quad (6)$$

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

$$H(h_j, j, \xi_i, \ell_{i,j}, k) = \begin{cases} \left[\chi_{a,j}(h_{i,j})\left(\xi_i \ell_{i,j}\right)^{\theta_j} + (1-\delta_j)\right] h_{ij} & \text{if } j = k, \\ \Gamma_{j,k} h_{ij} & \text{otherwise,} \end{cases} \quad (7)$$

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, $\chi_{a,j}(\cdot)$ 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[12]. All the off-diagonal elements of Γ are smaller than 1, which means changing occupations has a cost in terms of human capital. 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

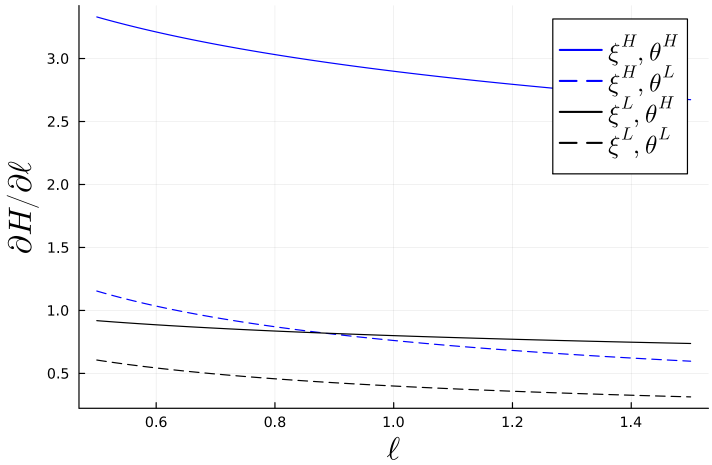
$$\frac{\partial H(h_j, j, \xi^H, \ell, j)}{\partial \theta_j} > \frac{\partial H(h_j, j, \xi^L, \ell, j)}{\partial \theta_j} \quad \forall\, \xi^H > \xi^L.$$

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 probability of staying in a given occupation will be driven by differences between the expected value of an individual for being employed at occupation j vs switching. Given that switching occupations always implies a loss in human capital, the difference in the value of staying vs switching will be increasing in human capital. In addition, since average human capital will tend to increase with age, the probability of switching will tend to decrease with

[12] 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



age, which is consistent with the evidence. Moreover, since the value of switching will be higher for occupations with higher θ_j, the probability of switching to an occupation with high θ_j will be higher than switching to an occupation with low θ_j.

From Equation 6, we can see that conditional on switching, individuals will be more likely to choose an occupation with higher expected value. This, together with Equation 7 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 of working age, at age $a = T_{ret}$, dynamic incentives to work disappear, and therefore workers only maximize their instant utility plus some continuation value which depends only on their human capital level. This simplifies the problem to:

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

$$\text{s.t.} \quad c_i = \lambda \left(A(h_{i,k}\ell_{i,k})^{\zeta} \right)^{1-\tau} + Tr, \tag{8}$$

where

$$V^{ret}(h_{i,k}) = \sum_{t=T_{ret}}^{T} \beta^{t-T} \left[u\left(\phi_{ret} A(h_{i,k}\bar{\ell})^{\zeta}, 0 \right) \right],$$

where T_{ret} is the age at which individuals retire and $\bar{\ell}$ is the average labor supply in the economy. This specification is meant to capture that when retired, individuals will not work, but will be able to consume a fraction ϕ_{ret} of their last period's pre-tax income.

3.4 Model Solution

Given the policy functions for labor supply $\ell(\boldsymbol{h}_i, j, a; \xi_i, \psi_i)$ and occupation switching probabilities $p(\boldsymbol{h}_i, j, a, k; \xi_i, \psi_i)$, as well as the conditional choice probabilities $p^{occ}(\boldsymbol{h}_i, j, a, k; \xi_i, \psi_i)$, I construct the transition matrices following Young (2010). Specifically, I use a non-stochastic simulation method to obtain $\{\Lambda_a\}_{a=a_0}^{T}$.

Each transition matrix Λ_a summarizes the joint evolution of the cross-sectional distribution of workers across human capital levels and occupations, incorporating optimal labor supply, occupation switching decisions, and the law of motion for human capital.

Given an initial distribution of workers at age a_0,

$$x(\boldsymbol{h}, a_0; \xi, \psi) = \{x(\boldsymbol{h}, j, a_0; \xi, \psi)\}_{j=1}^{J},$$

the distribution at subsequent ages is obtained by forward iteration:

$$x(\boldsymbol{h}, a; \xi, \psi) = \Lambda_a \, x(\boldsymbol{h}, a-1; \xi, \psi), \quad \forall a = a_0 + 1, \ldots, T.$$

This procedure yields the sequence of age-specific distributions

$$\{x(\boldsymbol{h}, j, a; \xi, \psi)\}_{j=1}^{J}{}_{,a=a_0}^{T},$$

which fully characterizes the evolution of a cohort over the life cycle.

Using these distributions, I compute aggregate moments for each occupation at every age. In a stationary environment with a constant inflow of entrants, these cohort profiles map directly into cross-sectional moments in the stationary equilibrium.

3.5 Stationary Recursive Equilibrium

Take as given the production-side parameters A and ζ, the tax schedule (λ, τ), the retirement replacement rate ϕ_{ret}, and the demographic structure.

A stationary recursive equilibrium consists of:

1. value functions $\{V(\boldsymbol{h}, j, a; \xi, \psi)\}_{a=a_0}^{T}$;
2. policy functions for labor supply $\{\ell(\boldsymbol{h}, j, a; \xi, \psi)\}_{a=a_0}^{T_{ret}}$;
3. occupation staying probabilities $\{p^{st}(\boldsymbol{h}, j, a; \xi, \psi)\}_{a=a_0}^{T_{ret}}$;
4. conditional choice probabilities $\{p^{occ}(\boldsymbol{h}, j, a, k; \xi, \psi)\}_{a=a_0}^{T_{ret}}$;
5. a sequence of age-specific distributions $\{x_a(\boldsymbol{h}, j, \xi, \psi)\}_{a=a_0}^{T}$;

such that:

(i) Given prices and policies, value and policy functions solve the worker's dynamic programming problem in Equations (3)–(6), and (8).

(ii) Occupation switching and choice probabilities are consistent with value functions and the extreme-value shocks.

(iii) The distribution evolves according to

$$x_a = \Lambda_a x_{a-1}.$$

(iv) The demographic structure is stationary: a unit mass of new workers enters at age a_0 each period, and workers exit at age T, so that the cross-sectional distribution across ages is constant over time.

(v) Firms are competitive and pay marginal productivity:

$$y(\boldsymbol{h}, j, \ell) = A(h_j \ell)^{\zeta}.$$

(vi) **Government budget constraint:** total tax revenues equal total expenditures,

$$\int T(y_i)\, di = \int \mathbf{1}\{a_i \geq T_{ret}\} \cdot \phi_{ret}\, y_i^{\text{last}}\, di + G^{other} + Tr,$$

where $T(y) = y - \lambda y^{1-\tau}$ and y_i^{last} denotes pre-retirement earnings.

This is a partial equilibrium model. However, under the assumption that tax revenues are used to finance retirement benefits and any residual revenues are spent unproductively from the point of view of workers (that is $Tr = 0$, and G^{other} is such that the government budget is balanced), the government budget constraint closes the model. Since firms are competitive and make zero profits, and there are no aggregate state variables, the equilibrium can be interpreted as a stationary general equilibrium with exogenous prices.

4 Calibration and parameter estimation

In this section, I describe the process used to calibrate the model and present the model fit for 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. This assumption also allows me to infer the correlation between learning ability and disutility of labor by looking at the between occupation variance in log hours at age 55. I use this insight to disentangle and separately estimate the effects of learning ability and preference heterogeneity in driving hours and earnings 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 $\{x(\boldsymbol{h}_i, j, 0, \xi_i, \psi_i)\}_{j=1}^{J}$ 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

$$e^j = \int x(\boldsymbol{h}, j, a_0; \xi, \psi) dF(\boldsymbol{h}, a_0; \xi, \psi),$$

as the initial mass of workers in each occupation. This has a very clear empirical counterpart which is the share of young people (age 23) 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:

$$\begin{bmatrix} \log(h_{ij}) \\ \log(\xi_i) \\ \log(\psi_i) \end{bmatrix} \sim N \left(\begin{bmatrix} \tilde{h}_j \\ \tilde{\xi} + \tilde{\xi}_j \\ \tilde{\psi} \end{bmatrix}, \begin{bmatrix} \sigma^2_{\tilde{h}_j} & \sigma_\xi \sigma_{\xi,j} \sigma_{\tilde{h}_j} \rho_{\tilde{h}_j \xi} & \sigma_{\tilde{h}_j} \sigma_\psi \rho_{\tilde{h}_j \psi} \\ \sigma_\xi \sigma_{\xi,j} \sigma_{\tilde{h}_j} \rho_{\tilde{h}_j \xi} & \sigma^2_\xi \sigma^2_{\xi,j} & \sigma_\xi \sigma_{\xi,j} \sigma_\psi \rho_{\xi\psi} \\ \sigma_{\tilde{h}_j} \sigma_\psi \rho_{\tilde{h}_j \psi} & \sigma_\xi \sigma_{\xi,j} \sigma_\psi \rho_{\xi\psi} & \sigma^2_\psi \end{bmatrix} \right). \tag{9}$$

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, dispersion and correlation with initial wages of log cognitive test scores by occupation should provide a good approximation of the differences in the mean ($\tilde{\xi}_j$), variance ($\sigma_{\xi,j}$) and correlation with initial human capital ($\rho_{\tilde{h}_j,\xi}$) 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. Just like the mean of learning ability is scaled by a constant $\tilde{\xi}$, all variance and covariance terms of ξ will be scaled by the parameter σ_ξ, which I will estimate to match the variance of hours at age 23.

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_{pre}:

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

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 $\rho_{\tilde{h}_j,\xi}$ is the correlation between initial log human capital and log learning ability. All these parameters are occupation-specific, which is why they are indexed by j.

In fact, the result of minimizing the moment condition will in these cases yield analytical solutions for the elements of Θ_{pre}. 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_{struct};

$$\Theta_{struct} = \left\{ \theta_1, \ldots, \theta_J, \chi, \sigma_\xi, \sigma_\psi, \rho_{\xi\psi}, \tilde{\psi}, \alpha_s, \mu^{st} \right\},$$

where θ_j is the human capital accumulation technology for each occupation, χ is a parameter in the human capital accumulation function that denotes a general learning technology, σ_ξ is the scale parameter for the standard deviation of learning ability, σ_ψ is the standard deviation of labor disutility, $\rho_{\xi\psi}$ is the correlation between learning ability and labor disutility, $\tilde{\psi}$ is the mean of labor disutility, and α_s and μ^{st} are, respectively, the scale and location parameters for the Gumbel shock to the value of staying in the same occupation.

I will refer collectively to the elements of these two sets as $\Theta = \{\Theta_{pre}, \Theta_{struct}\}$. 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 $\mathcal{G}_{pre}(\Theta_{pre})$ and $\mathcal{G}_{struct}(\Theta_{pre}, \Theta_{struct})$, respectively. Let us also define $\bar{\mathcal{G}}_{pre}$ and $\bar{\mathcal{G}}_{struct}$ 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

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

where W_{pre} is a weighting matrix. With the solution $\hat{\Theta}_{pre}$ I can then obtain the estimate $\hat{\Theta}_{struct}$ by solving

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

where W_{struct} 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 $\left[\mathcal{G}_{pre}(\Theta_{pre}) - \bar{\mathcal{G}}_{pre}\right]$ will yield exact analytical solutions for Θ_{pre} given that each moment in the set $\mathcal{G}_{pre}(\Theta_{pre})$ depends on a single element of Θ_{pre}. This means I just need to solve for each element in Θ_{pre} such that each of the relevant moment conditions are satisfied. On the other hand, the moment conditions for the parameters in Θ_{struct} will not yield exact analytical solutions.

The choice for the moments $\bar{\mathcal{G}}_{pre}$ is straightforward, since these moments do not depend on endogenous choices, and therefore I can pin down exactly which element in $\bar{\mathcal{G}}_{pre}$ will be affected by which element in Θ_{pre}. Moreover, these moments will have exact empirical counterparts, which are defined in table 5. However, the choice for the moments $\bar{\mathcal{G}}_{struct}$ 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 $\mathcal{G}_{struct}(\hat{\Theta}_{pre}, \Theta_{struct})$ and $\bar{\mathcal{G}}_{struct}$ is much more convoluted than the mapping between $\mathcal{G}_{pre}(\Theta_{pre})$ and $\bar{\mathcal{G}}_{pre}$, I rely on a global optimizer to find the values of Θ_{struct} that minimize the distance between the model moments and the empirical moments. For a detailed account of which moments are matched to the set $\mathcal{G}(\Theta_{pre})$, and which moments are considered to be more informative about each parameter in Θ_{struct}, see Table 5. Additional details of the estimation process are provided in Appendix E.

Table 5. *Estimated Parameters*

Estimated Parameters (Θ_{pre})			
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
$\rho_{h_j,\xi}$	Corr(human capital, learning ability)	Corr(wages, test scores) by occ., age 23	NLSY
Estimated Parameters (Θ_{struct})			
Parameter	**Description**	**Moment**	**Source**
θ_j	Human capital accumulation technology	Occupation-specific avg. wage growth	ACS
χ	General learning technology	Full sample 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
$\rho_{\xi\psi}$	Correl. learning ability and labor disutility	Between Occ. $V(\ln(hours))$	ACS
$\tilde{\psi}$	Mean log disutility of labor	Mean hours worked	Normalization
α_s	Scale parameter occ. switching shocks	Occupation Switching Prob., age 23	Normalization
μ^{st}	Location parameter occ. switching shocks	Occupation Switching Prob., age 55	Normalization

The remaining parameters in the model are calibrated externally, and their values and sources are provided in Table 6. I end this subsection by briefly discussing these choices. 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. The parameters that are most important in determining the behavior of labor supply over the life cycle, such as the discount factor, 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., 2025). In the particular case of λ, I picked the value to account for the overall level of after-tax income, and guarantee an average tax rate of 21%. However, it is worth noting that under the log utility assumption, the amount of labor supplied (and therefore human capital accumulation) is not sensitive to differences in λ. 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, 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. $\tilde{\psi}$ 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. $\tilde{\xi}$ was 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 O*NET. 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 normalization coefficient in the concavity of human capital accumulation function C_j is set to the mean initial human capital in each occupation.

μ^{sw} is set to zero as a normalization, since the relevant object for the decision of staying vs switching is the difference $\mu^{st} - \mu^{sw}$ and μ^{st} is estimated together with α_s to match an occupation switching probability of 0.04 at age 23 and 0.01 at age 55.[13] The scale for the individual occupation shocks conditional on switching α_{occ} is set to 10, to make individuals mostly choose the most preferred occupation.[14] Finally, when picking the parameters of the production function, ζ was set to be 1 so

[13] There are no reliable estimates of two-digit occupation switching rates over a one-year horizon, conditional on employment, so these targets cannot be taken directly from the data. To construct them, I use the fact that job-to-job transitions occur at a rate of approximately 10% per year, and assume that about a quarter of these transitions involve an occupation switch: the average of the two age-specific targets, 0.04 at age 23 and 0.01 at age 55, is 0.025, roughly a quarter of the 10% job-to-job rate. The quantitative results are robust to this assumption. I obtain similar results when I recalibrate the model assuming occupation switches are assumed to account for either 50% more or 50% less of job-to-job transitions.

[14] Setting α_{occ} to a large value also helps guarantee that the nested logit satisfies the regularity condition for

that production is linear and all the convexity in the compensation to labor comes from future wage growth, and TFP was arbitrarily set to 7.25, so that the units of human capital can be interpreted in terms of the Federal Minimum Wage. 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.

Table 6. *Externally calibrated Parameters*

Parameter	Description	Moment / Value	Source
Γ	Occupation human capital transferability	Skill requirements by occupation	O*NET
σ	Relative Risk aversion	1.0	BGP
$\tilde{\xi}$	Mean log learning ability	1.5	Normalization
μ^{occ}	Location parameter, occupation choice	0	Normalization
μ^{sw}	Location parameter, switching shock	0	Normalization
α_{occ}	Scale parameter, occupation choice	10	-
ζ	Returns to scale in production	1.0	-
A	TFP	7.25	Normalization
$1/\eta$	Frisch Elasticity of labor supply	0.3	Bick et al. (2025)
β	Discount factor	0.98	Huggett et al. (2011)
δ_c	Depreciation rate of human capital cognitive occs.	0.025	-
δ_m	Depreciation rate of human capital manual occs.	0.005	-
λ	Overall tax level	1.47	Normalization [15]
τ	Tax progressivity	0.181	Heathcote et al. (2017)
$\rho_{h_0\psi}$	Correl. initial human capital and leisure pref.	-0.3	Bick et al. (2025)
α_{age}	Age decay of learning ability	0.3	-
C_j	Normalization, concavity of $H(\cdot)$	$\bar{h}_j$	ACS
a_0	Initial period	23	-
T_{ret}	Retirement period	65	-
T	Terminal period	83	-

consistency with random utility maximization (RUMS). In this two-stage structure, the Gumbel variance of the inner nest (which occupation to switch to) is $\pi^2/(6\alpha_{occ}^2)$, and the variance of the outer nest (whether to stay or switch) is $\pi^2/(6\alpha_s^2)$. The regularity condition requires $\alpha_s \leq \alpha_{occ}$, i.e., that the inner nest be no more dispersed than the outer nest. By fixing $\alpha_{occ} = 10$ to a large value and estimating α_s to match moderate switching probabilities, this condition is satisfied in the calibration.

4.3 Identification strategy for variance of learning ability (σ_ξ) and leisure preferences (σ_ψ) and correlation between them ($\rho_{\xi\psi}$)

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_ret, 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, variation in the learning ability parameter does not affect labor supply once learning becomes negligible. In the model, this is governed by the age profile of learning, controlled by α_{age}, which determines how quickly learning ability declines over the life cycle. In particular, I calibrate α_{age} so that learning has already become quantitatively irrelevant by age 55. As a result, the variance of hours worked at age 55 is driven almost entirely by heterogeneity in leisure preferences, whereas at younger ages, variation in hours reflects both learning ability and leisure preferences. This feature allows me to separately identify σ_ξ and σ_ψ. Specifically, I identify σ_ψ by matching the variance of hours in the data at age 55, where learning effects are negligible. Given this, σ_ξ is identified by matching the higher dispersion of hours observed among younger workers, in particular at age $a = a_0$. Although both parameters are estimated jointly, this identification strategy exploits the life-cycle profile of learning: if hours were flat over the life cycle, variation would be attributed primarily to leisure preferences, whereas a steep declining profile would imply a larger role for learning ability.

This structure also has implications for the between-occupation variance of hours at age 55, which I use to identify the correlation between learning ability and leisure preferences, $\rho_{\xi\psi}$. At age 55, learning is quantitatively negligible, so differences in hours worked are driven solely by

[15] In this model income is measured in thousands of dollars (mean earnings $\bar{y} = \$31.30$). λ is calibrated to deliver an average tax rate of 21% at mean earnings, consistent with combined federal income and employee payroll taxes. λ is computed with the formula $\lambda = (1-\bar{\tau})\cdot\bar{y}^{\tau}$, where $\bar{\tau} = 0.21$ is the target average rate and $\tau = 0.181$ is the progressivity parameter.

heterogeneity in leisure preferences. In the absence of correlation between ξ_i and ψ_i, the distribution of leisure preferences is identical across occupations, implying no systematic differences in hours across occupations and, therefore, zero between-occupation variance[16]. In contrast, when ξ_i and ψ_i are positively correlated, workers with higher learning ability are also those with lower disutility of labor. Since high-ξ_i individuals sort into high-θ_j occupations earlier in the life cycle, these occupations will be populated by workers with systematically lower ψ_i at later ages. As a result, even when learning no longer affects incentives, average hours will differ across occupations, generating positive between-occupation variance. I therefore identify $\rho_{\xi\psi}$ by matching the between-occupation variance of hours at age 55.

4.4 Targeted Moments Θ_{pre}

In Table 7 I present some summary statistics for the Estimated Parameters Θ_{pre}, which are the parameters that define the initial distribution of human capital, learning ability and disutility of labor across occupations. The table shows the mean and standard deviation of log human capital and log learning ability by occupation, as well as the correlation between log human capital and log learning ability. The full table with the estimated values by occupation is provided in Table 17 in Appendix F.

Table 7. *Initial Distribution Parameters (across occupations)*

Parameter	**Mean**	**Std. Dev.**
Mean log h_0	2.531	0.213
Std log h_0	0.624	0.056
Mean log ξ	2.224	0.200
Std log ξ	0.522	0.044
Corr(h_0, ξ)	0.136	0.161

4.5 Targeted Moments Θ_{struct}

Table 8 presents the estimated values of the structural parameters Θ_{struct}. The occupation-specific parameters θ_j are reported in Table 19 in Appendix H.

[16] With log utility and no learning incentives, the FOC for labor depends only on preferences, so if ψ is uncorrelated with occupation, hours are equalized across occupations.

Table 8. *Estimated Structural Parameters*

Parameter	Description	Estimate
χ	General learning technology	0.010
σ_ξ	Scale for std. dev. of learning ability	2.096
σ_ψ	Std. dev. of labor disutility	0.782
$\rho_{\xi\psi}$	Correl. learning ability and labor disutility	-0.289
$\tilde{\psi}$	Mean log disutility of labor	0.216
α_s	Scale parameter occ. switching shocks	0.499
μ_{st}	Location parameter occ. switching shocks	5.336

4.6 Evaluating Model Fit

In this subsection I report the fit of the model's aggregate moments to the data. Table 9 shows the fit of the scalar targeted moments, including the variance of log hours at ages 23 and 55, and the average experience-wage profile. Table 18 in Appendix G reports the fit of occupation-specific wage growth, which is one of the key moments used to identify the human capital accumulation technology θ_j for each occupation.

Table 9. *Targeted Scalar Moments: Model vs. Data*

Moment	Model	Data	% Diff.
Var. log hours (age 55)	0.0358	0.0343	4.2554
Var. log hours (age 23)	0.0430	0.0457	-5.9906
Avg. EWP	1.3726	1.3875	-1.0745
Between-occ. var. log hours (age 55)	0.0027	0.0027	0.2486

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. The model does a good job at matching the negative slope of the variance of hours profile for all ages.

In Figure 7**(b)** I show the between-occupation variance of hours worked. The value at age 55 is targeted in the estimation; the rest of the profile (ages 23–54) is untargeted. The model is also able to match the decreasing profile of the between-occupation variance of hours over age.

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

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

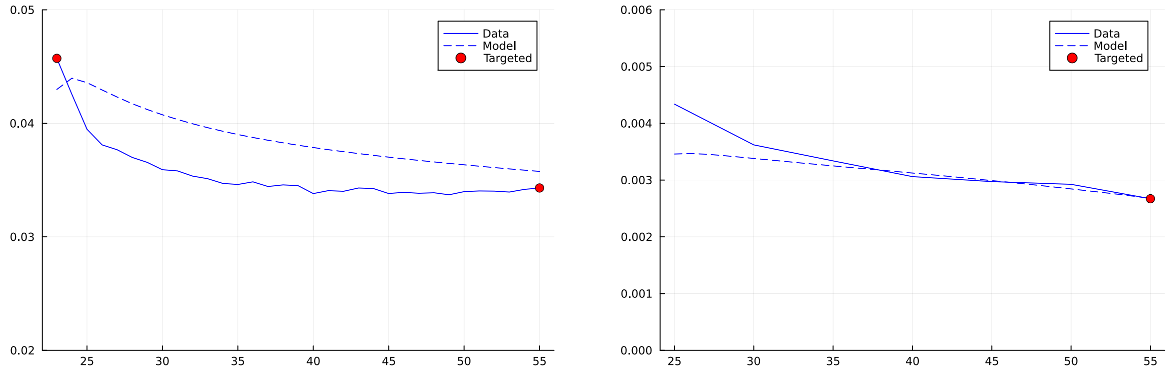


Figure 8**(a)** and 8**(b)** shows the fit of the model to the level of log earnings and log labor by age. The model can match the general shape of log earnings, but predicts higher earnings at young ages than the data. This is due to the fact that in the model the level of log labor at younger ages is higher than it is in the data.

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

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

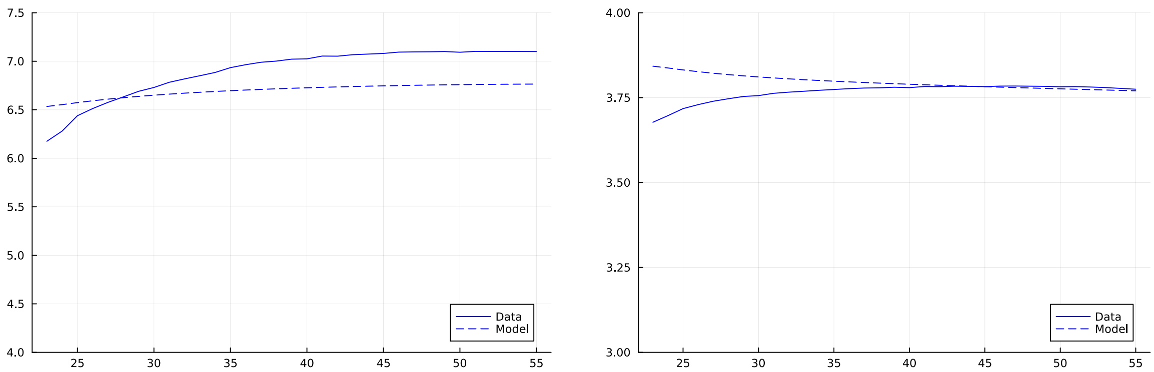
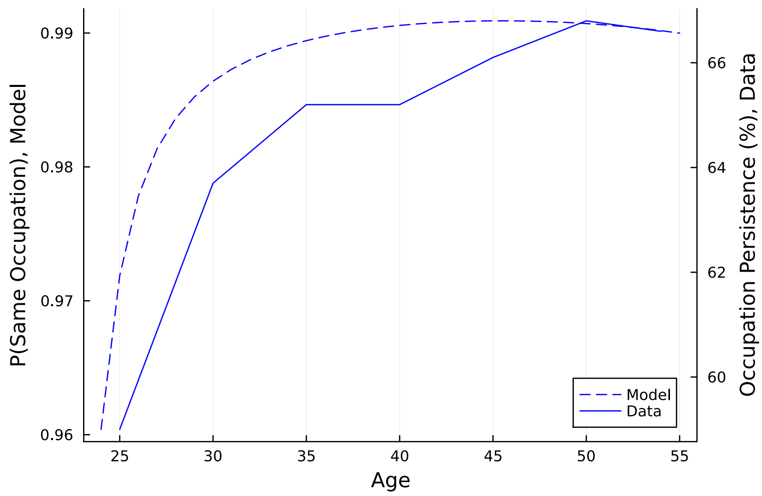


Figure 9 overlays the model-implied occupation persistence (the probability that a worker remains in the same occupation from one period to the next) with the analogous profile in the CPS matched panel (dashed: model, left axis; solid: data, right axis). Only the switching probabilities at ages 24 and 55 are targeted in the estimation, which pin down the level of persistence at the start and at the end of the career. The rise in persistence over the intervening ages, and where it levels off, are not targeted: this life-cycle shape is a fundamental feature of the model, generated by the same mechanisms that discipline the rest of the life-cycle results, rather than an outcome imposed by the calibration. I do not target the *level* of occupation persistence in the data beyond these two anchor ages: self-reported occupation codes are well known to generate spurious switching, so measured CPS persistence understates the true rate at which workers remain in an occupation (Kambourov and Manovskii, 2009a). This is why the two series are plotted on separate axes – the comparison is about shape, not level. What stands out is that the model's untargeted mid-career path: a steep rise in persistence through the late 20s and early 30s, followed by a plateau, closely mirrors the shape of the CPS profile.

Figure 9. *Occupation Persistence over the Life Cycle: Data and Model (Untargeted)*



In Table 10, I present four untargeted moments: the correlation between average workweek and wage growth by occupation, the correlation between average wage growth and learning ability by occupation (proxied in the data by cognitive test scores), the cross-sectional covariance between hours and wages, and the variance of log earnings at age 55. The model matches the sign of the two correlations and of the covariance, but it overstates their magnitude in every case.

This overstatement is expected. Learning ability ξ is a structural parameter of the model: it enters directly, and without error, into both the wage-growth process and the occupational choice margin, so the model's correlation between wage growth and learning ability is measured without noise by construction. In the data, cognitive test scores are only a noisy proxy for learning ability, and occupation codes and workweek are themselves measured with error – self-reported occupation codes are known to generate spurious switching, as discussed above, and workweek is subject to analogous reporting noise. Under classical measurement error, this noise attenuates the corresponding sample correlations and covariance toward zero relative to their true values. The model's untargeted correlations and covariance being uniformly stronger than their empirical counterparts is therefore consistent with, rather than evidence against, the mechanism the model proposes.[17]

Notably, the model also matches the *level* of the variance of log earnings at age 55 fairly closely (0.576 in the model versus 0.665 in the data), even though human capital accumulates (mostly) deterministically and workers face no idiosyncratic human capital shocks as in Huggett et al. (2011) and Bick et al. (2025). Instead, part of this earnings dispersion is generated by the occupation-specific taste shocks ε^{occ} that drive occupation switching over the life cycle. Because switching occupations entails a loss of occupation-specific human capital, these taste shocks expose workers to the risk of human capital losses even though human capital accumulation is otherwise deterministic. In this sense, taste-driven occupational mobility provides a microfoundation for part of the reduced-form risky human capital accumulation process used in Huggett et al. (2011) and Bick et al. (2025).

What the model does not fully capture is the *growth* in earnings dispersion over the life cycle: the variance of log earnings rises by 0.083 from age 23 to age 55 in the model, versus 0.190 in the data, so the model accounts for only about 44% of the observed increase. This shortfall is a direct consequence of ruling out risky human capital accumulation. Because human capital accumulates deterministically here, taste-shock-driven occupation switching is the only channel through which earnings dispersion can grow with age; there is no additional source of earnings risk that accumulates over the life cycle as in Huggett et al. (2011) and Bick et al. (2025). Introducing such a channel would not only mechanically raise the model's earnings-variance growth toward the

[17] As a back-of-the-envelope check, consider the standard attenuation formula for a correlation between two variables each measured with classical (i.i.d., mean-zero) error: $\rho_{data} = \rho_{model}/\sqrt{(1+\lambda_1)(1+\lambda_2)}$, where $\lambda_i \equiv \sigma^2_{e_i}/\text{Var}(x_i^{true})$ is the noise-to-signal variance ratio for variable i and the model's correlation is taken as the noise-free benchmark. Bick et al. (2025) calibrate classical measurement error in log hours with standard deviation $\sigma_{mh} = 0.10$, which they show accounts for 13% of the cross-sectional variance in annual hours – equivalent to a reliability ratio of 0.87, or $\lambda_{hours} \approx 0.15$. Taking this as an external estimate of hours-reporting noise and applying it to the correlation between wage growth and workweek (model $= 0.745$, data $= 0.551$) implies a reliability ratio of about 0.63 ($\lambda_{occ} \approx 0.59$) for the occupation-classification noise diluting wage growth by occupation – a moderate magnitude, consistent with the well-documented unreliability of occupation codes (Kambourov and Manovskii, 2009a). Holding this occupation-classification noise fixed and applying the same logic to the correlation between wage growth and learning ability (model $= 0.878$, data $= 0.479$) implies a reliability ratio of about 0.47 for cognitive test scores as a proxy for learning ability. All three implied magnitudes are plausible relative to the measurement-error literature on self-reported hours and occupation codes, suggesting the gap between model and data in these untargeted moments is quantitatively consistent with known reporting error rather than evidence of a deeper misspecification.

data, it would also be expected to weaken the two correlations discussed above: idiosyncratic human capital shocks that are uncorrelated with the deterministic, occupation-specific returns θ_j driving wage growth in the model would inject noise into both the wage-growth-workweek and wage-growth-learning-ability relationships, pulling the model's currently overstated correlations toward their more attenuated empirical counterparts. Risky human capital accumulation could therefore close part of both gaps documented in this section simultaneously. However, adding a stochastic human capital process to a model that already tracks human capital, occupation, learning ability, and preferences jointly would substantially expand the state space and make the model intractable; quantifying its exact contribution is left as a question for future research.

Table 10. *Untargeted moments for validation*

Statistic	**Model**	**Data**
Corr. Wage Growth and Workweek	0.7449	0.5510
Corr. Wage Growth and Learning Ability	0.8779	0.4786
Cov. Hours and Wages (cross-sectional, ages 23–55)	0.0407	0.0172
Var. log earnings (age 55)	0.5758	0.6650

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 and preferences for leisure. Using Shapley values to account for interaction effects, my findings indicate that in this model with occupational choice, differences in leisure preferences account for 80% of the variance in log hours at age 23, while learning ability explains only 8%. However, when we look at earnings at age 55, the picture reverses: learning ability accounts for 54% of the variance in log earnings, while leisure preferences explain only 28%.

Table 11. *Decomposing the variance of hours worked and earnings*

		Benchmark	$\sigma_\psi = 0$	$\sigma_\xi = 0$	$\sigma_\psi = \sigma_\xi = 0$	$S_\psi\%$	$S_\xi\%$
	$V(\ln(\text{hours}))$	0.044	0.007	0.039	0.005		
	% Bench / Shapley	100	16	88	12	80	8
	$V(\ln(\text{earnings}))$	0.376	0.29	0.355	0.162		
age = 23	% Bench / Shapley	100	77	94	43	37	20
	$V_B(\ln(\text{hours}))$	0.003	0.004	0.002	0.004		
	% Bench / Shapley	100	125	77	120	-34	14
	$V(\ln(\text{hours}))$	0.036	0.001	0.034	0.0		
	% Bench / Shapley	100	2	94	1	96	3
	$V(\ln(\text{earnings}))$	0.51	0.397	0.264	0.09		
age = 55	% Bench / Shapley	100	78	52	18	28	54
	$V_B(\ln(\text{hours}))$	0.003	0.001	0.001	0.001		
	% Bench / Shapley	100	24	31	20	43	37
	$V(\overline{\ln(\text{hours})})$	0.038	0.002	0.035	0.001		
	% Bench / Shapley	100	4	91	2	93	5
	$V(\overline{\ln(\text{earnings})})$	0.436	0.313	0.276	0.091		
Life cycle	% Bench / Shapley	100	72	63	21	35	44
	$\bar{V}_B(\ln(\text{hours}))$	0.003	0.001	0.001	0.001		
	% Bench / Shapley	100	47	41	40	27	33

Note: % Bench = value as % of benchmark. $S_\psi\%$, $S_\xi\%$ = Shapley shares (% of benchmark variance).

If we look at Table 11, we can see that in the absence of both learning ability and leisure preference heterogeneity, the variance of log hours is only 12% of the benchmark model at age 23 and 1% at age 55. This confirms that the vast majority of 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. The variance of earnings when there is no learning or leisure heterogeneity is still 43% of the benchmark at age 23, driven mechanically by the initial variance of wages estimated from the data. However, by age 55 this falls to only 18% of the benchmark, as the lack of heterogeneity prevents the accumulation of earnings differences over the life cycle.

Allowing for learning ability dispersion alone (setting $\sigma_\psi = 0$), the variance of log hours at age 23 remains at only 16% of the benchmark, but the variance of log earnings at age 55 reaches 78% of the benchmark. This suggests that even though learning ability explains only a small fraction of hours variance when workers are young, its interaction with human capital accumulation generates substantial earnings dispersion over the life cycle.

On the other hand, allowing for leisure preference dispersion alone (setting $\sigma_\xi = 0$), the variance of log hours at age 23 reaches 88% of the benchmark, while the variance of log earnings at age 55 is only 52% of the benchmark. The Shapley decomposition confirms this asymmetry: leisure preferences account for 80% of hours variance at age 23 but only 28% of earnings variance at age 55, while learning ability accounts for only 8% of hours variance at age 23 but 54% of earnings variance at age 55. This implies that learning ability's impact on earnings inequality at age 55 is roughly twice as large as the role of leisure preference heterogeneity.

Interestingly, the Shapley share for leisure preferences on the between-occupation variance of log hours is negative at age 23 (-34%), meaning preference heterogeneity actually *narrows* rather than widens between-occupation hours dispersion at labor market entry.[18]

The role of preference heterogeneity in between-occupation variance changes dramatically over the life cycle. While at age 23 preferences have a negative Shapley share (-34%), by age 55 this reverses to a positive contribution of 43%. This shift reflects evolving sorting patterns: early in the career,

[18] This dampening effect is evident from the counterfactual: removing preference heterogeneity actually *increases* the between-occupation variance of hours to 125% of the benchmark level. Because occupational choice is stochastic (nested logit) rather than a deterministic sort, every occupation draws from a similar mix of leisure-preference types rather than fully segregating by ψ, which pulls each occupation's average hours toward a common population-wide level regardless of its human capital return θ_j. At age 23, before between-occupation human capital differences have had time to compound, this compositional homogenizing effect is large relative to the still-small systematic signal from θ_j-driven sorting, so removing preference heterogeneity (setting $\sigma_\psi = 0$) leaves only the systematic, skill-driven component, and between-occupation dispersion widens. By age 55, decades of compounding have made the θ_j-driven divergence in human capital and hours far larger in absolute terms, so preference heterogeneity's homogenizing role becomes comparatively small relative to the total variance, and the usual direction reasserts itself: removing it now *reduces* the between-occupation variance of hours, consistent with the positive Shapley share of 43% at age 55 discussed below.

workers with heterogeneous preferences spread across occupations as they explore and switch, which smooths between-occupation differences. As workers age and settle into occupations, those with strong leisure preferences concentrate in occupations that accommodate their preferences, while those with weaker leisure preferences concentrate elsewhere. This increased sorting by preferences generates larger between-occupation hours differences later in life. Learning ability, in contrast, already plays a modest positive role in between-occupation hours variance at age 23 (14%), rising further to 37% at age 55. The benchmark between-occupation variance of hours remains close to 0.003 at both ages, in contrast to the sharp decline in the overall cross-sectional variance of hours documented above, indicating that between-occupation dispersion does not shrink appreciably even as within-occupation hours choices converge over the life cycle.

The lifetime measures in Table 11 summarize the overall contribution of each source of heterogeneity across the entire working life. Leisure preferences account for 93% of the lifetime variance of log hours, while learning ability accounts for only 5%. However, for lifetime earnings variance, learning ability dominates with a 44% Shapley share compared to 35% for leisure preferences. The average between-occupation variance of hours over the life cycle shows Shapley shares of 27% for preferences and 33% for learning ability, roughly comparable in magnitude with learning ability slightly ahead. These lifetime results reinforce the main finding: while leisure preferences are the primary driver of hours dispersion throughout the life cycle, learning ability heterogeneity—through its compounding effect on human capital accumulation—becomes the dominant source of earnings inequality over time.

5.1 Contribution of Hours to earnings inequality

To isolate the contribution of hours variation to earnings inequality, I conduct counterfactual experiments where hours are fixed rather than chosen optimally. Table 12 reports the results. In the first counterfactual, hours are fixed at the overall mean ($\ell = 1$) for all workers in all occupations. In the second, hours are fixed at occupation-specific means, allowing for between-occupation differences in average hours but eliminating within-occupation variation.

Table 12. *Earnings variance with fixed hours*

	Benchmark	$\ell = 1$	$\ell = \bar{\ell}_j$
Age 23			
$V(\ln(\text{earnings}))$	0.3761	0.1886	0.1968
% of Benchmark	100	50.1	52.3
Age 55			
$V(\ln(\text{earnings}))$	0.5096	0.2198	0.2515
% of Benchmark	100	43.1	49.4
Lifetime			
$V(\overline{\ln(\text{earnings})})$	0.4356	0.1853	0.2048
% of Benchmark	100	42.5	47.0

Note: $\ell = 1$ fixes hours at mean. $\ell = \bar{\ell}_j$ fixes hours at occupation averages.

When hours are fixed at the overall mean, only 43% of the lifetime variance of log earnings remains, indicating that hours variation—both the level chosen and its interaction with human capital accumulation—accounts for 57% of lifetime earnings inequality. Fixing hours at occupation-specific means retains slightly more variance (47%), as between-occupation differences in average hours are preserved. The difference between these two counterfactuals (4 percentage points) captures the contribution of between-occupation hours differences to earnings inequality, while the remaining 53% attributed to hours comes from within-occupation hours variation and its dynamic effects on human capital.

These results reveal that within-occupation hours variation is far more important than between-occupation hours variation in explaining lifetime earnings inequality. Specifically, within-occupation hours variation accounts for approximately 53% of lifetime earnings variance, while between-occupation hours differences explain only about 4.5%. This finding is consistent with the narrow range of occupation-specific mean hours in the data (approximately $\pm 10\%$ from the overall mean), which limits the scope for between-occupation hours differences to drive earnings inequality. In contrast, individual hours choices within occupations vary substantially based on workers' heterogeneous preferences and learning abilities, and these choices compound over time through their effect on human capital accumulation.

The contribution of hours to lifetime earnings inequality in this model is substantially larger than in previous work. Bick et al. (2025) conduct a similar counterfactual exercise and find that fixing hours at the mean leaves 82% of lifetime earnings variance, implying that hours variation explains only 18% of lifetime earnings inequality. In contrast, my model implies that hours variation accounts for 57% of lifetime earnings inequality—more than three times the contribution found in their framework. However, this comparison requires careful interpretation. The model in Bick et al. (2025) includes stochastic shocks to human capital accumulation, which generate additional earnings variation over the life cycle beyond what is determined by initial conditions and hours choices. My model abstracts from such shocks, making it analogous to the initial conditions component of their framework. As a result, the denominator (total lifetime earnings variance) is smaller in my model, mechanically increasing the share attributed to hours.

A back-of-the-envelope calculation helps put these results in perspective. In Bick et al. (2025), shutting down initial conditions heterogeneity leaves approximately 49% of lifetime earnings variance, implying that initial conditions account for about 51% of total variance. If my model captures the initial conditions component, scaling my estimate to a framework with shocks yields an implied hours contribution of $57\% \times 51\% \approx 29\%$ of total lifetime earnings variance, still substantially larger than the 18% found in Bick et al. (2025). Alternatively, if hours explain 18% of total variance in their framework and initial conditions account for 51%, then hours explain approximately $18\%/51\% \approx 35\%$ of the variance from initial conditions, compared to 57% in my model. This difference likely reflects the learning-by-doing mechanism: in my framework, hours choices directly affect human capital accumulation, amplifying their contribution to lifetime earnings inequality through a compounding effect over the life cycle.

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, assuming all other factors remain constant. 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 exclusively by leisure preferences, or part of it is caused by learning ability. In other words, I 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 learning ability heterogeneity. To achieve revenue equivalence of the tax systems I increase progressivity parameter τ and then adjust λ 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 is the only parameter of the tax system that affects labor supply. For the results for the base calibration, see Table 13. For the results in the calibration with no learning ability heterogeneity, see Table 14. Appendix J describes in detail how lifetime welfare and lifetime earnings inequality are computed.

As τ increases, λ rises in both calibrations in order to maintain revenue equivalence, over the whole range of progressivity considered ($\tau = 0.181$ to $\tau = 0.231$; see Tables 13 and 14). This reflects the fact that a more progressive tax schedule shrinks the effective tax base $\sum_i y_i^{1-\tau}$ by proportionally more than it shrinks aggregate pre-tax earnings $\sum_i y_i$: raising τ makes $y^{1-\tau}$ a more concave function of earnings, so with substantial earnings dispersion present in both calibrations, even a modest decline in aggregate output translates into a considerably larger decline in the effective tax base, requiring λ to rise in order to keep total revenue unchanged. This effect is of a similar magnitude in both calibrations: λ rises from 1.470 to 1.707 in the calibration with learning ability heterogeneity, and from 1.470 to 1.707 in the calibration without. In fact, λ is essentially identical between the two calibrations at every point on the grid, not just at the endpoints (1.470, 1.515, 1.609, and 1.707, respectively, in both), so the two calibrations differ little along this dimension. As discussed below, they instead differ mainly in how much this change in the tax system translates into reduced labor supply, and consequently into differences in inequality and welfare.

In both calibrations, increasing progressivity reduces incentives to work, particularly when people are older and have higher earnings. The resulting declines in output and average wage growth are of similar magnitude in both calibrations: output falls by about 2.1 percentage points in each, and average wage growth falls by about 1.1 to 1.2 percentage points, with the model without learning ability heterogeneity showing, if anything, a marginally larger decline in wage growth. The two calibrations therefore differ little in the aggregate labor-supply response to progressivity; as discussed below, they instead differ mainly in how this response translates into changes in earnings inequality and welfare.

Table 13. *Revenue-neutral increase in tax progressivity*

τ	λ	**Earnings Ineq.**	**Ouput**	**Welfare**	**Wage Growth**
0.181	1.47	100.0	100.0	100.0	100.0
0.191	1.515	99.3	99.6	97.2	99.8
0.211	1.609	97.9	98.8	91.9	99.4
0.231	1.707	96.7	97.9	86.9	98.9

Table 14. *Revenue-neutral increase in tax progressivity, no learning ability heterogeneity*

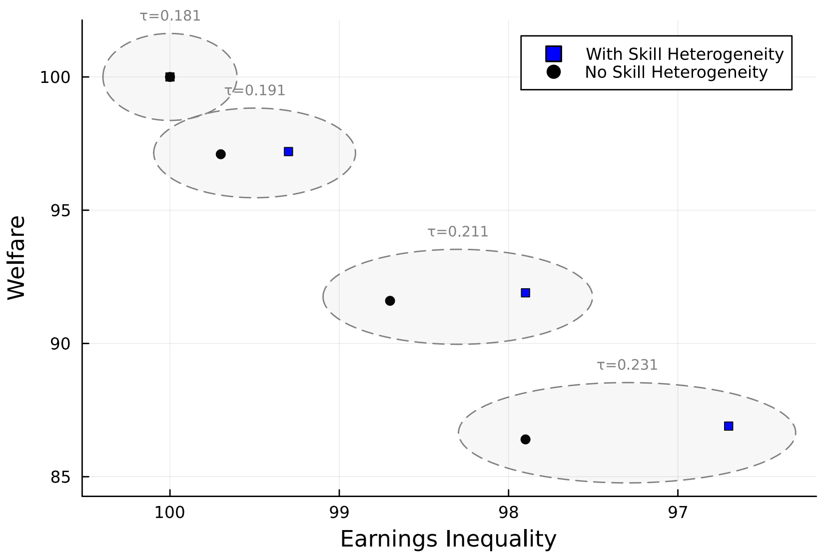
τ	λ	**Earnings Ineq.**	**Ouput**	**Welfare**	**Wage Growth**
0.181	1.47	100.0	100.0	100.0	100.0
0.191	1.515	99.7	99.6	97.1	99.8
0.211	1.609	98.7	98.7	91.6	99.3
0.231	1.707	97.9	97.9	86.4	98.8

In Figure 10, I illustrate the differences in the lifetime welfare cost of reducing inequality. Increasing tax progressivity in the model with learning ability heterogeneity is more effective at reducing lifetime earnings inequality than in the model without learning ability heterogeneity. This is because high-ability workers are more responsive to changes in progressivity and therefore adjust their labor supply by more, leading to a larger reduction in the growth of earnings inequality over the life cycle. Over the full range of progressivity considered ($\tau = 0.181$ to $\tau = 0.231$), lifetime earnings inequality falls by 3.3 percentage points in the model with learning ability heterogeneity, compared with only 2.1 percentage points in the model without learning ability heterogeneity.

Lifetime welfare, on the other hand, falls slightly more in the model without learning ability heterogeneity. When all hours dispersion is driven by heterogeneity in leisure preferences, differences in earnings largely reflect differences in workers' tastes. As a result, earnings differences overstate the underlying differences in welfare between low- and high-earning workers, making redistribution less valuable from a welfare perspective. In contrast, when heterogeneity in learning ability also contributes to earnings dispersion, greater tax progressivity provides some insurance against being born with low learning ability. Consequently, a given welfare cost generates a larger reduction in lifetime earnings inequality when workers differ in learning ability.

Combining these two effects, the welfare cost of reducing lifetime earnings inequality by one percentage point is substantially higher in the model without learning ability heterogeneity. The cost is approximately 4.0% of lifetime consumption per percentage point of inequality reduction in the model with learning ability heterogeneity, compared with approximately 6.5% in the model without learning ability heterogeneity.

Figure 10. *Equality-Efficiency Trade-off*



Notes: Each point in the graph corresponds to a different level of progressivity τ for each model. The values of progressivity tested in both models are: 0.181, 0.191, 0.211, 0.231. 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, as well as the decrease in occupation switching over the life cycle, without relying on age-specific structural parameters, providing a robust framework for understanding these dynamics.

Through the calibration and estimation of the model, I decompose the variance of hours worked and earnings into the portions attributable to learning ability and to leisure preferences. The decomposition reveals a sharp contrast: leisure preferences account for the vast majority of hours dispersion early in life (80%, versus only 8% for learning ability at age 23), yet it is learning ability that becomes the dominant force behind earnings inequality later in life, once amplified through human capital accumulation (54%, versus 28% for leisure preferences at age 55). More broadly, hours dispersion alone accounts for 57% of lifetime earnings inequality once its effect on future wage growth is taken into account, despite its modest direct, contemporaneous contribution to earnings. These findings highlight that understanding the source of hours dispersion is important for evaluating policies aimed at reducing lifetime earnings inequality.

The policy experiment conducted in this paper provides important insights into how progressive taxation impacts inequality and other economic aggregates. In particular, I show that increasing tax progressivity is more effective at reducing lifetime earnings inequality, and less costly in terms of lifetime welfare per percentage point of inequality reduced, in a world in which learning ability is an important source of hours dispersion than in a world in which all hours dispersion is due to leisure preferences: the welfare cost of reducing lifetime earnings inequality by one percentage point is approximately 4.0% of lifetime consumption in the former, compared with 6.5% in the latter. This difference arises even though the two calibrations imply similar aggregate responses of output and wage growth to progressivity. This suggests that the effectiveness of tax policies aimed at reducing inequality depends on the underlying sources of hours dispersion, rather than on how responsive aggregate output is to taxation.

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 and occupational sorting as key drivers 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 examining how they interact with occupational choice and learning ability heterogeneity.

More broadly, these results show that understanding the consequences of hours inequality requires taking a stand on why people work the hours they do. Hours differences that look similar in the cross-section can have very different implications for lifetime earnings inequality and welfare depending on whether they reflect differences in leisure preferences or differences in learning ability. In the context of tax policy, this implies that policies with similar effects on aggregate output can have substantially different distributional and welfare consequences.

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(3).

Bick, A., Blandin, A., and Rogerson, R. (2025). Hours Worked and Lifetime Earnings Inequality. *NBER Working Paper No. 32997.*

Blandin, A., Jones, J. B., and Yang, F. (2025). Marriage and Work among Prime-Age Men. *Federal Reserve Bank of Richmond Working Papers No. 23-02R.*

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.

Checchi, D., García-Peñalosa, C., and Vivian, L. (2022). Hours inequality. *CESifo Working Paper No. 10128.*

Erosa, A., Fuster, L., Kambourov, G., and Rogerson, R. (2022). Labor Supply and Occupational Choice. *NBER Working Paper No. 30492.*

Gethin, A. and Saez, E. (2026). Global Working Hours. *The Quarterly Journal of Economics*, page qjag030.

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. In Phelps, E. S., editor, *Designing Inclusion: Tools to Raise Low-end Pay and Employment in Private Enterprise*, 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.

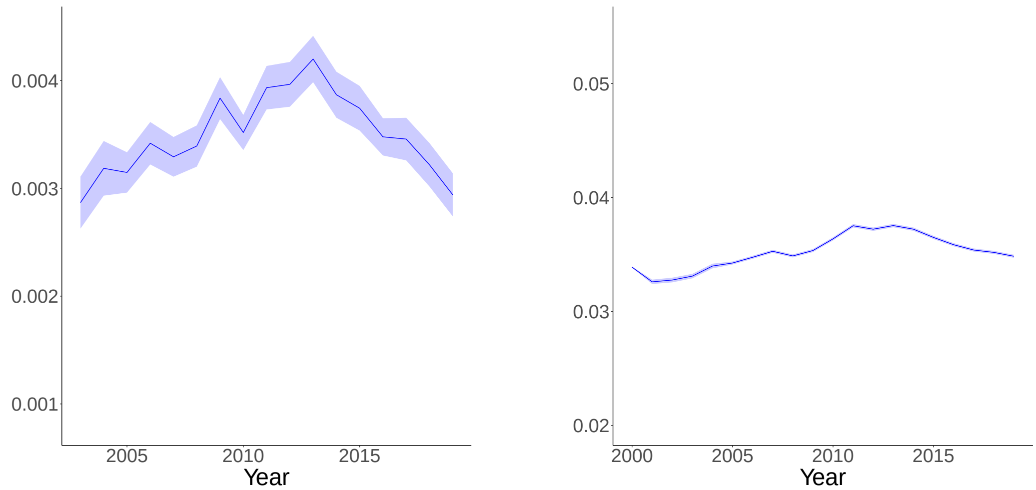
Young, E. R. (2010). Solving the incomplete markets model with aggregate uncertainty using the Krusell-Smith algorithm and non-stochastic simulations. *Journal of Economic Dynamics and Control*, 34(1):36–41.

A Appendix: Variance of Log hours worked by year

This appendix reports the two motivating cross-sectional facts from Figure 1 by calendar year rather than by age, to show that the patterns documented in the main text are not primarily driven by year effects. Figure 11(a) shows that the between-occupation variance of log hours has not declined over time. If anything, its evolution over calendar years runs somewhat opposite to the declining age profile documented in Figure 1, while its level remains relatively stable around 0.0035. Figure 11(b) similarly shows that the cross-sectional variance of log hours has remained broadly stable over the 2003–2023 period. These patterns suggest that the decline in hours dispersion with age observed in the ACS is not primarily driven by year effects. This interpretation is further supported by the fact that a similar declining age profile has been documented using true panel data from the NLSY79 and NLSY93.

Figure 11. *Variance of Log Hours Worked Over the Life Cycle*

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



Notes: Author's calculations from the American Community Survey (ACS), 2003-2023. Shaded areas represent 95% confidence intervals.

B Appendix: ACS descriptive statistics

Table 15 reports, for each of the 22 occupation categories used in the estimation, the average log wage, the mean and standard deviation of weekly hours worked, and the experience-wage profile

(EWP) estimate, all computed from the ACS sample.

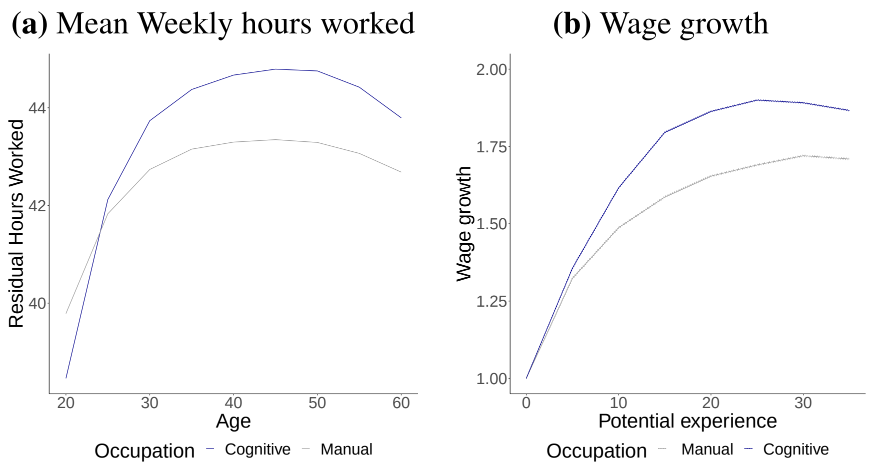
Table 15. *ACS descriptive statistics and EWP estimates*

	Occupation	log(wage)	Weekly Hours	std(hours)	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

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

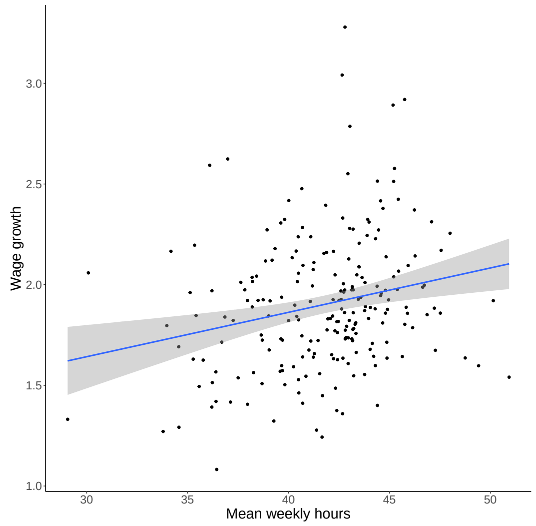
To check that the positive relationship between hours and wage growth documented in the main text is not an artifact of the occupation classification used there, I conduct two robustness checks using alternative levels of aggregation. First, I consider a much coarser classification, grouping occupations into cognitive and manual/routine categories. The results are reported in Figure 12. The positive relationship between hours and wage growth is preserved under this coarser classification: cognitive occupations exhibit both higher average hours and higher wage growth than manual/routine occupations.

Figure 12. *Employment outcomes over the life cycle by Occupation type*



Second, I consider a finer classification by computing the relationship between hours and wage growth at the occupation-by-industry level rather than by occupation alone. I do not use 3-digit occupations for this exercise because the resulting smaller cell sizes lead to a large number of age-occupation cells being excluded when computing wage growth. Figure 13 reports the correlation between EWP and hours worked across occupation-industry pairs. The positive relationship remains, confirming that the result is robust to using a more disaggregated classification.

Figure 13. *Correlation EWP and hours worked by Occupation-Industry pairs*



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 O*NET portal. O*NET gives a score in each of the 25 skills presented in Table 16 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 16. *List of O*NET 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 O*NET 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 O*NET. 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:

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

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

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

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

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

The parameters Θ_{pre} above are pinned down analytically from the exact mapping just described. The structural parameters Θ_{struct}, by contrast, are estimated by simulated method of moments (SMM), following the two-stage procedure described in Section 4. Letting $\mathcal{G}_{struct}(\hat{\Theta}_{pre}, \Theta_{struct})$ and $\bar{\mathcal{G}}_{struct}$

denote the vectors of model-implied and data moments, $\hat{\Theta}_{struct}$ minimizes the quadratic form

$$\left[\mathcal{G}_{struct}(\hat{\Theta}_{pre}, \Theta_{struct}) - \bar{\mathcal{G}}_{struct}\right]' W_{struct} \left[\mathcal{G}_{struct}(\hat{\Theta}_{pre}, \Theta_{struct}) - \bar{\mathcal{G}}_{struct}\right],$$

where W_{struct} is a diagonal weighting matrix with i-th diagonal entry $1/(\bar{\mathcal{G}}_{struct,i})^2$. This choice of W_{struct} makes the objective a sum of squared *percentage* deviations,

$$\sum_i \left(\frac{\mathcal{G}_{struct,i}(\hat{\Theta}_{pre}, \Theta_{struct}) - \bar{\mathcal{G}}_{struct,i}}{\bar{\mathcal{G}}_{struct,i}}\right)^2,$$

which places moments of very different scales and units – experience-wage profile (EWP) ratios around 1–2, variances around 0.03–0.05, occupation-switching probabilities around 0.01–0.04 – on comparable footing, rather than letting the largest-magnitude moments mechanically dominate the objective. Because the EWP by occupation contributes 22 individual moments, far more than any other target, I additionally downweight these 22 moments by a factor of 0.5 relative to the remaining scalar moments (the variance of log hours at ages 23 and 55, the average EWP, the between-occupation variance of log hours, average hours, and the two occupation-switching probabilities at ages 24 and 55), so that the cross-occupation wage growth pattern does not dominate the objective simply by virtue of being measured with more moments than the aggregate life-cycle targets.

The estimation combines a global optimization search with a subsequent local (Nelder-Mead) refinement, reaching a final loss of 0.03.

F Appendix: Estimation Results for Θ_{pre}

Table 17 reports the full set of estimated parameters Θ_{pre} summarized in Table 7 in the main text: the mean and standard deviation of initial log human capital, and of log learning ability, for each occupation.

Table 17. *Mean and standard deviation of initial log human capital and log learning ability by occupation*

Occupation	**Mean** $\log h_0$	**Std** $\log h_0$	**Mean** $\log \xi$	**Std** $\log \xi$	**Corr**(h_0, ξ)
Architecture and Engineering	2.964	0.638	2.411	0.496	0.422
Arts, Design, Entertainment, Sports, and Media	2.559	0.757	2.451	0.515	0.181
Building and Grounds Cleaning and Maintenance	2.263	0.612	2.110	0.572	0.156
Business Operations Specialists	2.781	0.618	2.495	0.572	0.224
Computer and Mathematical	2.973	0.673	2.562	0.493	0.162
Construction and Extraction	2.545	0.611	2.044	0.449	0.153
Counselors, Social and Religious Workers	2.482	0.601	2.273	0.558	-0.069
Education, Training, and Library	2.495	0.652	2.537	0.500	0.065
Farming, Fishing, and Forestry	2.239	0.603	2.243	0.407	-0.032
Food Preparation and Serving	2.216	0.578	2.164	0.521	-0.105
Healthcare Practitioners and Technical	2.633	0.593	2.426	0.596	0.068
Healthcare Support	2.382	0.568	1.947	0.508	0.194
Installation, Maintenance, and Repair	2.582	0.562	2.047	0.501	0.163
Legal	2.793	0.772	2.696	0.548	0.484
Life, Physical and Social Science	2.656	0.643	2.472	0.554	-0.146
Management, Business, Science and Arts	2.642	0.638	2.299	0.558	0.165
Office and Administrative Support	2.415	0.576	2.203	0.559	0.210
Personal Care and Service	2.236	0.692	2.137	0.507	-0.023
Production	2.499	0.575	2.069	0.516	0.208
Protective Services	2.538	0.603	2.014	0.477	0.265
Sales and Related	2.417	0.617	2.244	0.554	0.145
Transportation and Material Moving	2.368	0.599	2.032	0.493	0.078

G Appendix: Moment Fit for wage growth by occupation

Table 18 reports the full occupation-by-occupation comparison of the model's fitted experience-wage profile (EWP) against its data counterpart, summarized by the scalar moments in Table 9 in the main text.

Table 18. *Estimated Wage Growth by occupation*

Occupation	**Model**	**Data**	**% Diff.**
Architecture and Engineering Occupations	1.36	1.41	-3.45
Arts, Design, Entretainment, Sports, and Media Occupations	1.45	1.45	0.07
Building and Grounds Cleaning and Maintenance Occupations	1.28	1.26	1.89
Business Operations Specialists	1.38	1.38	-0.08
Computer and Mathematical Occupations	1.40	1.36	2.85
Construction and Extraction Occupations	1.25	1.38	-9.40
Counselors, Social and Religious Workers	1.19	1.18	0.09
Education, Training, and Library Occupations	1.59	1.59	-0.08
Farming, Fishing, and Forestry Occupations	1.25	1.23	1.16
Food preparation and Serving Occupations	1.34	1.16	15.89
Healthcare Practitioners and Technical Occupations	1.89	1.89	-0.28
Healthcare Support Occupations	1.20	1.19	0.53
Installation, Maintenance, and Repair Workers	1.39	1.37	0.91
Legal Occupations	1.67	1.68	-0.77
Life, Physical and Social Science Occupations	1.71	1.71	-0.38
Management, Business, Science and Arts Occupations	1.56	1.56	-0.04
Office and Administrative Support Occupations	1.43	1.43	0.26
Personal Care and Service Occupations	1.24	1.23	0.50
Production Occupations	1.26	1.39	-9.03
Protective Services Occupations	1.18	1.25	-5.63
Sales and Related Occupations	1.42	1.41	0.18
Transportation and Material Moving Occupations	1.37	1.34	2.24

H Appendix: Estimated Human Capital Accumulation Technology by Occupation

Table 19 reports the full set of estimated occupation-specific human capital accumulation parameters θ_j, summarized in Table 8 in the main text.

Table 19. *Estimated Human Capital Accumulation Technology θ_j by occupation*

Occupation	θ_j
Architecture and Engineering Occupations	0.790
Arts, Design, Entretainment, Sports, and Media Occupations	0.770
Building and Grounds Cleaning and Maintenance Occupations	0.300
Business Operations Specialists	0.774
Computer and Mathematical Occupations	0.793
Construction and Extraction Occupations	0.300
Counselors, Social and Religious Workers	0.698
Education, Training, and Library Occupations	0.825
Farming, Fishing, and Forestry Occupations	0.525
Food preparation and Serving Occupations	0.300
Healthcare Practitioners and Technical Occupations	0.840
Healthcare Support Occupations	0.758
Installation, Maintenance, and Repair Workers	0.875
Legal Occupations	0.757
Life, Physical and Social Science Occupations	0.831
Management, Business, Science and Arts Occupations	0.837
Office and Administrative Support Occupations	0.798
Personal Care and Service Occupations	0.706
Production Occupations	0.300
Protective Services Occupations	0.300
Sales and Related Occupations	0.791
Transportation and Material Moving Occupations	0.783

I Appendix: Model Fit for the Version with No Learning Ability Heterogeneity

The policy experiment in Section 6 compares the baseline calibration to a version of the model in which learning ability heterogeneity is shut down entirely ($\sigma_\xi = 0$), so that all dispersion in hours worked is attributable to differences in leisure preferences (see Table 14). This version is re-estimated separately from the baseline: since $\sigma_\xi = 0$ removes the identification strategy described in Section 4 that separately pins down σ_ξ and σ_ψ from the variance of hours at ages 23 and 55, the variance of hours is instead targeted as a single average over the full life cycle, and the correlation parameter $\rho_{\xi\psi}$ is dropped, as it is not defined when $\sigma_\xi = 0$.

Tables 20 and 21 report the fit of this version to the same targeted moments used in the baseline estimation. The estimation combines a global optimization search with a subsequent local (Nelder-Mead) refinement, reaching a final loss of 0.079, compared to 0.030 for the baseline calibration.

Table 20. *Estimated Wage Growth by occupation, version with no learning ability heterogeneity*

Occupation	**Model**	**Data**	**% Diff.**
Architecture and Engineering Occupations	1.32	1.41	-6.65
Arts, Design, Entretainment, Sports, and Media Occupations	1.37	1.45	-5.42
Building and Grounds Cleaning and Maintenance Occupations	1.34	1.26	6.67
Business Operations Specialists	1.35	1.38	-2.15
Computer and Mathematical Occupations	1.41	1.36	4.20
Construction and Extraction Occupations	1.42	1.38	3.55
Counselors, Social and Religious Workers	1.09	1.18	-8.33
Education, Training, and Library Occupations	1.55	1.59	-2.41
Farming, Fishing, and Forestry Occupations	1.37	1.23	11.27
Food preparation and Serving Occupations	1.36	1.16	17.70
Healthcare Practitioners and Technical Occupations	1.65	1.89	-13.02
Healthcare Support Occupations	1.32	1.19	10.88
Installation, Maintenance, and Repair Workers	1.29	1.37	-5.90
Legal Occupations	1.57	1.68	-6.31
Life, Physical and Social Science Occupations	1.60	1.71	-6.75
Management, Business, Science and Arts Occupations	1.60	1.56	3.09
Office and Administrative Support Occupations	1.35	1.43	-5.22
Personal Care and Service Occupations	1.39	1.23	12.78
Production Occupations	1.34	1.39	-3.59
Protective Services Occupations	1.33	1.25	6.18
Sales and Related Occupations	1.34	1.41	-5.47
Transportation and Material Moving Occupations	1.31	1.34	-2.72

Table 21. *Targeted Scalar Moments, version with no learning ability heterogeneity*

Moment	Model	Data	% Diff.
Avg. V(log hours), life cycle	0.0367	0.0355	3.4493
Avg. EWP	1.4115	1.3875	1.7286
Between-occ. var. log hours (age 55)	0.0026	0.0027	-3.5432
Avg. hours	1.0707	1.0000	7.0697
P(Switch, age 24)	0.0379	0.0400	-5.2328
P(Switch, age 55)	0.0100	0.0100	0.4989

J Appendix: Computing Welfare in the Policy Experiment

This appendix details how the lifetime welfare and lifetime earnings inequality statistics reported in Tables 13 and 14 and in Figure 10 are computed.

Lifetime welfare. For each combination of tax parameters $\boldsymbol{\tau} = (\tau, \lambda)$, the model solution delivers a value function $V(h, j, \xi, \psi, a_0; \boldsymbol{\tau})$ for every initial type (h, j, ξ, ψ), which is by construction the exact expected present discounted utility of an individual of that type over the remainder of the life cycle, integrating over all future realizations of occupational switching shocks and human capital accumulation. Since agents have log utility and no ability to save or borrow, a proportional increase in consumption in every period raises lifetime utility by an amount that does not depend on the path of consumption itself. This makes it possible to translate a change in V into a change in lifetime consumption-equivalent units. Specifically, for two tax systems $\boldsymbol{\tau}^0$ (the benchmark) and $\boldsymbol{\tau}^1$ (the reform), define, for every type (h, j, ξ, ψ),

$$\frac{\bar{c}^1(h, j, \xi, \psi, a_0)}{\bar{c}^0(h, j, \xi, \psi, a_0)} = \exp\left(\frac{V^1(h, j, \xi, \psi, a_0) - V^0(h, j, \xi, \psi, a_0)}{\sum_{a=a_0}^{T} \beta^{a-a_0}}\right),$$

which gives, for that type, the constant proportional change in consumption in every period that would leave lifetime utility unchanged under the benchmark tax system. The aggregate welfare change reported in the tables and in Figure 10 is then the population average of this type-specific consumption-equivalent change, computed using the initial distribution of types $F(h, j, \xi, \psi)$ at age

a_0,

$$\Delta W = \int \left[\frac{\bar{c}^1(h, j, \xi, \psi, a_0)}{\bar{c}^0(h, j, \xi, \psi, a_0)} - 1 \right] dF(h, j, \xi, \psi).$$

Since V is already the exact expected lifetime utility for each type, no simulation is required to compute ΔW: the value function returned by the model solution is used directly.

Lifetime earnings inequality. Unlike welfare, lifetime earnings inequality is a property of the entire distribution of *realized* life-cycle earnings paths, not an expectation, and cannot be recovered from the value function alone. It is instead computed by Monte Carlo simulation: a large number of individuals are drawn from the initial distribution of types at age a_0, and each is simulated forward through age T using the model's occupation- and human-capital-transition probabilities, accumulating post-tax earnings $\lambda y_{i,a}^{1-\tau}$ in every period along the simulated path. The reported statistic is the variance, across simulated individuals, of each individual's own average log post-tax earnings over their simulated life cycle.