Essays on Technological Change and Inequality

by

Siyu Shi

A Dissertation Presented in Partial Fulfillment
of the Requirements for the Degree
Doctor of Philosophy

Approved March 2023 by the
Graduate Supervisory Committee:

Gustavo Ventura, Chair
Alexander Bick
Natalia Kovrijnykh

ARIZONA STATE UNIVERSITY

May 2023

ABSTRACT

This collection of essays attempts to address the question: how does recent technological progress shape inequality in the labor market? In the first chapter, I document and investigate life-cycle profiles of skill premiums across cohorts. My empirical analysis shows that younger cohorts have steeper growth in the skill premium before age 40 but flatter growth after 40. I use a human capital investment model to account for the cross-cohort variation in skill premium profiles. The results indicate that the flattened growth after age 40 is caused by the drop in human capital (of high-skill workers) near the end of the life cycle. Besides, the magnitude of life-cycle growth in the skill premium is mainly driven by the relative skill price.

In chapter two and three, I study how technology usage affects earnings growth and earnings inequality over the life-cycle. In chapter 2, I construct a novel index to identify technology usage at the individual level using occupations as the proxy. I document technology usage patterns over the life-cycle and investigate its empirical relationship with labor earnings. I find that technology usage accounts for more than one-third of the growth in life-cycle inequality.

In chapter 3, I develop a life-cycle model with endogenous human capital investments and technology choices to quantify the relative importance of technology usage. The model features rich interactions between technology and human capital such that workers with high human capital are more likely to work with advanced technologies and vice versa. I find that technology usage contributes 31% of the growth in mean earnings and 46% of the growth in life-cycle inequality. I also evaluate policy implications of non-linear taxation on labor earnings. When tax progressivity on labor earnings is changed from US to European levels, the college attainment rate drops by 7 percentage points, and the growth in mean earnings decreases by 23%.

DEDICATION

To my parents.

ACKNOWLEDGEMENTS

I am indebted to Gustavo Ventura and Alexander Bick for their continuous guidance and support. I would like to thank Natalia Kovrijnykh for her helpful discussions and encouragement. I am also grateful for Bart Hobijn's guidance and advice. I am thankful for seminar and conference participants at ASU for their valuable comments and feedback. I also thank my peers for their helpful conversations and suggestions.

TABLE OF CONTENTS

CHAPTER Page

LIST OF TABLES

LIST OF FIGURES

Figure Page

Chapter 1

LIFE-CYCLE SKILL PREMIUMS ACROSS COHORTS

1.1 Introduction

This paper analyzes why the life-cycle profile of skill premiums varies across cohorts. There is a large body of literature that studies the life-cycle profile of earnings and how it changes across cohorts (Welch (1979), Berger (1985), Beaudry and Green (2000), Kambourov and Manovskii (2009a) and Jeong *et al.* (2015)). Literature has shown that the skill premium, which is the relative earning gap between high-skill workers and low-skill workers, has been increasing steadily since the 1960s because of skill-biased technological change, or equivalently the increasing relative demand for high-skill workers (Acemoglu *et al.* (2012)). However, there is little research that focuses on the combination of these two topics. I fill this gap by documenting life-cycle profiles of the skill premium for different cohorts.

I follow Kambourov and Manovskii (2009a) to construct the life-cycle profile of skill premiums and find that it varies significantly across cohorts. I divide the life cycle into two phases: before and after age 40. For younger cohorts, the first phase's growth becomes steeper, while the growth in the second phase flattens. In other words, the first phase's growth contributes to a more substantial proportion of the life cycle growth for younger cohorts. I further do a robustness check to show that the change in age patterns is still stark after controlling for time effects.

To investigate why the life-cycle profile changes across cohorts, I propose a human capital investment model based on Magnac *et al.* (2018). The wage rates (or rental rates of human capital) are determined through the "canonical model", a common workhorse in the skill-biased technological change literature (Katz and Murphy (1992)). For tractability, I assume that human capital accumulation is only for high-skill workers and that low-skill workers' human capital remains constant over time.[1] The life-cycle profile of skill premiums differs by cohorts for two reasons. First, cohorts (high-skill workers) are differentiated by the return of investment and depreciation, which determine the path of human capital accumulation. Second, the relative skill price, i.e. the log ratio of wage rates between high-skill and low-skill workers, is different across cohorts.

The human capital investment model is in the spirit of Ben-Porath (1967) but deviates from the existing literature in several aspects. First, the human capital production function is simplified such that the marginal return to investment is independent of the level of log human capital. Second, human capital investment does not require time allocation but only cause disutility. Third, I assume that there is no asset or capital market, so individuals do not borrow and save. Along with the functional form of utility, the model is able to generate a closed-form solution of investment decisions, and hence the life-cycle profile of human capital can be expressed explicitly as a function of age. Besides, depreciation is constant in the level of human capital, so the model could lead to a decline in human capital near the end of the life cycle, which is uncommon in the existing literature.

Using the implication of my model, I decompose the skill premium and find that both relative skill price and human capital are essential in explaining the life-cycle profile of skill premiums. The human capital profile changes notably across cohorts, which largely affects the shape of skill premium profiles. In particular, younger cohorts accumulate human capital faster in the first phase, but they also experience a drastic

[1] Since only the difference in human capital matters, normalization does not affect the result but only interpretation. Here I attribute all variation in the difference to high-skill workers.

drop in the second phase since their depreciation becomes larger. That's why the second phase's skill premium growth becomes flattened for younger cohorts. However, the magnitude of life-cycle growth in human capital does not change much across cohorts. So the magnitude of skill premium growth is largely driven by changes in the relative skill price. These two channels together account for the cross-cohort variation in skill premium profiles.

One contribution of my work is to extend the standard age-time-cohort framework by allowing interactions between age effects and cohort effects, which is the accumulation of human capital in the model. The baseline framework imposes a linearly additive structure on time, age, and cohort effects, which leads to a well-known colinearity problem. Most of the literature that studies life-cycle profiles usually make additional assumptions on either time effects or cohort effects for identification.[2] The closed-form solution derived from the model circumvents this problem so I can identify cohort-specific life-cycle profiles without such assumptions.

My work is also related to the literature that explores the relationship between skill-biased technological changes and skill premiums (Katz and Murphy (1992), Autor *et al.* (2008) and Acemoglu *et al.* (2012)). I extend the basic framework with the addition of human capital. Specifically, I separate human capital ratio from the observed skill premiums and focus on the ratio of wage rates. My estimation suggests that the elasticity of substitution between high-skill labors and low-skill labors is 2.85, and the annual growth rate of log technological change is 1.3 percentage points. Both results are consistent with the work of Acemoglu *et al.* (2012), which suggests that the standard skill-biased technological change hypothesis is robust to the inclusion of human capital.

[2]The most common assumption is to attribute any trend in the data to either time effects or cohort effects and set the other one to zero. See e.g. Lagakos *et al.* (2018).

Bowlus and Robinson (2012) and Bowlus *et al.* (2017) also study the relative skill price taking human capital into consideration, but their results are different from mine. This is caused by different implications of human capital investment models. Their identification strategy relies on the model from Heckman *et al.* (1998) in which human capital stays relatively constant near the end of the life cycle. So they identify fluctuation in earnings near the end of the life cycle as changes in wage rates. In contrast, in my model, human capital will suffer a sharp decline if the depreciation is high enough. Therefore, the difference in implications results in different interpretations of the relative skill price.

The paper is organized as follows. Section 1.2 presents empirical evidence about the life-cycle pattern of the skill premium by cohorts. In Section 1.3, I introduce the human capital model and use its implication to decompose the skill premium. Section 1.4 discusses the results. In Section 1.5, I conclude and discuss potential implications from my result.

1.2 Empirical Evidence

I begin by presenting the life-cycle profile of the skill premium for each cohort. In general, the skill premium profile for all cohorts is weakly increasing in age. That is, the wage differential between high-skill workers and low-skill workers is widening as workers get older. However, the shape of skill premium profiles varies significantly across cohorts. In particular, for younger cohorts, the growth before age 40 becomes steeper, but the growth after 40 is flattened. One potential explanation is that different cohorts experience different time effects throughout the life cycle. To explore this possibility, I show that the change in life-cycle profiles still stands out after controlling for time effects.

1.2.1 Data and Skill Premium

My analysis is based on the data from the Current Population Survey (CPS) with a concentration on the Annual Social and Economic Supplement (ASEC) over the years 1964-2019.[3] The measurement of earnings is hourly wages. The data harmonization process mostly follows Lemieux (2006) and the details are described in Appendix A. The only deviation is that I further limit my analysis to workers up to age 55. Casanova (2013) shows that life-cycle earnings profiles are largely affected by transitions into part-time work after age 55, either voluntary or involuntary. This transition issue could generate biased estimation on the skill premium among these old workers. Therefore I restrict observations below age 55.

The skill premiums are estimated from the Mincer regression. Following most literature (Katz and Murphy (1992) and Acemoglu *et al.* (2012)), I identify high-skill workers with college graduates and low-skill workers with the complement of college graduates. I run the following regressions for each age group j at period t:

$$\ln w_{i,t,j} = \gamma_0 + \omega_{t,j} C_{i,t,j} + X'_{i,t,j}\gamma_1 \tag{1.1}$$

where $w_{i,t,j}$ is the hourly wage for individual i of age group j at period t. $C_{i,t,j}$ is a dummy variable whether the individual has a college degree or not. The coefficient of interest is $\omega_{t,j}$, which captures the percentage wage gap between high-skill workers and low-skill workers within the age group j at period t. To adjust for the changing composition of the sample, I include $X_{i,t,j}$, a vector of control variables including sex, regions, marital status, and races.[4]

[3]The ASEC data starts from 1962. However, since the educational variable is not available in 1963, I begin my analysis from 1964.

[4]See more details in Appendix A.

1.2.2 Construction of Life-cycle Profiles

Though the CPS is repeated cross-sectional data, I use the method from Kambourov and Manovskii (2009a) and Kong *et al.* (2018) to construct pseudo panel data for synthetic cohorts. The rationale for building life-cycle profiles for each cohort is as follows. Since the CPS dataset is representative, I treat workers of age j in year t and workers of age $j+1$ in year $t+1$ as if they are from the same cohort. So the life-cycle profile of that cohort is given by the sequence $\{\omega_{t,j}, \omega_{t+1,j+1}, ...\}$.

To have less noisy life-cycle profiles of the skill premium, I follow Guvenen *et al.* (2015) to divide workers into five-year age bins (25-29, 30-34, ...). Similarly, I group years into 11 five-year periods.[5] Since both age bins and year bins have the same length of width, the logic of building life-cycle profiles above still works here but the time unit becomes five-year. I use the lower bound of the interval to refer to age bins or year bins. The cohort is indexed by the approximated birth year. For example, the 1939 cohort are workers between age 25 and 29 from 1964 to 1969.

I choose the CPS over standard longitudinal data sets like NLSY or PSID for two reasons. First, it allows my analysis to include more cohorts. For example, there are only two cohorts available in the NLSY. Second, the CPS data set gives a more accurate estimation of the skill premium comparable to other literature. As shown by Gouskova (2014), the PSID sample appears to be non-randomly selected on earnings, so the estimation on the skill premium is downward biased.

1.2.3 Life-cycle Profiles Vary Across Cohort

The life-cycle profiles are different across cohorts in two ways. First, younger cohorts go through a steeper growth before age 40. Second, younger cohorts have

[5] In Appendix A, I show that a more granular grouping does not significantly affect my empirical findings after smoothing.

a flatter growth after 40. From the life-cycle perspective, the magnitude of skill premium growth peaks with the 1954 cohort and then declines for the successive cohorts.

As shown in Figure 1, the life-cycle profile of skill premiums is non-decreasing in age for all six cohorts. The only exception is from the 1939 cohort where the skill premium slightly drops from 35 to 40. However, the growth pattern varies significantly across cohorts. For instance, the life-cycle profile of the 1944 cohort keeps increasing through the entire life cycle whereas profiles of the 1959 and 1964 cohort barely increase after age 40.

To illustrate the change in age patterns more formally, I summarize growth patterns in Table 1 for each cohort. The first column documents the life-cycle growth in the skill premium, i.e. the difference in the skill premium between age 25 and 50. To learn more about the structure of life-cycle growth, I divide the life cycle into two phases (before and after age 40) and document the growth in two periods separately. In the last column, I present the fraction of life-cycle growth that is accounted for by the growth in the first phase.

The first column in Table 1 shows that the magnitude of life-cycle growth changes significantly across cohorts. The life-cycle growth increases since the 1939 cohort and peaks with the 1954 cohort, who has a 32.4 percentage point increase in the skill premium through the life-cycle. After that, it declines for younger cohorts. The 1964 cohort only has a 20.8 percentage point of the life-cycle growth.

To understand the composition of the life-cycle growth, I break up the life-cycle into two phases: before and after age 40. The second column indicates that the growth in the first phase is relatively small for earlier cohorts. The 1939 cohort has the lowest growth of 5.2 percentage points in the first phase. The growth increases from the 1939 cohort and reaches to 28.1 percentage points for the 1954 cohort. Though the number declines after the 1954 cohort, it is still above 20 percentage points. The third column shows that the growth in the second phase shrinks for younger cohorts. For example, the growth of the 1939 cohort is 11.8 percentage points, whereas the number is close to zero for the 1959 and 1964 cohort.

These two patterns explain the increasing fraction shown in the last column, i.e. the first phase's growth accounts for a larger proportion of the life-cycle growth for younger cohorts. For the 1939 cohort, the first phase's growth contributes to 30.5% of the life-cycle growth. This number rises to 98.1% for the 1964 cohort.

1.2.4 Controlling for Time Effects

The life-cycle profile could change across cohorts because of different time effects that each cohort faces through their life cycle. Valletta (2016) shows that the growth of skill premiums slows down since the 1990s due to weakening in demand for high-skill workers. So an earlier cohort would experience a faster growth in the skill premium than a younger cohort does. In this subsection, I use a statistical model to show that the change in age patterns still stands out after controlling for time effects.

One classical framework to analyze the life-cycle profile is to decompose the skill premium of age group j at time t from cohort $t-j$ into three linearly additive terms: time, age, and cohort effects.

$$\omega_{t,j} = s_t^{time} + s_j^{age} + s_{t-j}^{cohort} + \epsilon_{t,j}$$

A well-know underidentification problem (see Deaton and Paxson (1994)) arises since a person's age added to the cohort year (birth year) gives the current year which means there is an exact linear relationship between time, age, and cohort effects. To avoid the problem of colinearity, I combine several cohorts into one broad cohort based on the order of birth year and index them by g. So the cohort effect is not specific to each cohort i but varies by broader category g. The detail of grouping is described in Appendix A.

Besides, this framework suggests that cohorts are only differentiated by a level term s^{cohort} after controlling for time effects. To better understand the difference between cohorts, I add an interaction term between age effects and cohort effects to capture potential changes in age patterns across cohorts. In particular, the skill premium of age group j at time t from the broad cohort g can be decomposed as follows:

$$\omega_{t,j,g} = s_t^{time} + s_j^{age} + s_g^{cohort} + s_j^{age} \cdot s_g^{cohort} + \epsilon_{t,j} \tag{1.2}$$

Figure 2 shows life-cycle profiles after controlling for time effects s_t^{time}. The difference in age patterns across cohorts is still prominent. In particular, the growth in the second phase becomes smaller for younger cohorts. The profile for the 1939 and 1944 cohort increases with age through the entire life cycle but the growth slows down after age 35. For the 1949 and 1954 cohort, the skill premium reaches to peak at age 40 and then starts to decline. The peak even comes earlier (age 35) for the 1959 and 1964 cohort.

In Table 2, I document growth patterns after controlling for time effects. As shown in the third column, the second phase's growth still keeps declining for younger cohorts, which is consistent with the pattern from the raw data in Table 1. Besides, the life cycle growth decreases substantially for younger cohort cohorts, as shown in the first column. This evidence indicates that the time effect itself cannot explain the cross-cohort variation in life-cycle profiles. Moreover, the change in profiles becomes more drastic after separating time effects. This suggests that life-cycle profiles are driven by factors that are different across cohorts.

1.3 A Model to Decompose the Skill Premium

In this section, I propose a human capital investment model from Magnac *et al.* (2018). The advantage of the model is that it generates a closed-form solution of life-cycle profiles of earnings that can be readily linked to the data. To better understand the determination of wage rates across skill groups, I embed human capital investment decisions within a CES production technology with two factors (high-skill and low-skill labor inputs). The model provides a theoretical framework to decompose the skill premium into the price (wage rate) and quantity (human capital).

1.3.1 Technology

Production of goods combines high-skill labor H_t and low-skill labor L_t measured in efficiency units (or equivalently human capital), using the following technology:

$$Y_t = [(A_{H,t}H_t)^{\frac{\sigma-1}{\sigma}} + (A_{L,t}L_t)^{\frac{\sigma-1}{\sigma}}]^{\frac{\sigma}{\sigma-1}} \tag{1.3}$$

where σ is the elasticity of substitution between high and low skill labors and $A_{H,t}$ ($A_{L,t}$) is factor-augmenting technological change for high-skill (low-skill) labor.

If the labor market is competitive, wage rates per efficiency unit are given by marginal products and can be obtained by taking first order conditions:

$$W_{L,t} = Y_t^{\frac{1}{\sigma-1}} A_{L,t}^{\frac{\sigma-1}{\sigma}} L_t^{-\frac{1}{\sigma}} \quad \text{and} \quad W_{H,t} = Y_t^{\frac{1}{\sigma-1}} A_{H,t}^{\frac{\sigma-1}{\sigma}} H_t^{-\frac{1}{\sigma}} \tag{1.4}$$

I refer to the log ratio of two wage rates as the *relative skill price*, which is given by:

$$\ln \frac{W_{H,t}}{W_{L,t}} = \frac{\sigma-1}{\sigma} \ln(\frac{A_{H,t}}{A_{L,t}}) - \frac{1}{\sigma} \ln(\frac{H_t}{L_t}) \tag{1.5}$$

This equation shows two competing forces that determine the relative skill price: the relative technological change $\frac{A_{H,t}}{A_{L,t}}$ and the relative supply $\frac{H_t}{L_t}$.

Following Goldin and Katz (2009) and Acemoglu *et al.* (2012), I assume that technological changes are skill-biased in the sense that $A_{H,t}$ always grows faster than $A_{L,t}$. If high-skill workers and low-skill workers are substitutes[6] ($\sigma > 1$), then increasing skill-biased technological change would increase the relative demand for high-skill workers and push up the relative skill price. Meanwhile, an increase in the relative supply could reduce the relative skill price with an elasticity of $\frac{1}{\sigma}$.

1.3.2 Human Capital Investment

The economy is populated by overlapping cohorts of individuals. A cohort is born each period and live J periods. Cohorts are indexed by their birth year i. For example, individuals of age group j at time t are from the $t-j$ cohort. There are two types of agents in each cohort: high-skill workers (H) and low-skill workers (L). The fraction of high-skill labors in cohort i is exogenously given by λ_i. Within high or low skill workers, individuals are homogeneous, so one can treat a cohort as a weighted average of two representative agents.

Human capital is heterogenous across skill groups. However, due to identification difficulty[7], I follow Keller (2014) to assume that the low-skill workers do not invest in human capital and normalize their human capital to 1. Hence the following human capital investment channel is only for high-skill workers.

Individual's Problem

A (high-skill) individual from cohort i maximizes expected lifetime utility by choosing the optimal investment decisions $\{x_{j,i}\}_{j=1}^{J}$ and consumption decisions $\{c_{j,i}^{H}\}_{j=1}^{J}$ given

[6] Ciccone and Peri (2005) show that the long-run elasticity of subsitution in the U.S. is around 1.5.

[7] In short, I cannot directly decompose the observed earnings into wage rate and efficiency units without further assumption. See more details in Appendix A

the initial human capital $h_{1,i}$:

$$\max_{\{x_{j,i},c^H_{j,i}\}_{j=1}^J} \quad E\left[\sum_{j=1}^{J}\beta^{j-1}\left(\ln c^H_{j,i} - \phi_i \frac{x^2_{j,i}}{2}\right)\right] \tag{1.6}$$

The individual's utility is the log of consumption $c^H_{j,i}$ net of investment cost adjusted by a cohort-specific parameter ϕ_i.

Since I focus on life-cycle earnings rather than consumption, I assume there is no asset or capital market so that individuals do not save or borrow and only consume their current earnings $w_{j,i}$, which is a product of human capital $h_{j,i}$ and wage rate per efficiency unit at time $j+i$:

$$c^H_{j,i} = w_{j,i} = W_{H,j+i}h_{j,i} \tag{1.7}$$

The following equation describes the technology of human capital production where ρ_i is the cohort-specific rate of return of investment, and δ_i is the depreciation of human capital.

$$\ln h_{j+1,i} = \ln h_{j,i} + \rho_i x_{j,i} - \delta_i \tag{1.8}$$

The human capital investment channel deviates from the existing literature mainly in two aspects. First, I assume that human capital investment does not require time allocation, which is different from Magnac *et al.* (2018). This additional assumption has two advantages. First, workers supply one unit of time inelastically so I could directly link the observed hourly wages to earnings in the model without adjusting for working time. Second, the cost of investment is only captured by the disutility term $\phi_i \frac{x^2_{j,i}}{2}$. The marginal disutility of investment is zero when $x_{j,i} = 0$. The constant rate of return of investment ρ_i implies that the marginal benefit of investment is always positive except for the last period since human capital will contribute to the continuation value (future earnings). Thus, individuals will make a positive amount of investment until the last period.

Second, the human capital production is different from the standard Ben-Porath technology. Unlike most literature[8], the marginal return to investment in terms of log human capital equals ρ_i, which is independent of the stock of log human capital and investment. This modification does not twist the spirit of the investment channel. Combining with the functional form of utility, it generates a closed-form solution which I will show later. Moreover, the depreciation δ_i is constant in the level of log human capital. This constant depreciation could lead to a drastic decline in human capital near the end of the life-cycle which is uncommon in the existing literature.

Optimal Investment Decisions

To understand how human capital accumulation works, I rewrite the individual's problem of cohort i resursively:

$$V_{j,i}(h_{j,i}) = \max_{x_{j,i} \in [0,\infty]} \quad \ln W_{H,j+i} + \ln h_{j,i} - \phi_i \frac{x_{j,i}^2}{2} + \beta E[V_{j+1,i}(h_{j+1,i})] \tag{1.9}$$

where I replace $\ln w_{H,j+i}$ with $\ln W_{H,j+i} + \ln h_{j,i}$ using equation (1.7). The log utility ensures that the flow of utility (net of investment cost) per period is just the sum of log wage rate and log human capital. The no consumption smoothing assumption greatly simplies the recursive problem such that only the dynamic of human capital matters.

The first order condition with respect to $x_{j,i}$ equates the marginal cost and the marginal utility of investment:

$$x_{j,i}\phi_i = \beta E[\frac{\partial V_{j+1,i}}{\partial \ln h_{j+1,i}}] \frac{\partial ln h_{j+1,i}}{\partial x_{j,i}} = \beta E[\frac{\partial V_{j+1,i}}{\partial \ln h_{j+1,i}}]\rho_i \tag{1.10}$$

[8]The most commonly used functional form is $h_{j+1} = (1-\delta)h_j + \alpha(nh_j)^\phi$ where n is the time allocated to human capital investment and $\phi \in (0,1)$. So the marginal return to investment is increasing in the level of human capital. See e.g. Huggett *et al.* (2011).

where the last equality comes from the assumption that the rate of return is constant in the level of log human capital as shown in equation (1.8).

Applying the envelope theorem[9] to equation (1.9) generates

$$\frac{\partial V_{j,i}}{\partial \ln h_{j,i}} = 1 + \beta E[\frac{\partial V_{j+1,i}}{\partial \ln h_{j+1,i}} \frac{\partial \ln h_{j+1,i}}{\partial \ln h_{j,i}}] = 1 + \beta E[\frac{\partial V_{j+1,i}}{\partial \ln h_{j+1,i}}] \tag{1.11}$$

Since the individual will not make any investment in the last period, the value function in the last period is $V_{J,i}(h_{J,i}) = \ln W_{H,J+i} + \ln h_{J,i}$ which implies $\frac{\partial V_{J,i}}{\partial \ln h_{J,i}} = 1$. Using this result, I can rewrite equation (11) as follows

$$\frac{\partial V_{j,i}}{\partial \ln h_{j,i}} = 1 + \beta + ... + \beta^{J-j} = \frac{1 - \beta^{J-j}}{1 - \beta} \tag{1.12}$$

Taking this equation to equation (1.10) yields the optimal investment decisions:

$$x_{j,i} = \frac{\rho_i \beta}{\phi_i} \frac{1 - \beta^{J-j}}{1 - \beta} \quad \text{for all} \quad 1 \leq j \leq J - 1 \tag{1.13}$$

This equation shows the key result that investment decreases with age. This is due to the fact that the present discounted value of one additional unit of (log) human capital only depends on the number of periods left and becomes smaller as individuals getting older, which is shown in equation (1.12).

It is also intuitive that the investment is decreasing in cost parameter ϕ_i and increasing in discount factor β and the rate of return ρ_i. Moreover, though individuals are forward-looking, future wage rates do not affect their decision today. That is, technological changes have no influence on human capital accumulation. This results from the log utility where substitution effects cancel out income effects.[10]

[9] Since the value function is age-dependent, I need to apply the chain rule to $\beta E[V_{j+1,i}(h_{j+1,i})]$.

[10] Kong *et al.* (2018) also impose no consumption smoothing assumption but individuals maximize earnings instead of log utility. So future wage rates would affect investment decisions in their model.

Life-cycle Profiles of Human Capital

Using optimal investment decisions from equation (1.13), I can derive the life-cycle profile of log human capital as follows:

$$\ln h_{j,i} = \ln h_{1,i} + \left(\frac{\rho_i \beta}{\phi_i(1-\beta)} - \delta_i\right) \cdot (j-1) + \frac{\rho_i \beta^{J+1}}{\phi_i(1-\beta)^2} \cdot \left(1 - \beta^{-(j-1)}\right) \quad (1.14)$$

The life-cycle profile consists of three components. The first term $\ln h_{1,i}$ determines the initial condition when entering the labor market at $j = 1$. The second term is a linear function of age and the third term governs the curvature of the growth.

In the absence of the depreciation, the last two terms generate a concave and increasing profile of log human capital. The reason is that investment is always positive and decreases with age as shown in equation (1.13). The speed of growth depends positively on the rate of return ρ_i and negatively on investment cost ϕ_i, which is also in line with investment decisions shown above.

However, if the depreciation is considerable, the model could generate a drastic decline near the end of the life cycle, which is uncommon in the literature.[11] The intuition is as follows. Since the value of human capital decreases with age, individuals make less investment toward the end of the life-cycle while the depreciation is constant in the level of log human capital. Therefore the depreciation will outweigh the magnitude of growth near the end of the life-cycle, leading to a decline in human capital.

1.3.3 Competitive Equilibrium

Before introducing the competitive equilibrium, I briefly describe the allocation for low-skill individuals. Since I normalize the human capital of low-skill workers to 1, the earnings for low-skill workers of age group j at time t is $w_{L,t,j} = W_{L,t}$. So their

[11]In Heckman *et al.* (1998), the human capital remains relatively stable near the end of life-cycle.

consumption decisions are given by $c^L_{j,i} = W_{L,j+i}$.

Definition A competitive equilibrium is an allocation of decisions $\{\{x_{j,i}\}^J_{j=1}, \{c^L_{j,i}\}^J_{j=1}, \{c^H_{j,i}\}^J_{j=1}\}^{\infty}_{i=-\infty}$ and wage rates $\{W_{H,t}, W_{L,t}\}^{\infty}_{t=-\infty}$ such that:

1. Given wage rates and the initial condition, individuals from cohort i choose $\{\{x_{j,i}\}^J_{j=1}, \{c^L_{j,i}\}^J_{j=1}, \{c^H_{j,i}\}^J_{j=1}\}$ optimally.

2. Given technological changes $A_{L,t}$ and $A_{H,t}$, prices equal marginal productivity: $W_{L,t} = Y_t^{\frac{1}{\sigma-1}} A_{L,t}^{\frac{\sigma-1}{\sigma}} L_t^{-\frac{1}{\sigma}}$ and $W_{H,t} = Y_t^{\frac{1}{\sigma-1}} A_{H,t}^{\frac{\sigma-1}{\sigma}} H_t^{-\frac{1}{\sigma}}$.

3. Labor market clears: $H_t = \sum_{j=1}^J \lambda_{t-j} \cdot h_{j,t-j}$ and $L_t = \sum_{j=1}^J (1-\lambda_{t-j})$ where $h_{j,t-j}$ follows the law of motion described in equation (1.8) given $\{\{x_{j,i}\}^J_{j=1}\}^{\infty}_{i=-\infty}$.

4. Good market clears: $Y_t = \sum_{j=1}^J \lambda_{t-j} \cdot c^H_{j,t-j} + \sum_{j=1}^J (1-\lambda_{t-j}) \cdot c^L_{j,t-j}$

Even though my model is built on a general equilibrium framework, the interaction between the human capital channel and the technology is unilateral. Wage rates $W_{H,t}$ and $W_{L,t}$ have no impact on (high-skill) individuals' investment decisions, though they are affected by human capital through the supply side.

1.3.4 Decomposition of the Skill Premium

From equation (1.7), I can rewrite the earnings for high-skill workers as indexed by age group j and time t:

$$w_{H,t,j} = W_{H,t} h_{j,t-j}$$

Using equation (1.5), the skill premium of age group j at period t can be expressed as follows:

$$\begin{aligned} \omega_{t,j} \equiv \ln \frac{w_{H,t,j}}{w_{L,t,j}} &= \ln h_{j,t-j} + \ln \frac{W_{H,t}}{W_{L,t}} \\ &= \underbrace{\ln h_{j,t-j}}_{\text{human capital}} + \underbrace{\frac{\sigma-1}{\sigma} \ln(\frac{A_{H,t}}{A_{L,t}}) - \frac{1}{\sigma} \ln(\frac{H_t}{L_t})}_{\text{relative skill price}} \end{aligned} \tag{1.15}$$

This decomposition shows that the cross-cohort variation of life-cycle profiles of the skill premium comes from two aspects: relative skill price and human capital. In the language of the age-time-cohort framework, the first one can be interpreted as time effects. The latter is a combination of age effects and cohort effects, which is different from the linearly additive structure in the baseline framework.

Different cohorts face different relative skill prices through their life cycles. Suppose that the relative skill price declined due to a negative shock, some cohorts experience this price drop during their early life-cycle while other cohorts experience this shock near the end of their life-cycle. So this shock affects life-cycle profiles differently across cohorts. Since the relative skill price is universal to all age groups, it can be treated as time effects.

The accumulation of human capital varies across cohorts. Equation (1.14) suggests that the life-cycle profile of human capital is a function of age. Furthermore, the function depends on cohort-specific parameters: the rate of return ρ_i, investment cost ϕ_i and the depreciation δ_i. Therefore the life-cycle profile of human capital is a mixture of age effects and cohort effects.

1.4 Quantitative Analysis

In this section, I use the implication from my model to decompose the skill premium by cohorts. The results indicate that both human capital and relative skill price are important in explaining the cross-cohort variation in the skill premium profiles. In particular, the magnitude of life-cycle growth in the skill premium mainly depends on the change in the relative skill price since the growth in human capital is similar across cohorts. However, the shape of skill premium profiles is largely affected by human capital profiles. Specifically, the growth of human capital after age 40 becomes smaller and even negative for younger cohorts since their depreciation becomes larger.

Hence, younger cohorts have flattened growth in the skill premium after 40.

1.4.1 Identification

Equation (1.15) provides a theoretical framework to break up the skill premium into the price and quantity. Given the functional form of human capital profiles, I can directly identify the evolution of human capital by substituting $\ln h_{j,t-j}$ with equation (1.14). Specifically, I regress the skill premium $\omega_{t,j}$ of age group j at period t on two age-dependent factors and a time dummy variable as follows:

$$\omega_{t,j} = \beta_{0,i} + \beta_{1,i} \cdot (j-1) + \beta_{2,i} \cdot \left(1 - \beta^{-(j-1)}\right) + \alpha_t + \epsilon_{t,j} \tag{1.16}$$

where β is the discount factor[12] and α_t is the time fixed effect representing the relative skill price $\ln \frac{W_{H,t}}{W_{L,t}}$.

Recovering Structural Parameters

The relationship between reduced-form coefficients and deep parameters are described below:

$$\beta_{0,i} = \ln h_{1,i}$$
$$\beta_{1,i} = \frac{\rho_i \beta}{\phi_i (1-\beta)} - \delta_i$$
$$\beta_{2,i} = \frac{\rho_i \beta^{J+1}}{\phi_i (1-\beta)^2}$$

From the estimation on $\beta_{2,i}$, I can back out the ratio of rate of return and investment cost $\frac{\rho_i}{\phi_i}$. This compound parameter can be interpreted as the net return of investment. After obtaining $\frac{\rho_i}{\phi_i}$, I can recover the deprecation for each cohort since $\delta_i = \frac{\rho_i \beta}{\phi_i(1-\beta)} - \beta_{1,i}$. These two parameters are essential in shaping cohort-specific life-cycle profiles of human capital.

[12]The annual discount factor is exogenously taken from the literature and chosen to be 0.98. Since my time unit here is five-year, I set $\beta = 0.98^5$.

Cohort-specific Parameters

One concern about identifying the life-cycle profile of human capital is that the number of observations might be inadequate. First, my panel observations are truncated in the sense that the complete life-cycle profile is not available for every cohort. For example, the oldest cohort comprises people between age 50 and 55 from 1964 to 1969 and that's the only observation from this cohort. So it is infeasible to identify two cohort-specific parameters for them. Second, since I divide workers into six age groups (25-29, ..., 50-55), the complete life-cycle only contains six observations. It is possible that the non-linear form cannot fully capture the life-cycle profile from six points.

To have a more robust estimation on human capital accumulation, I assume that cohort-specific parameters vary by the broad cohorts g that I use in section 2.4. In particular, I divide 16 cohorts into seven groups based on the order of birth and let β_1 and β_2 vary across groups. Besides, since I focus on life-cycle profiles, I further simplify the model by assuming the initial human capital is fixed across cohorts, i.e. $\ln h_{1,i} = \ln h_1$ for all i.

Relative Skill Price

To unpack the relative skill price α_t, I follow the literature (Acemoglu *et al.* (2012)) and assume that there is a log linear increase in skill-biased technological change, captured in the following form[13]:

$$\ln \frac{A_{H,t}}{A_{L,t}} = \eta \cdot t$$

where η is the log growth rate. So the relative skill price can be expressed as follows:

[13]Since I estimate the relative skill price as time fixed effects, the (log) skill price is normalized to 0 at $t = 1$. Therefore I drop the constant term in the regression.

$$\ln \frac{W_{H,t}}{W_{L,t}} = \frac{\sigma - 1}{\sigma}\eta \cdot t - \frac{1}{\sigma} \ln \frac{H_t}{L_t}$$

Given the estimated skill price $\hat{\alpha}_t$ from regression (1.16), I can further estimate the following equation:

$$\hat{\alpha}_t = \beta_3 \cdot t + \beta_4 \ln \frac{H_t}{L_t} + u_t \tag{1.17}$$

where $\beta_3 = \frac{\sigma-1}{\sigma}\eta$, $\beta_4 = -\frac{1}{\sigma}$, and u_t represents unobservable shocks. The coefficient β_3 shows the growth rate in the skill premium resulting from skill-biased technological change. The reciprocal of the absolute value of β_4 is the elasticity of substitution in the production function.

The relative supply $\frac{H_t}{L_t}$ can be constructed after obtaining the human capital. Specifically, from the estimation on (1.16), human capital can be backed out as

$$\ln \hat{h}_{t,j} = \hat{\beta}_{1,i} \cdot (j-1) + \hat{\beta}_{2,i} \cdot \left(1 - \beta^{-(j-1)}\right)$$

The relative supply is then given by

$$\frac{H_t}{L_t} = \frac{\sum_{j=1}^{J} (\lambda_{t-j} \cdot \hat{h}_{j,t-j})}{\sum_{j=1}^{J} (1 - \lambda_{t-j})} \tag{1.18}$$

where λ_{t-j} is the fraction of high-skill workers in cohort $t-j$ that can be directly taken from the data.

Equation (1.17) is analogous to the so-called "canonical model" that studies the relationship between the skill premium and skill-biased technological change.[14] However, I extend the framework by allowing for human capital formation. In the canonical model, the dependent variable is the skill premium whereas I use the ratio of wage rates which is the skill premium after controlling for human capital. Moreover, the labor supply in the regression is adjusted for human capital changes as shown in equation (1.18).

[14] See e.g. Katz and Murphy (1992), Acemoglu *et al.* (2012) and Autor (2017).

1.4.2 *Human Capital Profiles*

I first investigate how human capital profiles change across cohorts. My estimation shows that both the net return of investment and depreciation becomes larger for younger cohorts. This pattern leads to a sharp decline in human capital near the end of the life cycle for younger cohorts.

Table 3 shows the structural parameters for seven broad groups. Since the estimations of the first and the last group are obtained from incomplete life-cycle observations, their results are not comparable to the other cohorts'. For instance, the depreciation for the first group is negative. This is caused by data limitation that I cannot observe the skill premium before age 40 from the first group. That is, there is not enough variation to tell the accumulation and the depreciation apart.

As shown in the table, both the net return and depreciation become higher for younger groups. For the 1929 and 1934 cohort, their net return of investment is 0.017 and the depreciation is 0.002. The net return of the sixth group (1969/1974 cohort) increases to 0.062 and the depreciation also rises to 0.137.

Figure 3 presents the life-cycle profiles of human capital based on the estimation. As discussed above, the net return determines how fast (high-skill) workers could accumulate their human capital mainly in the early stage of the life cycle. For example, the net return from the 1959 cohort is almost twice as large as the 1949 cohort's. Consequently, the growth of log human capital from 25 to 40 of the 1959 cohort is 5 percentage points higher than the growth of the 1949 cohort.

The depreciation also plays a vital role in governing the path, especially near the end of life-cycle. As shown in Figure 3, the life-cycle growth of the 1929 cohort continues after age 40 while the log human capital of the 1939 and 1944 cohort barely increases after 40. Furthermore, for the 1959 cohort and its successive cohorts, the human capital profile starts to decline after age 40 and the magnitude of the drop is considerable. The significant change after age 40 is accounted for by the increasing depreciation as shown in the second column of Table 3.

1.4.3 Relative Skill Price

Figure 4 shows the estimated relative skill price α_t which declines from 1964 to 1979 and then increases steadily afterward.[15] As shown in Figure 4, the relative supply grows fastly in the 1970s. This rapid increase in the relative supply outweighs the growth rate of skill-biased technological change so the relative skill price declines during this period. After that, the growth of relative supply slows down, so the relative skill price keeps increasing.

To better understand the evolution of the relative skill price, I fit equation (1.17) to obtain the following estimate[16]:

$$\alpha_t = \underset{(0.002)}{0.064} \cdot t - \underset{(0.013)}{0.351} \cdot \ln \frac{H_t}{L_t} \tag{1.19}$$

This first coefficient implies that the growth of relative skill price driven by demand shifting toward high-skill workers is 1.3 log points per year.[17] The point estimate on the relative supply term suggests that the elasticity of substitution between high-skill workers and low-skill workers is $1/0.351 = 2.85$. Both results are in line with the work from Acemoglu *et al.* (2012), where they fit the skill premium through the "canonical model" and find that the annual growth rate is 1.6 log points and that the elasticity of substitution is 2.94 using the CPS data from 1963 to 2008.

[15] The downward trend before 1980 is also documented by Acemoglu *et al.* (2012).

[16] Standard errors are shown in parentheses

[17] Again, the time unit in my analysis is five-year so the annual growth is the coefficient divided by 5.

The similarity of results indicates that the standard skill-biased technological change hypothesis is robust to the addition of human capital. The inclusion of human capital alters the basic framework in two aspects. First, in the "canonical model", the dependent variable is the ratio of earnings which also reflects the ratio of human capital. I separate changes in human capital and use the ratio of wage rates between high-skill and low-skill workers as the outcome variable. Second, the labor supply is constructed with adjustment for human capital that is estimated from skill premiums profiles. Surprisingly, these two modifications do not significantly change the estimation result compared to Acemoglu *et al.* (2012), so the skill-biased technological change hypothesis is not affected by the incorporation of human capital.

My analysis of the relative skill price is not the first to take human capital into account. Bowlus and Robinson (2012) also decompose the skill premium in a similar way but their estimation is different from mine. Their estimated relative skill price starts declining from the mid-1990s, which is inconsistent with standard skill-biased technological change hypothesis.

This discrepancy is caused by different implications from different human capital models. Their identification relies on the implication that there is a period near the end of the life cycle where human capital is constant. So any variation in observed earnings during that period is treated as changes in skill prices from which they construct the time series of price.[18] In contrast, the human capital in my model could decline near the end of the life cycle if the depreciation is high. That is, I interpret fluctuations in observed earnings near the end of the life cycle as a combination of human capital changes and relative skill price changes. This explains why we have different conclusions on the path of the relative skill price.

[18] This approach is based on the human capital theory proposed by Heckman *et al.* (1998) and used by many papers. See e.g. McKenzie (2006), Lagakos *et al.* (2018) and Schulhofer-Wohl (2018).

1.4.4 Decomposition

Now I put human capital profile and relative skill price together and see how they contribute to the skill premium profiles. My decomposition shows that the magnitude of the life-cycle growth in the skill premium is largely affected by the change in the relative skill price since human capital growth is similar across cohorts. However, the shape of skill premium profiles is mainly governed by human capital accumulation. In particular, the drop in human capital near the end of the life-cycle offset the growth in the relative skill price, which leads a flattened growth in the skill premium for younger cohorts.

I formally document the growth contributed by relative skill price and human capital separately for each cohort in Table 4. On the one hand, the life-cycle growth in human capital does not change substantially across cohorts as shown in the second column. The 1939 and 1944 cohort have a 13.8 percentage points life-cycle growth in human capital, and the growth slightly drops 2 percentage points for the next four following cohorts. On the other hand, the life-cycle growth in the relative skill price varies drastically over cohorts as shown in the third column. For example, the 1939 cohort goes through an increase of 3.9 percentage points in the relative skill price, whereas the increase jumps to 16.9 percentage points for the 1954 cohort. Therefore, the magnitude of skill premium growth is largely determined by the progress of the relative skill price.

The fourth column in Table 4 shows the fraction of growth that can be accounted for by the relative skill price, which also suggests how important the relative skill price is from the life-cycle perspective. The fraction increases from the 1939 cohort (22.0%) to the 1954 cohort (59.7%), though it slightly drops after the 1954 cohort.

In Table 5, I break up the growth into two phases and study the roles of human capital and relative skill price separately. The result suggests that the flattened growth in the second phase is due to the drop in the human capital profiles.

From the 1939 cohort to the 1954 cohort, the relative skill price is crucial in explaining skill premium profiles. The human capital profile does not change much in both periods. However, the relative skill price that different cohorts face is quite different. For example, for the 1939 cohort, the change in the relative skill price between 25 and 40 is -6.5 percentage points. This number increases to 13.3 for the 1954 cohort. Similarly, the growth after 40 becomes smaller for successive cohorts, decreasing from 10.4 percentage points to 3.6 percentage points. These results suggest that the life-cycle profile is largely affected by the relative skill price before the 1954 cohort.

Human capital contributes more to the variation in skill premium profiles after the 1954 cohort because the human capital profile changes dramatically. The 1959/64 cohort have a faster growth before 40 compared to previous cohorts. Though the growth of the relative skill price slows down, the growth in skill premiums before 40 does not change because of the faster growth in human capital. Besides, the 1959 and 1964 cohort also suffer a sizeable decline (5.2 percentage points) in human capital near the end of the life cycle which explains why the life-cycle profile becomes flattened after 40.

1.5 Conclusion

In this paper, I document how the life-cycle profile of skill premiums varies across cohorts and study the cross-cohort difference through a human capital investment model. The decomposition indicates that both relative skill price and human capital are crucial in explaining the cross-cohort variation in skill premium profiles. In particular, the relative skill price mainly determines the extent of life-cycle growth in the skill premium. Besides, human capital accumulation varies significantly across cohorts which largely affects the shape of skill premium profiles.

The decomposition also has implications on skill-biased technological change hypothesis. I separate the ratio of wage rates from the observed skill premium and fit it with the canonical model. The estimation result is consistent with the literature, which indicates that the skill-biased technological change hypothesis is robust to the addition of human capital.

The identification in this paper largely relies on the closed-form solution from the human capital investment model which imposes several strong assumptions that are uncommon in the literature. Several potentially important channels are ignored from these simplifications. First, I assume there is no asset and capital market allowing for borrowing and saving for individuals. The omit of this channel will greatly simplify the problem as human capital investment is the only source of intertemporal choice. Second, the ongoing skill-biased technological change has no impact on investment decisions because of special functional forms. Kong *et al.* (2018) show that if individuals could foresee the rise in the price, they will invest more in human capital, which is another missing margin in my model. Adding these channels will not overturn the basic mechanism of human capital accumulation but it will generate more interesting features and make the accumulation process less mechanical.

Though the human capital model might not be a perfect explanation, the key insight here is to decompose the skill premium into the price and the quantity. The results of the quantity have several implications worth discussing. For example, my work shows that cohorts are differentiated by the rate of return to investment and the depreciation but I did not answer why they are different. One potential explanation is that the school quality changes over time so that high-skill workers experienced different college education which affects their accumulation later. This can also be linked with the work from Goldin and Katz (2009) where they emphasize the role of the slowdown in the quantity and quality of schooling in the U.S. wage inequality.

Another possible direction is to interpret the quantity beyond the scope of human capital accumulation. Acemoglu and Restrepo (2020b) propose a task-based framework to explain the rising skill premium. They find that the essence of skill-biased technological change is new tasks replacing old tasks. A related question is how can we connect the formation of efficiency units to this replacement process. Does a more frequent replacement of tasks strengthen or harm the accumulation of efficiency units? Besides, one can also link this to the recent occupational trends (e.g. polarization) in the labor market under this framework.

Table 1: Life-cycle Skill Premium Growth Patterns

Cohort	Life-cycle growth	First phase	Second phase	$\frac{\text{First phase's growth}}{\text{Life-cycle growth}}$
1939	0.170	0.052	0.118	30.5%
1944	0.236	0.134	0.102	56.8%
1949	0.305	0.236	0.069	77.4%
1954	0.324	0.281	0.043	86.7%
1959	0.234	0.226	0.008	96.6%
1964	0.208	0.204	0.004	98.1%

Note: The life cycle growth is the difference in the skill premium between age 25 and 50. The first phase's growth is the difference between age 25 and 40. The second phase's growth is the difference between age 40 and 50.

Table 2: Growth Patterns Controlling for Time Effects

Cohort	Life-cycle growth	First phase	Second phase
1939, 1944	0.168	0.140	0.028
1949, 1954	0.134	0.159	-0.025
1959, 1964	0.060	0.096	-0.036

Note: The life cycle growth is the difference in the skill premium net of time effects s_t^{time} between age 25 and 50. The first phase's grwoth is the difference between age 25 and 40. The second phase's growth is the difference between age 40 and 50.

Table 3: Cohort-specific Structural Parameters

Group	Cohort	Net return $\frac{\rho_i}{\phi_i}$	Depreciation δ_i
1*	1914, 1919, 1924	0.008	-0.018
2	1929, 1934	0.017	0.002
3	1939, 1944	0.023	0.026
4	1949, 1954	0.025	0.036
5	1959, 1964	0.046	0.085
6	1969, 1974	0.062	0.137
7*	1979, 1984, 1989	0.024	0.016

*: The life-cycle observations for these cohorts are incomplete so they are not comparable to other cohorts

Table 4: Decomposition of the Life-cycle Growth

Cohort	Skill Premium	Human Capital	Relative Skill Price	$\frac{\text{Relative Skill Price}}{\text{Skill Premium}}$
1939	0.177	0.138	0.039	22.0%
1944	0.217	0.138	0.079	36.4%
1949	0.258	0.114	0.144	55.8%
1954	0.283	0.114	0.169	59.7%
1959	0.231	0.113	0.118	51.1%
1964	0.218	0.113	0.105	48.2%

Note: The first three columns show the difference in (log) skill premium/human capital/relative skill price between age 25 and 50.

Table 5: Detailed Decomposition of Growth Patterns

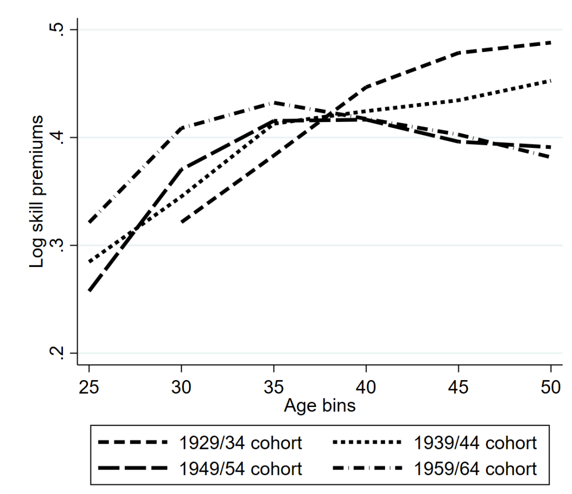
Cohort	Growth before 40		Growth after 40	
	Human Capital	Relative Skill Price	Human Capital	Relative Skill Price
1939	0.131	-0.065	0.007	0.104
1944	0.131	0.013	0.007	0.066
1949	0.122	0.090	-0.008	0.054
1954	0.122	0.133	-0.008	0.036
1959	0.165	0.091	-0.052	0.027
1964	0.165	0.064	-0.052	0.041

Note: The growth before 40 is the difference in (log) human capital/relative skill price between age 25 and 40. The growth after 40 represents the difference between age 40 and 50.

Figure 1: Life-cycle Profiles of Skill Premiums by Synthetic Cohorts

Note: Author's calculation from IPUMS CPS ASEC, 1964-2019. The skill premium is estimated from Mincer regressions, measuring the percentage wage gap between high-skill and low-skill workers.

Figure 2: Skill Premium Profiles after Controlling for Time Effects



Note: This figure shows estimated age effects, cohort effects and the interaction term by broad category g.

Figure 3: Life-cycle Profiles of Log Human Capital

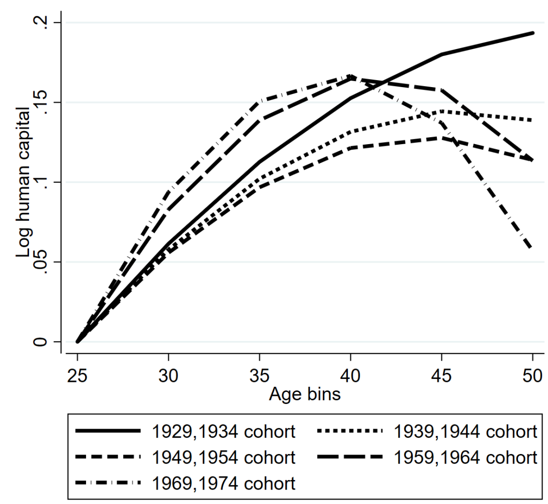
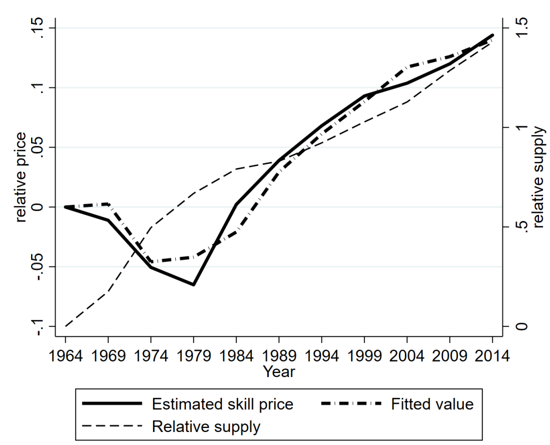


Figure 4: Relative Skill Price



Note: The solid line represents α_t estimated from equation (1.16) and the dashdot line represents fitted values from equation (1.19). The relative supply is constructed based on equation (1.18) and is normalized to 0 in 1964.

Chapter 2

AN EMPIRICAL INVESTIGATION OF TECHNOLOGY USAGE ON EARNINGS

2.1 Introduction

There is a long literature that studies the effects of technology on labor market outcomes like inequality (Burstein *et al.* (2019)) and employment (Acemoglu and Restrepo (2020a)). However, most of the work either focuses on a specific technology (like computers) or treats technology as a broad category. In this paper, I look into technology usage at a granular level and provide new empirical evidence on its impact on individual earnings.

The empirical challenge is to quantify technology usage at the individual level. To overcome this obstacle, I construct an index *distance to the frontier* to approximate information technology usage using occupations as proxy following Gallipoli and Makridis (2018). This index, which is based on the importance of IT-related knowledge, tasks, and skills, measures how far the technology used in one specific occupation is behind the most IT-intensive technology (frontier technology). This index can be interpreted as the relative position in the technology distribution as the frontier technology moves forward.

Using the constructed index, I document technology usage patterns across education and over the life-cycle. I find that technology usage patterns vary significantly by education: college workers are more likely to work with advanced technologies. In addition, there is a considerable gap in technology level between college and non-college workers throughout life-cycle. However, there is barely no changes in technology level over the life-cycle for both educational groups.

Next, I present empirical evidence to show a strong positive relationship between technology usage and earnings. I include the technology index in an otherwise standard Mincer regression and the estimated coefficient on the technology index is positive and statistically significant. In particular, the earnings difference between workers in the 75th percentile of the technology index and the workers in the 25th percentile is 19% after controlling for observables. This correlation even becomes stronger at the occupation level. Furthermore, I find that the observed variation in technology usage accounts for 38% of the growth in life-cycle earnings inequality.

This reduced-form analysis might underestimate the impact of technology because it fails to capture rich interactions between technology usage and human capital. The reason is that I document a strong correlation between technology usage and education: the share of college workers increases in the level of technology. Moreover, there is a considerable gap in technology level between college workers and non-college workers over the life-cycle. These facts suggest that technology choices and human capital investments could be jointly determined even from the beginning of the life-cycle. Therefore technology could generate effects on earnings through the interplay with human capital, which cannot be directly measured by the reduced-form analysis.

The paper is organized as follows. In Section 2.2, I describe how to measure technology usage at the individual level using occupations as a proxy and document technology usage patterns. Section 2.3 investigates the relationship between technology usage and labor earnings at different levels. Section 2.4 concludes.

2.2 Technology Usage Measurement and Patterns

In this section, I investigate technology usage behavior over the life-cycle and across educational groups through a novel measurement of technology usage. The measurement is based on how intensively workers use information technology at the occupational level. I find there is a significant gap in average technology level across educational groups over the life-cycle. In addition, the average technology level conditional on education does not fluctuate much over the life-cycle.

2.2.1 Measurement of Technology

The empirical challenge to study technology usage patterns is the lack of a direct measure at the individual level. To overcome this obstacle, I construct an index *distance to the frontier* to approximate technology usage using occupations as the proxy based on Gallipoli and Makridis (2018). The index is based on how intensively people use information technologies in daily work. The rationale behind this measure is inspired by well-documented facts that information technologies can greatly improve productivity at different levels.[1]

This index measures how far one technology (occupation) is behind the frontier technology, i.e., the most advanced technology. Since the frontier technology is evolving over time, this index can be interpreted as the relative position in the moving technology distribution.

I draw detailed information from Occupational Information Network (O*NET) data set on how intensively workers use information technologies. The O*NET is a comprehensive database of worker attributes and job characteristics. The survey interviews a random sample of workers in each occupation. Interviewees answer questions on a scale from 1 ("not important") to 6 ("extremely important") that measures the importance of some specific knowledge, tasks, or skills. A large literature has used the O*NET database to analyze the labor market outcomes using the task approach

[1] Stiroh (2002) shows that the usage of information technology improves productivity at the industry level. Bloom *et al.* (2012) shows a similar result at the firm level. Akerman *et al.* (2015) find that the adoption of broadband internet improves the productivity of skilled workers.

(See Autor *et al.* (2003) and David and Dorn (2013)).

I construct the index *distance to the frontier* by extracting values of characteristics related to IT technology. Specifically, I consider a set of knowledge, tasks and skills associated with IT technology and sum up the levels of importance (from 1 to 6). After that, I normalize the sums of all occuaptions to the interval $[-1, 0]$. The details of the construction are shown in Appendix B.

This index, as implied by its name, describes how far the technology used in one specific occupation is behind the frontier technology. By construction, the occupation that requires the most intensive IT activities is considered to be the frontier technology and its distance to the frontier is 0. Table 6 shows a sample of representative occupations and their distances in each distance quintile. For instance, janitors are the most common occupation in the first distance quintile (bottom of the technology distribution) and computer scientists are the most common occupation in the 5th distance quintile.

I assume the index is time-invariant over the period of the analysis, i.e. the distance of an occupation relative to the frontier is fixed even though the frontier technology is evolving over time. For instance, consider an occupation with the task of inputting and editing text. Workers used IBM MT/ST, a stand-alone word processing device, in the 1970s and switched to computer softwares like WordPerfect or Microsoft Word in the 1990s. Since both technologies were up-to-date at their time, the relative distance of this occupation does not change. Meanwhile, the absolute level of technology increased over time because computer softwares are more efficient than typewriters.

To justify this assumption, I provide empirical evidence to show that there are no significant changes in task intensity and skill composition so this measurement is robust over time. The O*NET data set is only available from 2003 so I use the information from the fourth edition of the Dictionary of Occupational Titles (DOT) conducted in 1977, which is the predecessor of the O*NET, to check how IT-related task intensity changes across time. I construct an index based on a similar combination of skills and tasks for each occupation in the DOT and compare it with the indices from the O*NET in 2003 and 2021 separately.

Standard OLS regressions indicate that the technology index in 1977 has strong explanatory power on the index in 2003 as well as in 2021 with corresponding R-squared of 0.62 and 0.63.[2] Figure 5 also shows the scatter plots of indices between different periods. Though there are some occupations become more or less IT intensive over time, the above empirical evidence suggest that the skill composition and task intensity from which I infer relative technology level do not change in general.

2.2.2 Technology Usage Patterns

Utilizing the constructed index, I document technology usage patterns across education and over the life-cycle. I find a huge variation in technology usage by education: the fraction of college workers increases with technology level. In addition, there is a considerable gap in technology level between college and non-college workers throughout life-cycle. However, the life-cycle technology usage profile is relatively stable as the mean technology level barely changes over the life-cycle for both educational groups. Specifically, the change in the mean distance between age 23 and 60 for non-college workers is 0.04 whereas the gap in average technology level across education is around 0.25.

[2]The explanatory power of the index in 2003 on the index in 2021 is even higher, with a R-squared of 0.74.

Data source The analysis draws information from the Current Population Survey (CPS) Annual Social and Economics Supplement (ASEC) over the period 1968-2019. I restrict the sample to full-time full-year male workers with earnings above 50% of the federal minimum wage in that year. Self-employed workers are also excluded.[3] I harmonize occupational codes in both CPS and O*NET to the 2010 SOC code and link the constructed index from the O*NET to the CPS sample.

Technology usage by education The distribution of technology usage varies significantly across educational groups as shown in Figure 6 panel (a). I divide workers into two educational groups: with college degrees and without college degrees. College workers are largely concentrated on the right tail of the distribution whereas non-college workers mainly work with less advanced technologies with a distance of less than -0.6.

Panel (b) shows that the relative share of college workers increases with the technology level. At the bottom of the technology distribution (distance less than -0.8), around 90% of the workers don't have a college degree. For example, the relative share of college workers in janitors (with a distance of -0.95 as shown in Table 6) is around 5%. The share of non-college workers decreases with distance and less than 30% of non-college workers are in the top 10% technologies. The increasing share of college workers suggests there could be a selection mechanism of technology choices based on education.

Life-cycle profiles of technology usage Next, I look at technology usage patterns over the life-cycle. I construct the life-cycle profiles by extracting the age coef-

[3]Similar criteria are applied in the literature on earnings inequality. See Storesletten *et al.* (2004) and Guvenen (2007) for example.

ficients ($\beta_{i,t}^{\text{age}}$) from the following statistical model:

$$y_{j,c,t} = \beta_j^{\text{age}} + \beta_t^{\text{year}} + \beta_c^{\text{cohort}} + \epsilon_{c,j,t} \tag{2.1}$$

where $y_{j,c,t}$ is the statistic of interest from cohort c of age j at time t. Due to the linear relationship between age, year, and cohort ($c = t - j$), it is impossible to identify three terms separately without further assumptions. The common way to deal with this problem is to normalize either the time effects β_t^{year} or the cohort effects β_c^{cohort} to zero and attribute the trend to the other factor.

To control for both age effects and cohort effects, I lump three adjacent cohorts into one aggregate cohort which gives me extra degrees of freedom to identify three terms separately.[4] The implicit assumption of this linear statistical model is that the time effects (or cohort effects) only interact with the age profile through the additively separable form.

Two features stand out from the life-cycle profiles of technology usage by education as shown in Figure 7. First, there is a considerable gap in technology level between college and non-college workers even from the beginning of the life-cycle. Specifically, the mean distance of college workers is 0.27 higher than non-college workers at age 25. This difference is 1.3 times the standard deviation of the distance in the entire sample.

Second, the life-cycle profiles of technology usage are relatively flat, especially for college workers. For non-college workers, the growth of mean distance from age 23 to 60 is 0.04, which is equivalent to 20% of the standard deviation of the distance in the sample. The growth of mean distance is only 0.02 for college workers over the same period. Put differently, the gap in technology level across education is relatively constant throughout life-cycle between college and non-college workers.

[4]The shape of age profiles does not change if I only control for year effects or cohort effects.

One additional caveat regarding the interpretation of life-cycle profiles: the age profile of mean distance represents the relative speed of technology upgrading since the frontier technology grows over time. By construction, the distance remains constant if one sticks to the same occupation over time, which implies that the worker adopts new technology at a pace that is consistent with the growth rate of the entire technology distribution.

I also present the life-cycle profiles of different percentiles in Figure 8 to show that the distribution of technology usage is relatively stable over time. There are two things worth mentioning. First, the level of technology percentiles varies significantly by education. For example, the 90th percentile of technology usage for non-college workers is similar to the 50th percentile for college workers. This observation also confirms that there is a huge variation in technology usage by education.

Second, all age profiles are relatively flat over the life-cycle, which indicates that technology distribution conditional on education is relatively stable. However, one should not interpret that technology usage behavior at the individual level is stable as well because this figure is not informative about switching behaviors. Specifically, a worker in the 90th percentile of technology level at age 25 could switch to the 10th percentile at age 55.

2.3 Technology Usage and Earnings

The observation on technology usage patterns naturally begs the question: how does technology affect earnings? I present empirical evidence to show positive correlations between technology level and earnings at different levels, and quantify the contribution of technology to earnings inequality.

To study the relationship between technology usage and earnings at the individual level, I include the technology index to the Mincer regressions as described below:

$$\ln w_{i,t} = \beta_0 + \beta_1 n_{i,t} + \sum_t \beta_{2,t}\text{year}_t + \beta_3 \text{age}_{i,t} + \beta_4 \text{age}^2_{i,t} + X'_{i,t}\gamma + \epsilon_{i,t} \tag{2.2}$$

where $\ln w_{i,t}$ is log real annual earnings for individual i in year t, $n_{i,t}$ is the distance to the frontier technology constructed at occupational level, and $X_{i,t}$ is the set of control variables, including dummies of race, education, marital status and states.

Table 7 column (2) shows that the estimated coefficient on technology is 0.691 with a standard error of 0.002, which is statistically significant from zero. Since the distance takes value from the interval $[-1, 0]$, the result implies that workers in the frontier technology ($n = 0$) on average earn 69.1% more relative to workers in the least advanced technology ($n = -1$) after controlling for observables.

The comparison between column 1 and 2 indicates that the inclusion of the technology index increases the R^2 of the standard Mincer regression from 0.326 to 0.369 as shown. This result implies that technology usage contributes 4.3 percentage points of the overall variation in earnings. That is, the technology index increases the explanatory power of the standard Mincer regression by 13%.

Since the technology index is constructed at the occupational level, there is a perfect linear relationship between the technology index and occupation. Therefore one might wonder to what extent the variation is accounted for by the technology index instead of occupational fixed effects. In column (3), I replace the technology index with occupation dummies and find that the R^2 increases to 0.410. Compared to R^2 in the first two columns, it implies that the technology index is able to explain almost half of the variation across occupations.

To solve the collinearity problem, I run a two-step regression which allows me to disentangle the effect of technology usage from occupational fixed effects. I first run the Mincer regression with occupational dummies (OCC_j) as shown in Equation

(2.3). The first stage is to extract the occupational fixed effects λ_j. In the second step, I regress the estimated occupational fixed effects λ_j on the technology index n_j to examine to what extent the variation across occupations can be accounted for by the variation in the technology index.

$$\ln w_{i,t} = \beta_0 + \sum_j \lambda_j OCC_j + \sum_t \beta_{2,t}\text{year}_t + \beta_3 \text{age}_{i,t} + \beta_4 \text{age}_{i,t}^2 + X_{i,t}'\gamma + \epsilon_{i,t} \tag{2.3}$$

$$\hat{\lambda}_j = \beta_0' + \beta_1 n_j + \epsilon_j \tag{2.4}$$

Column (4) in Table 7 shows that the positive relationship between technology and earnings is also robust at the occupation level. The effect of technology even becomes stronger as the estimated coefficient on technology increases to 0.777 with a standard error of 0.063. The reason is that some high-paying occupations like managers or lawyers are not at the top of the technology distribution. Such occupations require interpersonal or leadership skills and do not involve a high intensity of technology usage. As a result, the coefficient on technology will be underestimated if not controlling for such skills. The two-step regression helps me to disentangle the impact of technology from other valuable skills of an occupation. Therefore its estimation is higher than the one from the modified Mincer regression.

More importantly, as shown in column (4), the R^2 in the second stage of the two-step regression is 0.473. This number implies that almost half of the variation across occupations ($\hat{\lambda}$) can be explained by the constructed index of technology usage. This is also quantitatively consistent with the comparisons in R^2 from column (1) to column (3). Specifically, the occuaptional fixed effects increases R^2 of the standard Mincer regression from 0.326 to 0.410 and the technology index contributes 4.3 percentage points.

Contribution to life-cycle inequality I conduct a simple accounting exercise to demonstrate how technology usage affects life-cycle earnings inequality. I find that the observed variation in technology usage accounts for 38% of the growth in life-cycle earnings inequality.

To isolate the impact of technology, I construct an alternative measurement of earnings as described below:

$$\ln \tilde{w}_{i,t} = \ln w_{i,t} - \hat{\beta}_1 n_{i,t} \tag{2.5}$$

where $w_{i,t}$ is the observed annual labor earnings for individual i at time t, $n_{i,t}$ represents the distance to the frontier and $\hat{\beta}_1$ is the estimated coefficient of the technology index in Table 7 column 2. I denote $\ln \tilde{w}_{i,t}$ as the *residualized earnings*, which rules out the part of earnings that can be explained by technology usage.

I compare the age profiles of life-cycle earnings inequality between the raw earnings ($\ln w_{i,t}$) and the residualized earnings ($\ln \tilde{w}_{i,t}$). In particular, I utilize the statistical model described in Equation (2.1) and use the variance of log earnings as the metric of inequality. The wedge between these two age profiles of earnings inequality can be understood as the variation accounted for by technology usage.

Figure 9 shows that the growth in life-cycle inequality drops significantly using the residualized earnings, which excludes the part explained by technology. Specifically, the level of raw earnings inequality increases 12.5 log points over the life-cycle but the growth decreases to 7.7 log points when using the alternative measurement of earnings. This means that the observed variation in technology usage directly contributes 38% of the growth in life-cycle inequality.

2.4 Concluding Remarks and Discussion

In this paper, I provide novel empirical evidence to show that technology usage at the individual level is positively correlated with labor earnings. I first document technology usage pattern across education and over the life-cycle. My account exercise also suggests that technology usage accounts for 38% of the growth in life-cycle earnings inequality.

Though my analysis provides a sketch of the relative importance of technology, the effects on life-cycle earnings might be underestimated. The reason is that the reduced-form analysis cannot capture how technology affects earnings through the interaction with human capital. As shown in Figure 6 and Figure 7, there is a positive correlation between technology and education that lasts throughout the life-cycle. These facts suggest that technology usage and human capital could be jointly determined from the very beginning of the life-cycle.

In addition, there are other empirical studies showing that education and technology usage are correlated. For instance, Riddell and Song (2017) find that education increases the probability of technology adoption. Mincer (1989) also provides empirical evidence on how technological change affects human capital adjustment. Therefore one needs a life-cycle model that can explain the joint distribution of technology usage and education to thoroughly quantify the contribution of technology to life-cycle earnings. Further study is needed to explore the interaction between human capital accumulation and technology decision.

Table 6: Examples of Occupation and Distance

Distance quintiles	1	2	3	4	5
Non-college workers					
Occupations	Janitors	Truck drivers	Supervisors of salers	Automotive technicians	Computer scientists
Distances	-0.95	-0.76	-0.57	-0.38	-0.10
College workers					
Occupations	Janitors	Clergies	Managers	Accountants	Computer scientists
Distances	-0.95	-0.65	-0.45	-0.35	-0.10

Note: The table presents occupations with most workers in each quintile of the distance by education.

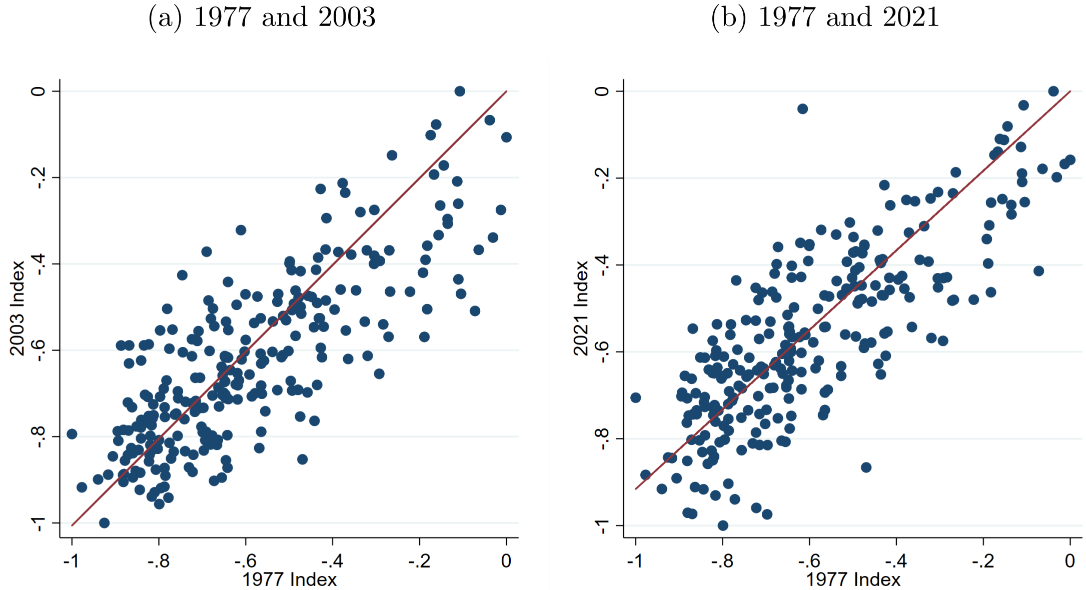
Table 7: Effects of Technology on Earnings

	Mincer regression			Two-step
	(1)	(2)	(3)	(4)
Technology (β_1)	✗	0.691	✗	0.777
		(0.002)		(0.063)
Occupation dummies	✗	✗	✓	✗
N		1262416		442
R^2	0.326	0.369	0.410	0.473

Note: Column (1) presents the estimation of the standard Mincer regression without the technology index. Column (2) shows the estimation of the modified Mincer regression with the technology index. Column (3) includes broad occupational dummies based on (2). Column (4) shows the results of the two-step regression and the R^2 is for the second step regression.

Source: CPS ASEC 1968-2019 and O*NET.

Figure 5: Correlation of Technology Indices over Time



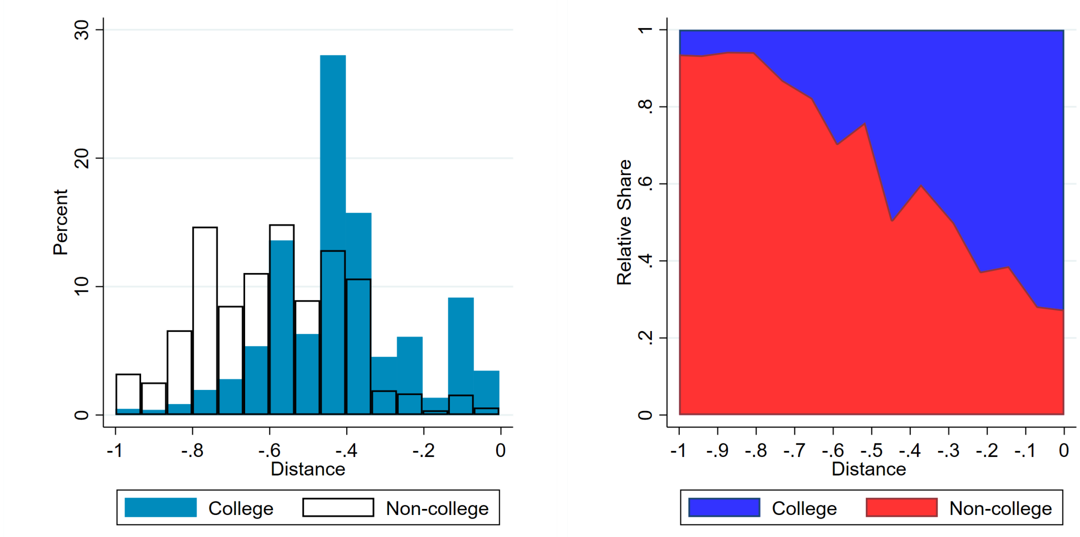
Note: The figure shows the correlation of occupational technology indices across different years.

Source: Author's calculation from the 4th edition of DOT (1977), O*NET 2003 and 2021.

Figure 6: Technology Distribution by Education

(a) Technology distribution by education

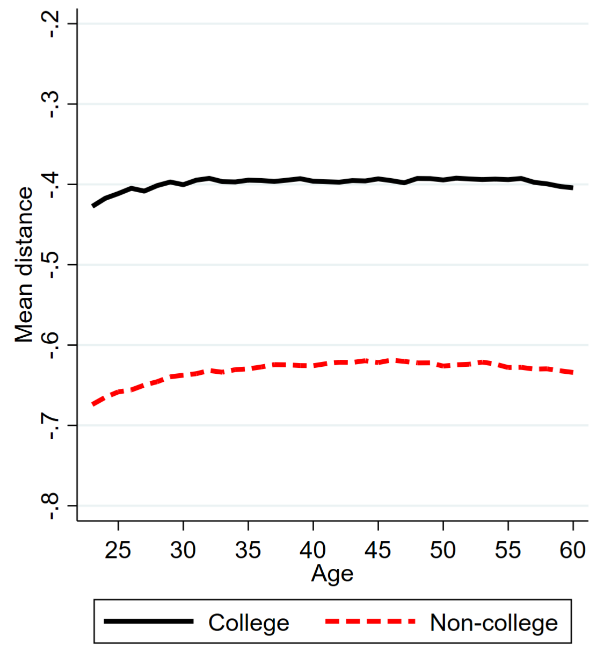
(b) Relative Share by education



Note: Panel (a) shows the distribution of technology usage by educational groups: workers with and without college degrees. Panel (b) shows the relative share of college workers and non-college workers by distance (technology level). The technology distribution is divided by 20 bins and the relative share is calculated in each bin.

Source: Author's calculation from CPS ASEC 1968-2019 and O*NET.

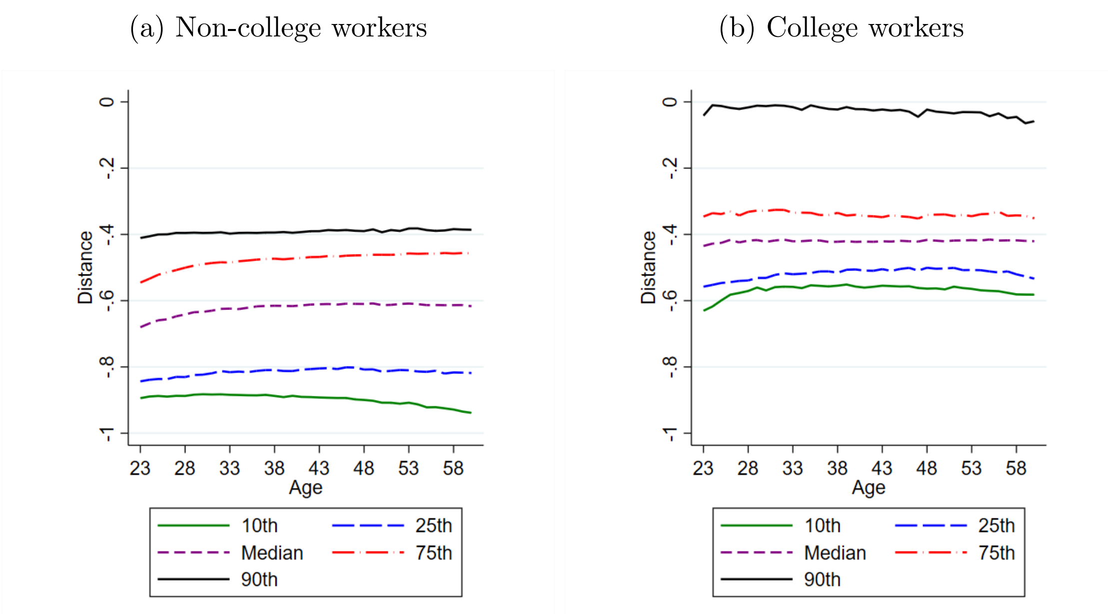
Figure 7: Life-cycle Technology Usage Profiles



Note: This figure presents the life-cycle profile of technology usage using the constructed index *distance to the frontier* measured at the occupation level. A higher distance means a more advanced technology.

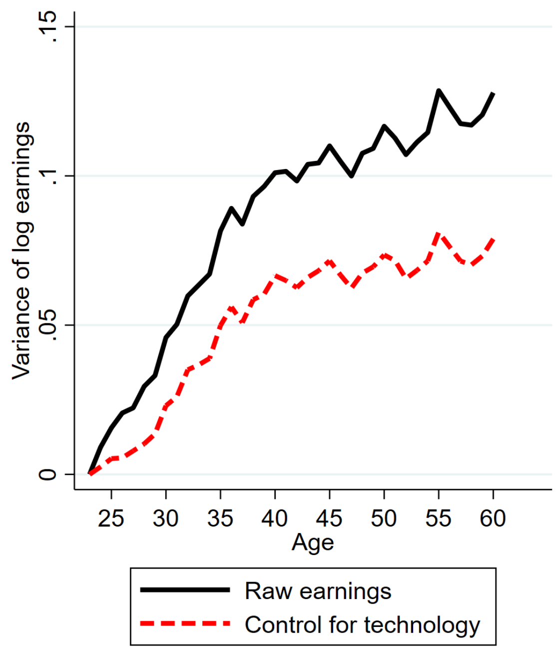
Source: Author's calculation from CPS ASEC 1968-2019 and O*NET.

Figure 8: Technology Profiles by Percentiles



Source: Author's calculation from CPS ASEC 1968-2019 and O*NET.

Figure 9: Life-cycle Earnings Inequality



Note: The figure shows the age profile of variance of log earnings estimated from Equation (2.1). The solid line represents the raw earnings $\ln w_{i,t}$ and the dotted line represents the residualized earnings $\ln \tilde{w}_{i,t}$, which excludes the part explained by technology. Both levels are normalized to 0 at age 23 for comparison purpose.

Source: Author's calculation from CPS ASEC 1968-2019 and O*NET.

Chapter 3

TECHNOLOGY USAGE AND LIFE-CYCLE EARNINGS

3.1 Introduction

Literature has well studied the impact of technology on labor earnings but very few papers have explored this question from the life-cycle perspective.[1] One important margin that arises from the life-cycle perspective is that workers at different life stages could react to the same technological change differently. For instance, when a new technology is developed, young workers would like to exert the effort to learn it while old workers might stick to the old technology as the cost of learning is relatively high. In this paper, I will address two questions: (1) what determines technology usage behavior over the life-cycle, and (2) how does technology usage affect earnings over the life-cycle?

As explored in chapter 2, technology usage at the individual level has significant impact on labor earnings. However, the strong correlation between education and technology usage could result in an underestimation of its impact. In particular, technology could generate ripple effects on life-cycle earnings through the interaction with human capital, which is not captured in the reduced-form analysis.

To quantify the relative importance of technology usage on life-cycle earnings, I develop a life-cycle model with a college decision, technology choices, and human capital investments. Individuals are heterogeneous in initial human capital and the cost of college education which determine their college decisions. College workers accumulate additional human capital at the college stage with the cost of forgoing

[1] Hudomiet and Willis (2021) studies the effect of computerization on near retirement workers.

four years of earnings. During the working stage, individuals maximize utility by choosing which technology to work with and making human capital investments. The model then is parameterized to match life-cycle profiles of technology usage and earnings as well as the college attainment rate.

The novelty of the model is to allow for rich interactions between technology and human capital, which are summarized in three mechanisms. The first mechanism is denoted *direct channel*, in which I assume the earnings function is the product of human capital and technology level. This assumption explicitly leads to the complementarity between these two terms. The second one is the *switching channel* where technology switching comes at a cost of loss in human capital. This assumption is built on Kambourov and Manovskii (2009c) where they find human capital is occupation-specific and partially transferable. Workers can only carry a fraction of human capital when switching to advanced technologies, and the loss of human capital depends on the distance between two technologies.

The last channel is the *catch-up channel*. Since the entire technology distribution is moving forward over time, one needs to learn new knowledge to stay updated with the current technology. I model this cost of learning as the *catch-up cost*, which is increasing with technology level and decreasing with human capital. This mechanism is in the spirit of Galor and Moav (2000) where the time required for learning the new technology diminishes with the level of ability. All three mechanisms will be shown to be important in matching technology usage patterns from the data.

Findings I find that technology usage contributes 31% of the growth in mean earnings and 46% of the growth in variance of log earnings from age 23 to 60. Specifically, the growth in mean earnings drops by 26 percentage points and the growth in life-cycle inequality drops by 5.6 log points after removing technology choice from the model. That is, the model boils down to a risky human capital model.

My model provides two key insights about technology usage. First, the increasing earnings inequality over the life-cycle is largely driven by the interaction between technology and human capital through a reinforcement mechanism. In particular, technology complements human capital through the direct channel. Thus, workers in advanced technologies have more incentives to invest in human capital. Meanwhile, the catch-up channel lowers the barrier of staying with advanced technology for people with high human capital so they are more likely to upgrade technology. To sum up, the model allows for a positive feedback loop between technology and human capital which amplifies earnings inequality over the life-cycle.

I conduct counterfactual experiments to quantify this reinforcement mechanism and its effects on life-cycle earnings. In particular, I shut down the catch-up channel by removing the catch-up cost of technology usage, and the growth in life-cycle inequality reduces 87%. The growth in inequality also decreases when I shut down the direct channel by equalizing the productivities across the entire technology distribution. However, these two channels have opposite effects on the growth of mean earnings over the life-cycle: the catch-up channel depresses earnings growth as it imposes barriers to technology upgrading but the presence of the direct channel slightly boosts the earnings growth. Specifically, shutting down the direct channel decreases the growth in life-cycle earnings by 18%, while the catch-up channel increases the growth in life-cycle earnings by 15%.

The second insight is that technology usage is a crucial determinant behind college attainment, which complements the standard human capital view in Becker (1962). In particular, I find technology provides additional incentives for college education through the interaction with human capital. When shutting down the direct channel, the fraction of college workers drops from 29.8% to 17.8%. Once the complementarity between human capital and technology does not exist, the opportunity of additional human capital accumulation during the college stage becomes less attractive so the college attainment rate declines.

Finally, I conduct a policy experiment to evaluate the role of non-linear taxation on labor earnings. When the progressivity in the economy increases from the U.S. level to the European level under tax neutrality, the growth in mean earnings decreases by 23% and the college attainment rate drops by 7 percentage points. This result confirms the consensus from the non-linear taxation literature that a progressive tax dampens the incentive to accumulate human capital. Guvenen *et al.* (2014), Blandin (2018), Badel *et al.* (2020) and Esfahani (2020) are examples of this line of work. Moreover, I also find a progressive tax suppresses the incentive of technology upgrading, which further reduces earnings growth.

However, the effect of a progressive tax on life-cycle inequality is relatively small compared to the above papers for two reasons. Although a progressive tax distorts the incentive to accumulate human capital and compresses the wage structure, the reinforcement mechanism is slightly strengthened instead. The reason is that a progressive tax has asymmetric second-order effects on technology usage by human capital. In general, all workers experience technology downgrading when switching to a progressive tax but the magnitude of the downgrade is larger for people with low human capital. Consequently, it generates a stronger correlation between human capital and technology, which largely offsets the reduction in earnings inequality brought by a more compressed wage structure.

Related literature To the best of my knowledge, this is the first paper to study technology usage patterns from the life-cycle perspective. Previous studies on individuals' technology choices only focus on a short period or infinite horizon. For example, Chari and Hopenhayn (1991) study technology adoption for agents that only live two periods, and Kredler (2014) extends their work to infinite-horizon. Jovanovic and Nyarko (1996) propose a theoretical framework to study the trade-off between learning by doing and adopting new technologies. My work applies important modeling elements from the above papers in a life-cycle framework. The model also shares similar intuitions with the literature on technology adoption from the firm's perspective, like Parente (1994) and Greenwood and Yorukoglu (1997). Specifically, the incentive of technology upgrading decreases with age as the benefit can only be enjoyed for a shorter period.

My paper broadens the understanding of earnings inequality by unveiling an important mechanism associated with technology. My paper not only incorporates key features from previous work, such as uninsurable earnings shocks and risky human capital accumulation, but also includes technology choices as another source of inequality over the life-cycle. It is closely related to Huggett *et al.* (2011), who find that the difference in initial conditions accounts for the bulk of the variation in earnings inequality. My analysis complements their findings by showing the interaction between technology and human capital as an amplifier of life-cycle inequality.

My work is also closely connected to the literature on occupational mobility. Since the technology index is constructed at the occupational level, technology switching can also be understood as occupational switching. In line with the work of Dillon (2018) and Liu (2019), I find that the opportunity of switching technologies helps mitigate negative earnings shocks. However, instead of focusing on earnings risk in detail, I focus on how technology switching affects earnings inequality, like Kambourov and Manovskii (2009b) and Cubas and Silos (2017). In particular, I conduct my analysis in a life-cycle framework and explicitly emphasize the interplay between technology switching and endogenous human capital investments.

3.2 A Life-cycle Model for Technology Usage

I develop a life-cycle model with a college decision, endogenous technology choice, human capital investments, and incomplete-markets to quantify how technology usage affects life-cycle earnings. The model allows for rich interactions between technology and human capital decisions. I will first ask the model to reproduce technology usage and earnings patterns over the life-cycle for both college and non-college workers and then shut down the technology channel to see what happens to earnings growth and earnings inequality. A tax system is also embedded in the model which allows me to study the role of technology if the economy switches from a proportional tax to a progressive tax.

3.2.1 Environment

Time is discrete. Each period a unit mass of individuals is born who live up to J periods. The population growth rate is μ. Individuals enter the economy with high-school degrees at age 18. They can spend four years in college or enter the labor market directly. During the working stage, they maximize expected lifetime utility by choosing which technology to work with in each period and making human capital investments. They will retire exogenously after age J_R.

I assume workers supply one unit of labor inelastically in each period. Individuals also borrow and save assets at the risk-free rate r to smooth consumption over the life-cycle. The model is in a partial equilibrium where I abstract away from the demand side of technologies and take the growth rate of the technology distribution as exogenous.

Technology and earnings Technology is chosen from the interval $[-1, 0]$ to closely follow the concept of the distance to the frontier in chapter 2. Earnings is a function of technology n, human capital h, productvity z and time t:

$$w = \exp(z) \cdot h \cdot \gamma^{(\eta \cdot n + t)} \tag{3.1}$$

where the component $\gamma^{(\eta \cdot n+t)}$ can be interpreted as the marginal productivity of working with technology n at time t.

The parameter γ stands for the growth rate of the technology distribution. If one stays at the same relative position in the technology distribution from t to $t+1$, his earnings would grow at the rate

$$\gamma = \frac{\exp(z) \cdot h \cdot \gamma^{(\eta \cdot n+t+1)}}{\exp(z) \cdot h \cdot \gamma^{(\eta \cdot n+t)}} \tag{3.2}$$

The parameter η captures the productivity difference within the technology distribution. The earnings ratio between workers in the frontier technology ($n = 0$) and workers in the least advanced technology ($n = -1$) equals γ^{η}. So η rescales the productivity gap for the interval $[-1, 0]$.

Human capital evolution I model human capital evolution in the spirit of Ben-Porath but the set-up is different mainly in two aspects. First, human capital accumulation is uncertain in the sense that the evolution is stochastic. One's investments can only affect the probabilities. Second, the accumulation process is stepwise such that one cannot skip intermediate levels.

Following Jung and Kuhn (2019), I assume the human capital levels are discrete and represented by an evenly spaced ordered set $[h_{min}, ..., h_{max}]$. During the working stage, individuals make human capital investments by choosing the effort $e \in [0, 1]$ which affects the law of motion of human capital evolution. The cost is captured by the disutility term ζe^2.

The evolution of human capital follows a Markov process with probabilities that depend on the effort e, age j, and education $s \in \{\text{College}, \text{Non-College}\}$. In particular, let h^+ (h^-) denotes the immediate successor (predecessor) of human capital level h, the probability that human capital increases to the next level is given by

$$P_s(h_{t+1} = h^+ | h_t = h, e, j) = \rho^{j-22} \cdot p_s \cdot e \tag{3.3}$$

where p_s is the baseline probability that varies by education.[2] Human capital depreciation is modeled by the term ρ^{j-22} with $\rho < 1$. When workers get older, it is less likely to climb up the skill ladder as the baseline probability is multiplied by a factor less than one. The probability that human capital decreases to the previous level is

$$P_s(h_{t+1} = h^- | h_t = h, e, j) = (1 - \rho^{j-22} \cdot p_s \cdot e)\alpha_s^{down} \tag{3.4}$$

where $\alpha_s^{down} \in [0, 1]$ and it is also education-specific. The level of human capital remains the same with probability

$$P_s(h_{t+1} = h | h_t = h, e, j) = (1 - \rho^{j-22} \cdot p_s \cdot e)(1 - \alpha_s^{down}) \tag{3.5}$$

The law of motion of human capital evolution is summarized in the following equation

$$h' = \begin{cases} h^+ & \text{with probability } \rho^{j-22} \cdot p_s \cdot e \\ h & \text{with probability } (1 - \rho^{j-22} \cdot p_s \cdot e)(1 - \alpha_s^{down}) \\ h^- & \text{with probability } (1 - \rho^{j-22} \cdot p_s \cdot e)\alpha_s^{down} \end{cases} \tag{3.6}$$

When the human capital level is h_{min} (h_{max}), the probability of human capital decrease (increase) is absorbed into the probability of staying.

The human capital accumulation process is stepwise. In order to reach the maximum level h_{max}, one needs to experience all its predecessor levels. If a worker falls from the human capital ladder, it would take some time to climb back to the original level. Put it differently, the loss cannot be reimbursed by an excess amount of investments in a short time.

[2]This assumption is to illustrate that the average learning ability is different across education, like in Kong *et al.* (2018). The detailed discussion is postponed to Section 3.3.2.

Cost of switching technologies I assume human capital is technology-specific (Chari and Hopenhayn (1991) and Kambourov and Manovskii (2009c)) and partially transferable (Jovanovic and Nyarko (1996) and Violante (2002)). The knowledge accumulated at old technologies cannot be completely applied in new technologies. The following equation shows the amount of human capital that can be transferred when switching to new technologies:

$$\tilde{h}(n, n', h) = \begin{cases} h & \text{if} \quad n \leq n \\ \left[1 - (n' - n)^2\right] h & \text{if} \quad n' > n \end{cases} \tag{3.7}$$

Equation (3.7) shows the switching cost is asymmetric such that it only occurs when people upgrade technology ($n' > n$). If the worker chooses technology downgrading ($n' \leq n$), he can keep the same human capital level after switching. The downward cost is eliminated to decrease the obstacle of technology downgrading, which is a common phenomenon in the data.

The cost of technology upgrading in terms of the human capital loss is increasing in the distance of the switch ($n' - n$). This functional form is built on the work from Jovanovic and Nyarko (1996) where they provide micro foundations using the Bayesian updating setup.

More experience can be carried to new technologies if they are highly correlated with the old ones. For example, most of the coding skills in Matlab can be directly applied to Python. However, the experience with Excel, a less-advanced technology relative to Matlab, can hardly be helpful to learn Python. The correlation of technology is interpreted as the distance of the switching ($n' - n$). So the loss in human capital is small if two technologies are close.

3.2.2 College Decisions

Workers are endowed with initial human capital h_0 and psychic cost of college education q. Both initial conditions are drawn from two independent log normal distributions:

$$h_0 \sim LN(\mu_{h_0}, \sigma^2_{h_0}) \quad \text{and} \quad q \sim LN(\mu_q, \sigma^2_q) \tag{3.8}$$

Given the combination of h_0 and q, workers are endogenously sorted into college path and non-college path. College workers spend four years to acquire the desired human capital level at the cost of disutility which depends on q then they enter the working stage. Another benefit of college education is that college workers are more likely to work with advanced technologies when entering the labor market relative to non-college workers after graduation. Non-college workers will directly enter the labor market with initial human capital h_0.

Non-college path If the worker does not attend college, he will directly enter the working stage at age 18 with initial human capital h_0. So the value as a non-college (NC) worker is

$$W_{NC}(h_0) = \int_n \int_{z_0} V_{NC}(a_0, h_0, n, z, 18) dF_z(z_0) dF_n^{NC}(n) \tag{3.9}$$

where $V_{NC}(a_0, h_0, n, z, 18)$ is the value as non-college worker at the working stage with asset level a_0, human capital h_0, technology n, productivity z at age 18. The initial productivity is drawn from the distribution $N(\mu_{z_0}, \sigma_{z_0})$ with CDF $F_z(z)$. Workers's initial technology is also determined stochastically and it is drawn from the distribution $F_n^{NC}(n)$.

College path If the worker decides to go to college, he chooses human capital investment x in the college. The production function of human capital is given by

$$h_c(h_0, x) = (h_0 \cdot x)^{\alpha_h} + h_0 \tag{3.10}$$

and the cost of investment is captured by the following disutility term

$$q(x + \mathbb{1}\{x > 0\}) \tag{3.11}$$

This disutility can be understood as the psychic cost of attending college.[3] The worker has to pay (1) the fixed cost of college $q \cdot \mathbb{1}\{x > 0\}$, and (2) the cost that is proportional of the investments $q \cdot x$. Since both terms are increasing in the cost parameter q, it is less costly for people born with lower q to attend the college and acquire human capital investments.

The value of the college education is presented as:

$$W_C(h_0, q) = \max_x \quad -q(x + \mathbb{1}\{x > 0\}) + \beta^4 \int_n \int_{z_0} V_C(a, h_c(h_0, s), n, z, 22) dF_z(z_0) dF_n^C(n) \tag{3.12}$$

Similarly, V_C stands for the value of a college worker at the working stage. This continuation value is discounted by β^4 since it takes four years to complete the college education. For simplicity, I abstract away from the consumption-saving problem during the college stage.

College workers' initial productivity level is drawn from the same distribution $F_z(z)$ as non-college workers. However, their initial technology choice is drawn from a different distribution $F_n^C(n)$ which has first-order stochastic dominance over $F_n^{NC}(n)$. That is, college workers on average work with more advanced technologies. I postpone the discussion of the details to Section 3.3.1.

[3]See Restuccia and Vandenbroucke (2013) for example.

College attainment The lifetime value of a worker with initial human capital h_0 and cost q is described as

$$W(h_0, q) = \max\{W_C(h_0, q), W_{NC}(h_0)\} \tag{3.13}$$

Given the combination of initial conditions, people choose either the college path or the non-college path that generates the highest lifetime value.

The cost of college is to forgo four periods of utility from working stage. The benefit of college education is mainly two-fold. First, workers can directly make human capital investments in the college stage and it is not subject to the stepwise procedure. That is, one with a very low q could accumulate a lot of human capital during college stage. Second, college workers are more likely to work with advanced technologies relative to high-school workers since they are exposed to new technologies in the college stage. This feature accounts for the difference in the initial technology conditions between the two educational groups.

3.2.3 Working Stage

In this subsection, I describe the value functions in the working stage by education types $m \in \{C, NC\}$. In short, both college and non-college workers face same idiosyncratic productivity shocks over the life-cycle. However, the transitions of shocks and human capital are different by education, which I will emphasize later.

Let $V_s(a, h, n, z, j)$ denote the value of a worker at age j working at technology n with education s, human capital level h, asset level a and productivity shock z at the beginning of the period. The value function is

$$V_s(a, h, n, z, j) = \int \max\{V_s^{stay}(a, h, n, z, j), V_s^{move}(a, h, n, \mathbf{Z}, j)\} F(\mathbf{Z}) \tag{3.14}$$

where $V_s^{stay}(a, h, n, z, j)$ denotes the value of staying at the same relative position and

$V_s^{move}(a, h, n, \mathbf{Z}, j)$ is the value of moving to new technologies. $\mathbf{Z}$ stands for the vector of technology-specific productivity shocks.

At the beginning of the period, workers first decide whether to stay with the same technology or move to new technologies. The decision is based upon the realization of the vector of shocks $\mathbf{Z}$ over the technology distribution. That is, the worker will know his productivity z_n if he moves to technology n. Each shock z_n is drawn from the same normal distribution $N(\mu_z, \sigma_z^2)$ independently. This vector of shocks only matters when switching to new technologies and does not affect the value of staying.

The value of staying is described below. If the worker chooses to stay, he will work with technology n this period and collect earnings based on current productivity level z and human capital h.[4] After that, the worker chooses the amount of effort e spent on human capital investments and asset level in the next period a' (or equivalently consumption level c). The taxes are summarized as $T(w, a)$ which I will explain in Section 3.5.

$$\begin{aligned} V_s^{stay}(a, h, n, z, j) = \max_{c, a', e} \quad & u(c) - \phi_s(n, h, j) - \zeta e^2 \\ & + \beta \int \sum_{h_{min}}^{h_{max}} V_s(a', h', n, z', j+1) P_s(h'|h, e, j) dF_s(z'|z) \\ \text{s.t.} \quad & a' + c = (1+r)a + w(h, n, z, j) - T(w, a) \\ & a' \geq \underline{a} \quad \text{and} \quad e \in [0, 1] \end{aligned} \tag{3.15}$$

The worker needs to pay a catch-up cost $\phi_s(n, h, j)$ when staying and this cost comes as the disutility term

$$\phi_s(n, h, j) = \phi_0 (1+n)^{\phi_1} h^{\phi_2} \delta_s^{j-23} \tag{3.16}$$

where $\phi_0, \phi_1 > 0$ and $\phi_2 < 0$. Since the entire technology distribution is progressing over time, staying at the same relative position also means technology upgrading so

[4]Here earnings is a function of age j instead of time t. I implicitly assume the baseline cohort enters the labor market at $t = 0$ so the time index conincides with age j.

he must update his knowledge to operate the new technology. The catch-up cost is also adjusted by a education-specific age factor δ_s to model that the learning cost varies over the life-cycle.

The catch-up cost is increasing in the technology level n and decreasing in human capital level h. That is, it is easier to update the latest knowledge for people with higher levels of human capital. This feature captures the spirit of Galor and Moav (2000) where time required for learning the new technology diminishes with the level of ability. This functional form is also needed to generate the difference in the level of technology between college and non-college workers.

For the continuation value, he will stay at the same relative position n in the next period. His human capital level will evolve stochastically with probability $P_s(h'|h, e)$ as described in Equation (3.6). This is one distinction between college and non-college workers in the working stage since the baseline probability is different.

Another distinction in the value function across education groups is the law of motion of productivity shock. The shock z evolves stochastically according to a mean-reverting AR(1) process as the following

$$z'(z) = \rho_s^z z + \epsilon_s^z \tag{3.17}$$

where $\epsilon_s^z \sim N(0, \sigma_{\epsilon_s}^2)$. So the difference comes from the size of innovation $\sigma_{\epsilon_C}^2$ ($\sigma_{\epsilon_{NC}}^2$) and the persistence of shocks ρ_C^z (ρ_{NC}^z).

This set-up is common in the literature of income process and earnings inequality.[5] In addition, it serves the purpose of increasing occupational mobility especially for technology downgrade. One driver behind technology switching in the model is that the worker draws an extremely good productivity shock for one specific technology. In the absence of this process, people would get stuck with technologies where they

[5] See Guvenen (2009) for a empirical investigation in this topic.

have high productivity levels. Thus workers will not switch to other technologies unless they draw a better productivity shock, which is less likely to happen since the current shock is already good enough.

The value of switching to a new technology is described below:

$$V_s^{move}(a, h, n, \mathbf{Z}, j) = \max_{n' \in [-1,0]} V_s^{stay}(a, \tilde{h}(n', n, h), n', z_{n'}, j) \tag{3.18}$$

where $z_{n'}$ is the technology-specific productivity shock from the vector $\mathbf{Z}$ and $\tilde{h}(n', n, h)$ is the amount of human capital that can be carried to new technology n'. When a worker decides to switch to a new technology n', he will suffer the loss in human capital and then the problem goes back to the "stay" case where he chooses human capital investments and smooths consumption.

The timing of the working stage is summarized in Figure 10. At the beginning of the period, workers first draw the vector of shocks $\mathbf{Z}$ over the technology distribution and then decide to stay or move. If one chooses to stay, he will collect labor income based on current state variables. If he decides to move, he also chooses which technology to work with in this period. Then, his human capital level is determined according to Equation (3.7) and the productivity level is $z_{n'}$.

After collecting labor income, workers choose effort e to invest in human capital, smooth consumption by choosing asset holding tomorrow a', and then enter the next period. The value function is evaluated after the realizations of human capital and shock.

The value function in the last period of working stage is

$$\begin{aligned} V_s^{stay}(a, h, n, z, J_R) = \max_{a'} \quad & u(c) - \phi_s(n, h, J) + \beta V_s^R(a', J_R + 1) \\ \text{s.t.} \quad & a' + c = (1 + r)a + w(h, n, z, J_R) - T(w, a) \end{aligned} \tag{3.19}$$

In the last period of the working stage, workers decide how much to save for the retirement period and do not make any human capital investments. The continuation value V_s^R only depends on savings a' and age.

3.2.4 Retirement Stage

Individuals retire after age J_R and get no labor income. They only live off their accumulated assets plus social security benefits net off taxes. The problem of retirement at age $j > J_R$ is described below:

$$\begin{aligned} V_s^{retire}(a, j) &= \max_{a'} u(c) + \beta V_s^{retire}(a', j+1) \\ \text{s.t.} \quad a' + c &= (1+r)a - T(0, a) + b_s^{ss} \end{aligned} \tag{3.20}$$

Notice that workers in the retirement stage no longer receive labor earnings so the first argument in the tax function is zero. Workers also receive social security benefits after retirement. The benefit is also education-specific and on average college graduates receives more benefit than high-school graduates: $b_C^{ss} = \kappa b_{NC}^{ss}$ with $\kappa > 1$.

3.2.5 Tax System

The tax system $T(w, a)$ in the model consist of two parts: income tax T^{inc} and social security T^{ss}. Individuals' labor earnings and capital income are taxed at a flat rate τ and the social security system taxes labor earnings at the rate τ_{ss} for individuals at the working stage. So the tax function can be presented as

$$T(w, a) = \tau(w + ra) + \tau_{ss} w \tag{3.21}$$

After retirement, agents receive fixed social security benefits b_C^{ss} or b_{NC}^{ss} in each period. The social security system is pay-as-you-go, i.e., it finances the benefits from taxes collected from individuals during the working stage. Government also consumes G for non-productive purpose to balance the budget.[6]

[6]See the formal definition of the stationary equilibrium in the Appendix.

3.2.6 *Sources of Life-cycle Inequality*

The sources of earnings inequality over the life-cycle mainly come from three aspects: human capital (h), technology n, and productivity shocks z. In this subsection, I discuss these three sources and their associated mechanisms, and explain how they affect earnings inequality over the life-cycle.

Interaction between technology and human capital

Technology interacts with human capital mainly in three channels. The first channel is the *direct channel*, i.e., earnings is a function of technology and human capital as shown in Equation (3.1). This set-up explicitly assumes that technology and human capital are complements. As a result, the marginal benefit of human capital investments increases with technology so people in advanced technologies have more incentives to accumulate human capital. This idea dates back to the insight of Schultz (1975) where technological progress complements ability in the formation of human capital. What's more, the incentive of technology upgrading also varies by human capital due to the complementarity.

The *catch-up channel* indicates that the cost of technology usage negatively depends on the level of human capital as described in Equation (3.16). This equation indicates it is easier to stay with advanced technologies for workers with high human capital. Since this cost applies to all workers regardless of switching or not, it also imposes barriers to technology upgrading. To sum up, this catch-up channel lowers the cost of technology usage for people with high human capital.

The last channel is the *switching channel* where the technology upgrading comes with the loss of human capital. Since the switching cost is proportional as shown in Equation (3.7), workers with high levels of human capital will suffer more human capital when switching to better technologies. Thus they are less likely to make a huge step toward frontier technology. This channel works in the opposite direction as the catch-up channel since it discourages people with high human capital to upgrade technology.

The first two channels generate a positive correlation between human capital and technology which amplifies earnings dispersion over the life-cycle. On one hand, the direct channel provides more incentives for human capital investments for workers in advanced technologies. On the other hand, workers with high levels of human capital are more likely to switch to advanced technologies due to the catch-up channel. Consequently, this reinforcement mechanism between human capital and technology will magnify the dispersion in earnings through the interaction between these two components and the correlation will become stronger over the life-cycle. Meanwhile, the switching channel reduces earnings dispersion as it depresses technology upgrading, especially for people with high human capital.

Idiosyncratic shocks

Another important source of inequality comes from idiosyncratic productivity shock z. I follow the standard set-up in the literature to model income risks as an AR(1) process. However, the introduction of technology decisions alleviates the dispersion brought by the shocks. The reason is that the opportunity of switching technologies in each period helps workers mitigate bad shocks.

In the standard AR(1) income process, one might experience a sequence of persistent negative shocks because of bad luck. In my model, due to the presence of technology decisions, one can easily "reset" his productivity level by switching to another technology with high productivity shock so the above scenario will not happen. That is, the opportunity of switching technologies makes shocks less persistent, which lowers the level of dispersion generated by productivity shocks.

3.3 Parameterization and the Benchmark Economy

This section describes how I set the parameters in the model and discusses the properties of the benchmark economy. I first choose a collection of parameters exogenously, either taken from the literature or directly identified from the data. The rest of the parameters are jointly calibrated to match the life-cycle profiles for both college and non-college workers and other statistics, which mostly come from empirical exercises from chapter 2. The parameters are listed in Table 8.

3.3.1 Parameters Chosen from External Source

Demographics The life-cycle starts from age 18 to 75 but I only focus on the life-cycle statistics from age 23 to 60. Individuals retire after age 64 and live another 10 periods. The annual population growth rate is 1.2%, which is the geometric average over the period 1959–2007 from the Economic Report of the President (2008). I assume it is a small open economy where the interest rate is set exogenously to be 0.047 so the after-tax interest rate is 4%.

Tax and social security In the benchmark model, I set the flat tax rate τ on income to be 0.15, which is the approximation of the tax rate in the U.S. once itemizations, deductions and income-contingent benefits are considered. The tax rate of social security on labor earnings is 0.1, which is close to the average rate in the period of analysis.

I assume the social security benefits for college workers are 17% higher than non-college workers:

$$b_C^{ss} = 1.17 b_{NC}^{ss} \tag{3.22}$$

This number is borrowed from Guner *et al.* (2021) where they document how social security benefits vary across household types and educational types.

Technology The Mincer regression with the technology index is used to identify parameters in the earnings function. Taking log of the earnings function in Equation (3.1) generates

$$\ln w \quad = z + \ln h + (\ln \gamma\eta) \cdot n + \ln \gamma \cdot t \tag{3.23}$$

where n is the distance to the frontier and t represents year. Notice that this is analogous to Mincer regression used in the empirical analysis.

Equation (3.23) implies that $\ln \gamma$ corresponds to the coefficient of year in the Mincer regression and $\ln \gamma\eta$ maps to the coefficient of the technology index. Since I use year dummies in the Mincer regression, I further run a linear regression on the estimated year dummies and estimate the annual growth rate of the technology distribution is 0.5%, i.e., $\gamma = 1.005$. That is, if one stays with the same technology over time, all else equal, the natural growth rate of his earnings is 0.5%.

After pinning down γ, the parameter η is identified to match the coefficient of the technology index in the Mincer regression. Setting $\eta = 111$ means that the earnings gap between the most advanced technology ($n = 0$) and the least advanced technology ($n = -1$) is 0.77 in the model, which is consistent with the empirical findings in chapter 2.

Initial distributions of technology I take the initial technology distributions $F_n^{NC}(n)$ and $F_n^{C}(n)$ as exogenous and infer them directly from the data. Specifically, I fit the technology distribution at age 18 (23) with Beta distribution for non-college (college) workers. The advantage of Beta distribution is that it has a limited support

$[0,1]$. After rescaling, it can be mapped to the interval of technology index $[-1,0]$. Figure 11 shows the fitted distributions and the raw distributions from the data.

3.3.2 Parameters Chosen Internally

The rest of the parameters except for the discount factor β are jointly chosen to match (1) the fraction of college workers, (2) life-cycle profiles of mean earnings, mean distance and the variance of log earnings for both college and non-college. I denote the set of 24 parameters as Γ

Formally, the parameterization strategy is to minimize the distance between moments generated by the model and moments from the data. The minimization problem is described below:

$$\min_{\Gamma} \sum_{s=NC,C} \left[\sum_{j=23}^{60} \left((\frac{A^m_{j,s} - A^d_{j,s}}{A^d_{j,s}})^2 + (\frac{B^m_{j,s} - B^d_{j,s}}{B^d_{j,s}})^2 + (\frac{C^m_{j,s} - C^d_{j,s}}{C^d_{j,s}})^2 \right) \right] + (\frac{\omega^s - \omega^d}{\omega^d})^2$$

where $A^m_{j,s}$ is the mean log earnings of workers at age j from $s \in \{C, NC\}$ educational group simulated by the model and $A^d_{j,s}$ is the counterpart from the data. $B^m_{j,s}$ and $C^m_{j,s}$ stand for variance of log earnings and mean distance respectively. ω^m is the fraction of college workers in the model and ω^d is the counterpart from the data.

Lastly, I set the discount factor β to match the ratio between median asset and median labor income. The target ratio is 2.5, which is taken from the Survey of Consumer Finances (SCF) 2013.[7] The discount factor β is chosen to be 0.988 and it generates the ratio between median asset and median labor income of 2.6 in the model.

Human capital process Human capital levels are discrete and represented by an evenly spaced ordered set $[h_{min}, ..., h_{max}]$. The lowest level is normalized to 1 and

[7]Labor income w corresponds to earnings and asset a corresponds to wealth in the SCF .

the highest level is 17.6. I set the number of human capital levels to be 41, which is the same length as the working stage. The rationale is that it would take the whole working stage to climb from the lowest level to the highest level since the accumulation of human capital is stepwise. The rest of the parameters are set to match mean earnings profiles and earnings dispersion profiles.

The parameterized values indicate that college workers have a higher baseline probability of human capital increase. This is in line with the results from the literature on college attainment where they find the average learning ability is higher among college workers. As a result, college workers on average accumulate human capital faster than non-college workers.[8] In addition, the parameter that governs human capital decrease (α^{down}) is also higher for college workers, which is to match the depreciation near retirement since the depreciation rate is the same across education.

Productivity shocks The size of shocks drawn over the technology distribution in each period is $\sigma_z = 0.132$ and this applies to both education groups. The parameterized size of innovation of AR(1) process for college and non-college workers are 0.143 and 0.131 respectively. The persistence parameter for college and non-college workers are 0.95 and 0.92. These values are in the ballpark of the empirical estimation by Guvenen (2009). In addition, the values suggest that college workers experience larger and more persistent shocks relative to non-college workers, which is also supported by findings from Guvenen (2009).

One caveat in interpreting productivity socks is that the realized shocks are the combination of technology decisions and the AR(1) process. As discussed in Section

[8]In Keller (2014) and Kong *et al.* (2018), the learning ability affects the marginal return to effort in the human capital production function. People with high learning ability would be sorted into the college path and they will make more investments during the college stage. As a result, the average learning ability of college workers is higher.

3.6.1, one can easily "reset" his productivity by switching to new technology. In fact, the opportunity of switching technologies can help workers to avoid a sequence of negative shocks.[9] So the realized sequence of shocks is less persistent than the parameters of the AR(1) process suggest.

Catch-up cost The parameters associated with catch-up cost mainly affect technology upgrading. In particular, the parameters ϕ_0 and ϕ_1 determine how fast workers upgrade technology over the life-cycle. The parameter ϕ_2 is the key to generate the technology usage gap between college and non-college workers since the average human capital level is different across education groups.

Initial distribution The initial distributions of human capital h_0 and q are crucial to pin down the college attainment rate. The distribution of q also affects how college workers accumulate human capital during the college stage, and generate the variation in human capital within college workers. Moreover, the college cost q also generates heterogeneity in human capital within college workers.

3.3.3 Understanding Technology Switching

Before showing the model's performance, I first discuss the mechanism of technology switching and how it varies by education and age. In Figure 12, I present kernel density estimation of switching probabilities conditional on workers who switch to other technologies from the simulated economy.[10] For illustration purpose, I only

[9]This view is close to the literature on occupational mobility, e.g., Dillon (2018) and Liu (2019).

[10]I compare the moments related to technology switching between the simulated economy and the data in Appendix C. In general, the model understates the probability of switching relative to the data. The reason is that workers might switch occupations for non-pecuniary reasons in reality, which are not captured in my model.

focus on workers in the 3rd quintile group of the distance ($-0.63 < n < -0.53$).[11]

In general, technology switching is asymmetric such that the distribution is left-skewed, i.e., people are more likely to upgrade technology. The reason is that technology upgrade directly delivers a higher utility as it increases earnings and hence consumption. However, the magnitude of upgrade is smaller compared to downgrades. Figure 12 panel (a) shows that young workers are more likely to upgrade technology compared to old workers. Panel (b) conveys a similar message between college workers and non-college workers but the difference is relatively small.

To better understand the distribution of technology switching, I investigate the key equation:

$$V_s^{move}(a, h, n, \mathbf{Z}, j) = \max_{n' \in [-1,0]} V_s^{stay}(a, \tilde{h}(n', n, h), n', z_{n'}, j) \tag{3.24}$$

This equation governs how far a worker would like to switch (n') given the vector of productivity shocks $\mathbf{Z}$. In Figure 13, I plot $V_s^{stay}(a, \tilde{h}(n', n, h), n', z_{n'}, j)$ as a function of n' and hold productivity shocks $z_{n'}$ constant for all $n' \in [-1, 0]$ for comparison purpose.

The value function is hump-shaped in n'. The value first increases with n' since technology level is positively correlated with earnings. However, two downward forces stop workers from upgrading. First, technology upgrade leads to the loss in human capital that is proportional to the distance of switching $n' - n$ as shown in Equation (3.7). In addition, workers have to pay the catch-up cost $\phi_s(n', h', j)$ in the new technology n'. Moreover, since they suffer human capital loss, it also exacerbates the catch-up cost as it decreases with h'. These two channels together explain why the value function decreases with n' above a certain threshold level. Therefore we see

[11] Though technology switching largely depends on the current technology level, the intuition on switching can also be applied to other quintile groups.

workers prefer a short step of technology upgrade over a long step in Figure 12.[12]

3.3.4 The Benchmark Economy

In this subsection, I examine the quantitative properties in the benchmark economy and compare them with the data counterparts. The parameterized model is able to match targeted life-cycle profiles of earnings and technology usage for both educational groups. In addition, the college attainment rate generated by the model is 29.8%, which is quite close to the average college attainment rate (29.4%) over the period 1968-2019 in the CPS sample.

Figure 14 shows that the model is able to match earnings profiles for both college and non-college workers. In particular, non-college workers' earnings growth over the life-cycle is 60% while the magnitude of growth is about 150% for college workers. College workers on average experience steeper earnings growth because they have a higher baseline probability of human capital increase as shown Table 8. This is the abstraction that college workers on average have higher learning ability relative to non-college workers.

Panel (c) and (d) show that the model generates increasing earnings inequality over the life-cycle for both educational groups. For non-college workers, the growth in life-cycle inequality is minor. The earnings dispersion profile slightly deviates from the data for college workers at the beginning of the life-cycle due to the timing of graduation. In the model, workers who choose the college path will graduate in four years and enter the labor market at age 23 uniformly. In reality, there is a substantial amount of students finishing bachelor degrees in more than four years so the timing

[12]The actual switching behaviors are more complicated because shocks vary across technologies. One may switch to a lower-ranked technology because he draws an extremely good shock z for that technology.

of entering the labor market also varies, which explains the dip in earnings dispersion profile as shown in the data. Other than that, the model is successful in replicating the growth in life-cycle inequality.

Figure 15 presents the model's performance on technology usage. The average distance profiles for both college and non-college workers are within the 95% confidence interval from the data. The model generates hump-shaped mean distance profiles for both college and non-college workers. The intuition is straightforward. At the early stage of the life-cycle, individuals have the incentive to upgrade technology since they can enjoy the benefit for the rest of the life-cycle. When approaching the end of the life-cycle, the cost of technology upgrades outweights the benefit of working with advanced technologies. Consequently, workers gradually stop climbing up the technology ladder as shown in Figure 12 panel (a).

Since I did not match the life-cycle profile of technology dispersion (variance of the distance), I examine untargeted moments for validation: the relative share of college workers over the technology distribution. Figure 16 suggests that the model can replicate the joint distribution of technology usage and education. In particular, the relative share of non-college workers decreases with the technology level. The only unmatched part is that there are fewer non-college workers at the top of the technology distribution.

The decreasing relative share is mainly driven by the catch-up channel. Equation (3.16) suggests that staying at a higher technology position requires more effort and hence leads to higher disutility. Since this catch-up cost decreases with human capital level, it implies that college workers on average face smaller cost as their human capital level is higher. So they are more likely to climb up the technology distribution.

College decisions The college attainment decision is characterized by the combination of initial human capital h_0 and psychic q. Figure 17 shows the college decisions over the joint initial distribution. It is not surprising that people with higher cost q are less likely to attend college since it is directly associated with the disutility term during the college stage as shown in Equation (3.12). Moreover, people with low q would accumulate more human capital.

Given the same level of q, individuals with higher initial human capital are less likely to attend college. The reason is that the time cost of college education exceeds the benefit of human capital investments. If one skips the college stage, he directly enters the labor market and gains earnings based on his initial human capital. If he decides to attend college, he must forgo four periods of the working stage. Even though he could accumulate additional human capital during the college stage, it cannot offset the sacrifice of four periods of earnings.

3.4 Technology and Life-Cycle Earnings

In this section, I first conduct counterfactual experiments to shut down each interaction channels associated with technology separately and evaluate their effects on life-cycle earnings. Then I completely remove the choice of technology usage from the model and quantify its overall impact.

Results show that technology usage accounts for 31% of the growth in mean earnings and 46% of the growth in earnings inequality. Moreover, I find that the model generates a reinforcement mechanism between technology and human capital which amplifies earnings growth and earnings inequality over the life-cycle.

3.4.1 *Catch-up Channel*

The first experiment is to shut down the catch-up channel by reducing catch-up cost. Since the entire technology distribution is moving forward, individuals have to pay catch-up cost $\phi_s(n,h,j)$ (in disutility term) to stay at the same relative position over time with the functional form

$$\phi_s(n,h,j) = \phi_0(1+n)^{\phi_1} h^{\phi_2} \delta_s^{j-23} \tag{3.25}$$

To reduce the catch-up channel by 50%, I set ϕ_0 to be half of the parameter in Table 8. $\phi_0 = 0$ means completely shutting down the catch-up channel, i.e., the disutility term associated with technology usage disappears.

Figure 18 panel (a) suggests that reducing the catch-up channel increases earnings growth over the life-cycle. In particular, as shown in Table 9, the magnitude of earnings growth increases by 27% after shutting down the catch-up channel. The steeper growth is mainly driven by the change in technology usage patterns as shown in panel (c). Without the catch-up cost, workers face fewer barriers when switching to advanced technologies so they climb up the technology ladder at a faster pace. Consequently, the mean distance profile keeps increasing over the life-cycle even near retirement. Since technology level is positively associated with earnings, this leads to steeper earnings growth over the life-cycle.

Panel (b) in Figure 18 suggests that turning down the catch-up channel greatly reduces the growth in life-cycle inequality and the quantitative evaluation is presented in Table 9. In the benchmark economy, the earnings inequality keeps increasing over the life-cycle and it is accompanied by a stronger correlation between technology and human capital as shown in panel (d). This observation confirms the reinforcement mechanism discussed in Section 3.6.1 where workers with high human capital are more likely to work with advanced technologies and vice versa. Therefore the increasing correlation amplifies the earnings dispersion over the life-cycle through the positive feedback loop.[13]

Reducing the catch-up channel weakens the influence of human capital on technology, which undermines the reinforcement mechanism and hence lowers the growth in earnings inequality. In the benchmark economy, the catch-up cost decreases with human capital so it is easier to upgrade technology for people with high human capital. So workers would be more stratified in the technology distribution on the basis of human capital. Once the catch-up cost is removed, human capital will not facilitate technology upgrading so there will be more people with low human capital switching to advanced technologies. Indeed, panel (d) shows that the correlation between technology and human capital is almost zero when the catch-up channel is reduced by 50%. The correlation even becomes negative after shutting down the catch-up channel.[14] This suggests that the amplification mechanism is weakened and therefore the growth in life-cycle inequality decreases.

College decisions Table 9 shows that the college attainment rate drops 12.7 percentage points after shutting down the catch-up channel. To further understand the change in the attainment rate, Figure 19 compares college decisions between the benchmark model and the catch-up channel experiment. The black dots denote individuals who will go to college in the benchmark case ($\phi_0 = 3.1$) but decide not to attend college after shutting down the catch-up channel ($\phi_0 = 0$). In general, the

[13]In Figure 29, I also show that the changes in life-cycle inequality is not driven by the compositions effect, i.e. the change in the college attainment rate. The life-cycle inequality conditional on each educational groups decreases when shutting down the catch-up channel.

[14]Due to the switching channel, people with high human capital are less likely to switch since the loss in human capital is proportional. Therefore it forms a negative correlation between human capital and technology.

threshold levels of cost q for college education decreases, especially for people with low human capital.

When catch-up cost is eliminated, people value human capital less because it is not beneficial for technology upgrading as discussed above. As a result, college education becomes less attractive and the college attainment rate drops.

Moreover, the decline in the threshold level of q becomes larger for people with low initial human capital. This is because people with high human capital are less likely to be subject to the catch-up cost when upgrading technologies in the benchmark economy. On the contrary, people born with low initial human capital are more likely to be deterred from upgrading because they cannot afford the catch-up cost due to low human capital. Therefore people with low human capital would like to attend college to accumulate additional human capital even though their cost q is relatively high.

Once the catch-up cost is eliminated, people with low human capital will face no barriers of technology upgrading so they can directly enter the labor market and climb up the technology ladder. Consequently, only people with extremely low cost q would like to attend college as they can accumulate a huge amount of human capital.

3.4.2 Direct Channel

The direct channel means that earnings function is the product of technology level n and human capital h as described below

$$w = \exp(z) \cdot h \cdot \gamma^{(\eta \cdot n + t)} \tag{3.26}$$

This functional form explicitly generates a complementarity between human capital and technology.

The parameter η governs the productivity difference within the technology distribution. To reduce the direct channel by 50%, I set η to be half of the calibrated value, which means that the earnings gap between the frontier technology and the least advanced technology shrinks 50%. Similarly, I shut down the direct channel by setting $\eta = 0$. In this extreme case, all technologies have the same productivity level as the frontier technology ($n = 0$). This also implies that technology does not complement human capital.

One caveat with the experiment of the direct channel is that lowering the parameter η also increases the level of earnings for people who do not use the frontier technology. This income effect might affect technology and human capital decisions at the aggregate level. To control for this possible channel, I multiply earnings function by a factor less than one such that the mean earnings at age 23 in each counterfactual is the same as the benchmark economy.

Figure 20 shows that shutting down the direct channel reduces the growth of life-cycle inequality and it also flattens the mean earnings profile. Specifically, Table 9 shows that the growth in mean earnings over the life-cycle decreases from 84% to 51%. Moreover, the growth in life-cycle inequality decreases by 6.8 log points. In Figure 30, I also show that the changes in life-cycle earnings profiles at the aggregate level is not mainly driven by the composition effect.

The intuition of flattened earnings inequality profile is similar to the experiment of the catch-up channel, i.e., the reduction in η also undermines the reinforcement mechanism. Specifically, shutting down the direct channel first eliminates the dispersion brought by technology usage and then compresses earnings dispersion through the complementarity term. Moreover, it closes the channel from technology to human capital since the incentive of human capital accumulation will not depend on the technology level now.

College decisions Figure 21 compares the college decisions between the benchmark model and the direct channel experiment. The black dots denote individuals who will go to college in the benchmark case ($\eta = 111$) but decide not to attend college after shutting down the direct channel ($\eta = 0$). In general, the threshold levels of initial human capital and cost q for college education both decreases, which implies college education is less attractive once technology has less impact on earnings.

The reduction in η affects the value of college education mainly in two aspects. First, as the data suggested in Figure 11, college workers on average work with better technologies relative to non-college workers at the beginning of the life-cycle. The reduction in η weakens this initial advantage in technology because now people have higher earnings at the lower part of the technology distribution, which directly decreases the benefit of college education.

Second, since the earnings gap across technologies shrinks, the importance of the interaction between technology and human capital also decreases. As a result, workers have less incentive to accumulate human capital so more people would skip the college stage and enter the labor market directly.

3.4.3 Switching Channel

The last interaction channel is the switching channel, where workers suffer human capital loss when switching to better technologies as shown in Equation (3.7). As shown in Table 9, shutting down switching channel increases the mean earnings growth to 81% and the growth in life-cycle inequality to 15.5 log points.

Once the switching cost is removed, workers would upgrade technology more frequently so they experience steeper earnings growth over the life-cycle. In addition, the reinforcement mechanism becomes stronger so the life-cycle inequality also increases. The intuition is the following. Since the loss in human capital is proportional to human capital, people with high human capital are less likely to make a huge step of technology upgrading. Once this barrier is removed, they would upgrade technology more intensively so the correlation between technology usage and human capital becomes stronger, which leads to a higher level of inequality and a steeper growth in life-cycle inequality.

The college attainment rate does not change significantly because the switching cost is proportional to human capital so it is not in favor of any specific educational groups.

3.4.4 Initial Advantage

The above experiment indicates that technology plays an important role in determining college decisions. In this subsection, I disentangle the impact of technology on college decisions and conclude that the initial advantage in technology distribution is the key determinant. Once this advantage is eliminated, the college attainment rate drops from 29.8% to 22.5%.

The empirical analysis shows that college workers on average work with better technologies relative to non-college workers even at the beginning of the life-cycle and it is modeled as the difference in initial technology distributions presented in Figure 11. I shut down this channel by assuming that college workers also draw initial technology choices from the same distribution as non-college workers.

The last row in Table 9 shows that the elimination of the initial advantage greatly reduces the college attainment rate. Because of the composition effect[15], the magnitude of earnings growth decreases and the earnings inequality profile decreases over the life-cycle at the aggregate level as presented in Figure 22.

[15] From Figure 14, we know that non-college workers have flatter mean earnings profile and earnings inequality profile.

Figure 23 shows the change in college decisions over the joint distribution of initial conditions. Again, the black dots represent people who go to college in the benchmark case but decide not to attend college once the initial advantage is eliminated.

The initial advantage largely benefits people born with high levels of human capital and provides additional incentives for college education for them. In the experiment, the threshold level of cost q increases for people with higher levels of h_0. Since the cost of college education only depends on the parameter q, this implies that the benefit of college actually decreases for those people once the initial advantage is eliminated.

On the contrary, the threshold level does not change for workers with low levels of initial human capital. Put it differently, the initial advantage is not the key determinant of going to college for them. Instead, the additional human capital accumulation during the college stage is the main reason.

3.4.5 All Together

Lastly, I turn down all interaction channels associated with technology and evaluate how life-cycle earnings change. In particular, I do not allow workers to switch technologies and shut down the catch-up cost associated with technology usage. Besides, I shut down the direct channel by equalizing productivity levels across all technologies such that the mean earnings at age 23 is the same as the benchmark economy.

The model boils down to a risky human capital investments model where life-cycle earnings are only determined by endogenous human capital investments (at college and during the working stage) and idiosyncratic shocks. The difference between life-cycle earnings profiles can be interpreted as the contribution of technology usage.

As shown in Figure 24 and Table 9, after removing technology usage, the growth in mean earnings decreases by 26 percentage points (31%). In addition, the growth in earnings inequality decreases by 5.7 log points (46%) over the life-cycle, which is larger than the number (38%) obtained from the reduced-form analysis in Section 2.3. Moreover, the fraction of college workers drops from 29.8% to 18.2%.

One caveat with the final result is that it is not additive because each experiment might be intertwined with other channels. For example, shutting down the direct channel also implicitly assumes that the initial advantage is eliminated. In addition, the effect of the catch-up channel on life-cycle inequality is much larger than the overall impact. This is because the catch-up experiment does not isolate the effects of the switching channel, which in turn increases the growth in life-cycle inequality.

3.5 Policy Analysis: Non-linear Taxation

In this section, I evaluate the effects of non-linear taxation on life-cycle earnings. In particular, I replace the proportional tax on labor earnings in the benchmark economy with a progressive tax. That is, the marginal tax rate and the average tax rate of labor earnings increases with labor earnings. I reparameterize the model with the progressivity level in the U.S. to match the moments as discussed in Section 3.3.2, and then explore how different levels of progressivity affect life-cycle earnings.[16]

The policy experiments show that a progressive tax reduces the college attainment rate, and lowers mean earnings and earnings growth over the life-cycle. However, the effects on life-cycle inequality are relatively small, which is contrary to the recent view in the literature that progressive taxation compresses the wage structure and hence decreases earnings inequality, like Erosa and Koreshkova (2007) and Guvenen *et al.* (2014). The reason is that a progressive tax has second-order effects on technology usage through the catch-up channel, which instead strengthens the reinforcement mechanism. As a result, it leads to a slight increase in earnings inequality, which

[16] The quantitative analysis in Section 5 is robust under the progressive tax at the U.S. level.

partially offsets the reduction in earnings inequality brought by a compressed wage structure.

3.5.1 Progressive Tax System

In the baseline model, individuals' labor earnings and capital income are taxed at a flat rate τ. I now replace the proportional tax rate on labor earnings with progressive taxes and leave the tax rate on capital income unchanged.

I borrow the progressive tax system pioneered by Feldstein (1969) and later popularized by Benabou (2002). In particular, the average tax rate on labor earnings is given by

$$\tau(w) = 1 - \lambda(w/\bar{w})^{-\tau_p} \tag{3.27}$$

where $\bar{w}$ is the mean labor earnings in the economy. The average tax rate of the individual with mean labor earnings is $1-\lambda$. This tax rate increases with labor earnings w in a concave pattern since $\tau_p > 0$.

The parameter λ controls for the level of the tax rate and the parameter τ_p stands for the progressivity in the tax schedule. In the case of $\tau_p = 0$, the average tax rate will not depend on labor income, i.e., it boils down to the standard proportional tax.

3.5.2 Tax Progressivity and Earnings over the Life-cycle

I conduct policy experiments to explore how progressivity affects earnings over the life-cycle. The results in Table 10 indicate that a more progressive tax system leads to a lower college attainment rate and a smaller earnings growth over the life-cycle. However, the effects on life-cycle inequality are relatively small compared to the literature.

Recall that the parameter τ_p in Equation (3.27) governs the progressivity of the non-linear tax system. A higher τ_p means the tax system is more progressive. In the following counterfactual analysis, I fix the total amount of taxes collected by the government by adjusting the tax rate λ accordingly. The benchmark economy is reparameterized to the progressivity level in the U.S. ($\tau_p = 0.05$) following the estimation from Guner *et al.* (2014), who use data on federal tax returns in 2000.[17]

I consider three counterfactual scenarios of progressivity as shown in Table 10. The first scenario is $\tau_p = 0.10$, a number estimated by Heathcote *et al.* (2020) where they additionally include federal government transfers alongside taxes. The second scenario with $\tau_p = 0.15$ stands for the level of progressivity in European countries, like U.K. or Germany.[18] Lastly, I evaluate the effects when the economy switches from a progressive tax to a proportional tax ($\tau_p = 0$).

Earnings growth

Figure 25 panel (a) shows that the mean earnings profile becomes flatter as the tax system becomes more progressive. This result is consistent with the common view in the literature that progressive taxes distort the incentive to accumulate human capital.[19] Since the marginal tax rate increases with earnings, the marginal benefit of human capital investments decreases as a larger fraction of income would be taxed. So it suppresses the human capital accumulation over the life-cycle. This is confirmed by the observation in panel (b) where the mean human capital profile becomes flatter with more progressive taxes.

In addition to human capital, the progressive taxes also suppress the incentive of technology upgrading, and intuition is the same as the argument for human capital accumulation. Panel (c) suggests that the mean distance profile shifts downward

[17] I only change the parameters related to the catch-up cost, human capital probabilities and the initial distribution of h_0 and q.

[18] See Heathcote *et al.* (2020) table 2 for details.

[19] See Guvenen *et al.* (2014), Krueger and Ludwig (2016), and Badel *et al.* (2020) for example.

when taxes become more progressive, which implies that people on average use less advanced technologies over the life-cycle. In particular, the average distance drops more than -0.05 at age 60 when switching from $\tau_p = 0.05$ to $\tau_p = 0.15$. The magnitude is equivalent to 0.4 times of the standard deviation of the distance at age 23. Since earnings are a function of human capital and technology, panel (b) and (c) in Figure 25 together imply a flatter growth in life-cycle earnings if the tax system becomes more progressive.

One potential reason behind the flattening of earnings profile is the composition effect, i.e., the decline in the college attainment rate. Since non-college workers have a flatter mean earnings profile, the drop in the college attainment rate naturally leads to a flatter earnings profile at the aggregate level. To rule out this possibility, I also look at the life-cycle profiles for both college and non-college workers respectively and the results show that the progressive taxes do disincentivize human capital accumulation and technology upgrading for both educational groups.

As presented in Figure 31 and Figure 32 panel (a), the growth in life-cycle earnings decreases with the progressivity for both educational groups. Panel (b) and (c) show that workers have less incentive to accumulate human capital and upgrade technology when facing a more progressive tax regardless of education. Therefore, the flattening of the earnings profile is not solely driven by the change in the college attainment rate.

Table 11 shows that a progressive tax also reduces mean labor earnings and income in the economy. In particular, the mean labor earnings drops 6% when the economy switches to a progressive tax at European levels ($\tau_p = 0.15$). The decline in mean income is around 5%, which is smaller than the change in labor earnings. The reason is that the progressive tax is on labor earnings but not on the return to savings so the incentive of saving is not distorted. In addition, the mean labor earnings increases

2.8% if the economy switches to a proportional tax given the total amount of taxes fixed.

Earnings inequality

The last column in Table 10 suggests that the growth in earnings inequality is also affected by a more progressive tax but the magnitude is smaller. This result is contrary to the recent findings in the literature that progressive taxes compress wage structure and hence lower earnings inequality like in Guvenen *et al.* (2014) or Badel *et al.* (2020). In particular, Esfahani (2020) finds that increasing the progressivity parameter τ_p from 0.13 to 0.17 reduces the growth in life-cycle inequality around 30% whereas my results suggest the decline is less than 5%.

Why earnings inequality is not critically affected by the progressive tax? The answer is that the reinforcement mechanism that generates increasing earnings inequality is not changed by a progressive tax. In fact, the correlation between technology and human capital even becomes stronger when the tax system is more progressive. Specifically, the average correlation between human capital and technology over the life-cycle is 0.28 when $\tau_p = 0.05$ and it increases to 0.34 when $\tau_p = 0.15$. Indeed, the wage structure is compressed by progressive taxes as earnings profiles become flatter. The progressive tax also strengthens the positive correlation between human capital and technology, which in turn increases the level of inequality. Overall, the first force (compressed wage structure) slightly outweighs the second force (stronger correlation) so the reduction in earnings inequality is not significant.

The reason behind the stronger correlation is that the progressive tax has asymmetric effects on technology usage through the catch-up channel. In particular, a progressive tax depresses technology upgrading and the effects are stronger for workers with low human capital. In Table 12, I present the average distance over the life-cycle by human capital level. Specifically, I divide workers into five groups based on the level of human capital at age 60. The last column documents the changes in the distance when switching from the benchmark economy to the progressive tax. For the first human capital quintile group, the decline in the average distance is 0.05 whereas the change in the last human capital quintile is -0.03. In short, the drop in the distance is larger for people with low human capital. Put it differently, the progressive tax is in favor of workers with high human capital in terms of technology usage. As a result, workers would be more stratified in the technology distribution on the basis of human capital so the correlation becomes stronger.

This asymmetric effect is driven by the fact that the catch-up cost is a decreasing and convex function of human capital as shown in Equation (3.16) with the calibrated parameter $\phi_2 = -1.3$. As the progressive tax suppresses human capital accumulation, it also increases the catch-up cost of technology usage. Since the cost is convex and decreasing in human capital, the increase in the cost is larger for people with low human capital, which makes it much harder to upgrade technology relative to people with high human capital. Therefore we see workers from the first two quintile groups experience a larger technology downgrade when switching to a progressive tax.

College decisions

The progressive tax also suppresses the incentive to attend college as shown in the first column of Table 10. The college attainment rate drops from 29.8% to 22.5% when τ_p increases to 0.15. This result is also qualitatively consistent with findings from Esfahani (2020). In short, the progressive tax discourages human capital accumulation, which also includes investments during the college stage.

To better understand the patterns in college attainment, I present the change in college decisions over the joint initial distributions when the economy switches to a progressive tax at European level ($\tau_p = 0.15$) in Figure 26. The black dots denote workers who would go to college when the progressive tax is at the U.S. level but decide to skip the college stage when the progressive increases to the European level.

As shown in Figure 26, the threshold level of cost q for college education decreases after the tax becomes more progressive. The reason is the following. A progressive tax distorts the incentive of human capital accumulation even during the college stage, which further lowers the value of the college stage. Since the cost of education only depends on the cost parameter q, given the same initial human capital condition h_0, one needs a lower initial condition on cost q to attend college. Put differently, a more progressive tax makes college education less profitable from the life-cycle view so the fraction of college attainment rate declines.

3.6 Final Remarks

In this paper, I thoroughly quantify the contribution of technology to earnings through the lens of a life-cycle model with a college decision, endogenous technology usage, and human capital investments. The novelty of the model is to allow for rich interactions between human capital and technology. In particular, human capital facilitates technology upgrading through the *catch-up channel*. The *direct channel* makes human capital accumulation investments contingent on technology as it leads to the complementarity between these two factors in earnings. Moreover, the *switching channel* captures the barrier to technology upgrading in terms of the loss of human capital.

My model suggests that technology usage accounts for 31% of the growth in mean earnings and 46% of the growth in earnings inequality over the life-cycle. Furthermore, counterfactual experiments suggest that both catch-up channel and direct channel are crucial in generating increasing earnings inequality over the life-cycle. Specifically, these two channels build up a reinforcement mechanism between technology and human capital where workers with high human capital are more likely to work with advanced technologies and vice versa. The interaction between these two terms amplifies the earnings dispersion over the life-cycle.

Furthermore, a progressive tax on labor earnings has relatively small effects on life-cycle inequality, which is contrary to the recent findings from the literature. Though the progressive tax schedule compresses the wage structure by distorting the incentive to accumulate human capital, it also slightly strengthens the reinforcement mechanism between technology and human capital, which offsets the reduction in life-cycle inequality.

Table 8: Parameterization

Category	Meaning	Parameter
Externally chosen parameters		
Demographic	population growth rate	$\mu = 0.0012$
	rate of return on asset	$r = 0.047$
	life expectancy and retirement age	$J = 75$, $J_R = 64$
Tax	proportional tax rates on income	$\tau = 0.15$, $\tau_{ss} = 0.1$
Technology	growth rate of the technology distribution	$\gamma = 1.005$
	productivity difference within the technology distribution	$\eta = 111$
Initial distribution of tech	approximated by Beta distribution from the data	
Internally chosen parameters		
Preference	discount factor	$\beta = 0.988$
	disutility of human capital investments	$\xi = 0.25$
Human capital	human capital grid	$h_{min} = 1$, $h_{max} = 17.6$
	baseline probability of human capital increase	$p_C = 0.35$, $p_{NC} = 0.23$
	human capital decrease parameter	$\alpha_C^{down} = 0.15$, $\alpha_{NC}^{down} = 0.07$
	depreciation	$\rho = 0.99$
	human capital production at college stage	$\alpha_h = 0.35$
Productivity shocks	size of innovation	$\sigma_z = 0.132$, $\sigma_C^{\epsilon} = 0.143$, $\sigma_{NC}^{\epsilon} = 0.131$
	persistence of shocks	$\rho_C^z = 0.95$, $\rho_{NC}^z = 0.92$
Catch-up cost	disutility associated with technology usage	$\phi_0 = 3.1$, $\phi_1 = 1.5$, $\phi_2 = -1.3$
	age adjustment in disutility	$\delta_C = 0.994$, $\delta_{NC} = 0.999$
Initial distributions	initial human capital h_0	$\mu_{h_0} = 1.01$, $\sigma_{h_0} = 0.1$
	psychic cost of college q	$\mu_q = 2.91$, $\sigma_q = 0.5$
	initial productivity z	$\mu_{z_0} = 0$, $\sigma_{z_0} = 0.025$

Note: This table presents parameters used in the benchmark economy. The first set of parameters is chosen from external sources. The second set of parameters is jointly determined to match the life-cycle profiles of mean earnings, variance of log earnings and mean distance for both college and non-college workers as well as the average college attainment rate.

Table 9: Life-Cycle Earnings under Counterfactual Experiments

	% of college workers	Mean earnings growth	Growth in life-cycle inequality (log points)
Benchmark	29.8	1.84	12.3
Catch-up channel			
reduced by 50%	26.2	2.03	7.2
reduced by 100%	17.1	2.11	1.5
Direct channel			
reduced by 50%	21.8	1.58	7.4
reduced by 100%	17.8	1.51	5.5
Switching channel			
reduced by 100%	30.3	1.81	15.5
Eliminate the initial advantage	22.5	1.85	7.8
Remove technology usage	18.2	1.58	6.6

Note: Column 1 shows the fraction of people attending college in each scenario. Column 2 shows the aggregate mean earnings growth between age 60 and 23. The last column shows the change in the variance of log earnings, measured as log points, between age 23 and 60. To reduce the catch-up (direct) channel by 50%, I set ϕ_0 (η) to be 50% of the original level. ϕ_0 or η is set to 0 to completely shut down each channel. To remove technology usage, I do not allow workers to switch technologies and shut down all interaction channels.

Table 10: How Progressivity Affects Life-Cycle Earnings

	% of college workers	Mean earnings growth	Growth in life-cycle inequality (log points)
Progressivity tax			
$\tau_p = 0.05$ (Benchmark)	29.8	1.84	12.3
$\tau_p = 0.10$	25.9	1.74	12.0
$\tau_p = 0.15$	22.5	1.65	11.5
Proportional tax: $\tau_p = 0$	33.3	1.91	12.5

Note: This table presents how earnings change with respect to the progressivity (τ_p) of the tax schedule. The benchmark model is parameterized with $\tau_p = 0.05$. Column 1 shows the fraction of people attending college in each scenario. Column 2 shows the mean earnings growth from age 23 to 60 at the aggregate level. The last column shows the change in the variance of log earnings, measured as log points, between age 23 and 60. Total taxes collected by the government are constant in each scenario.

Table 11: How Progressivity Affects Aggregate Earnings

	Mean labor earnings	Mean income
Progressive tax		
$\tau_p = 0.05$ (Benchmark)	100	100
$\tau_p = 0.10$	97.0	97.6
$\tau_p = 0.15$	94.0	95.2
Proportional tax: $\tau_p = 0$	102.8	102.1

Note: This table presents how mean labor earnings and mean income (labor earnings and return on capital) change with respect to the progressivity (τ_p) of the tax schedule under the tax neutrality condition. I normalize the mean labor earnings and mean income to 100 in the benchmark economy. Total taxes collected by the government are fixed in each scenario.

Table 12: Change in Technology Usage by Human Capital Quintile

HC quantile	Average distance over the life-cycle		
	$\tau_p = 0.05$	$\tau_p = 0.15$	change
1	-0.66	-0.71	-0.05
2	-0.55	-0.62	-0.07
3	-0.53	-0.58	-0.05
4	-0.52	-0.55	-0.03
5	-0.50	-0.53	-0.03

Note: This table presents the average distance over the life-cycle by human capital quintile. I divide all workers into five groups based on the level of human capital at age 60 and calculate the average distance from age 23 to 60 in the benchmark economy and under the progressive tax ($\tau_p = 0.15$). The last column shows the change in the average distance, i.e. the difference between the second and the third column.

Figure 10: Timeline of the Working Stage

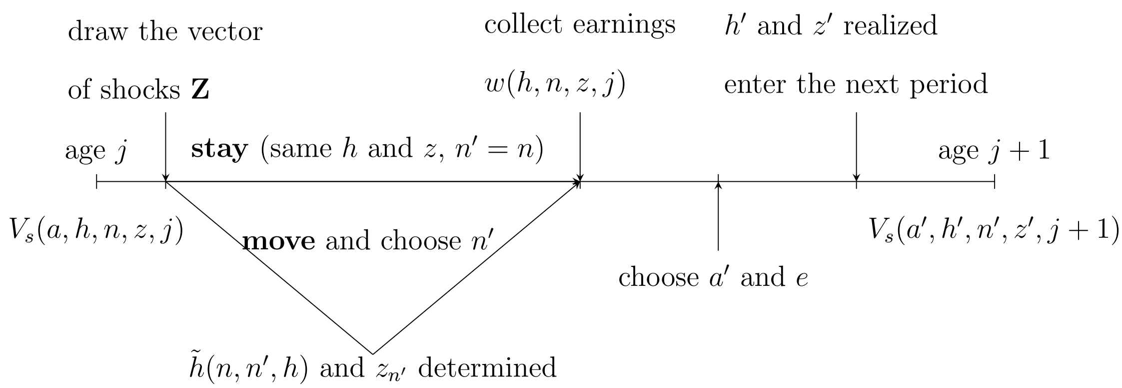
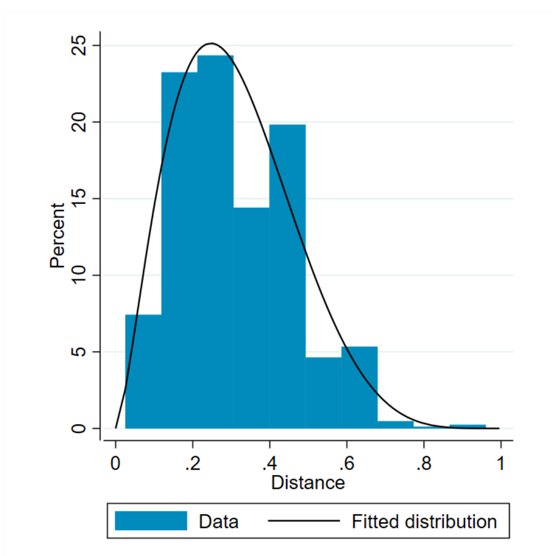
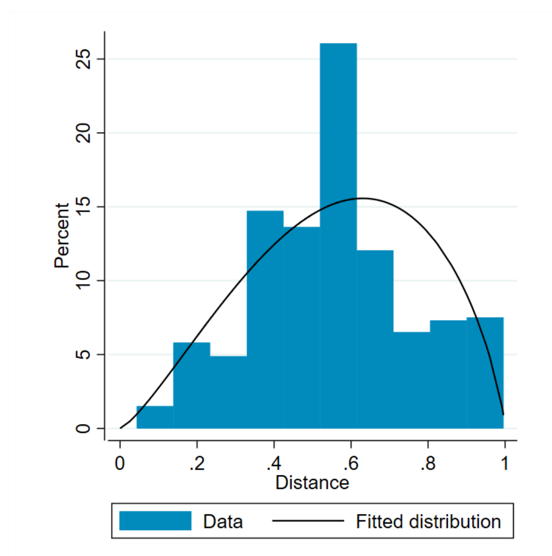


Figure 11: Initial Technology Distributions (college and non-college)

(a) Non-college at age 18



(b) College at age 23

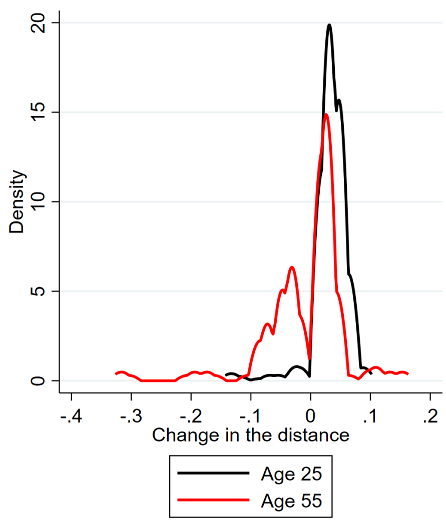


Note: This figure shows the initial distribution in terms of the distance for college and non-college workers. The solid lines represent the fitted Beta distribution used for the model as $F_n^{NC}(n)$ and $F_n^C(n)$.

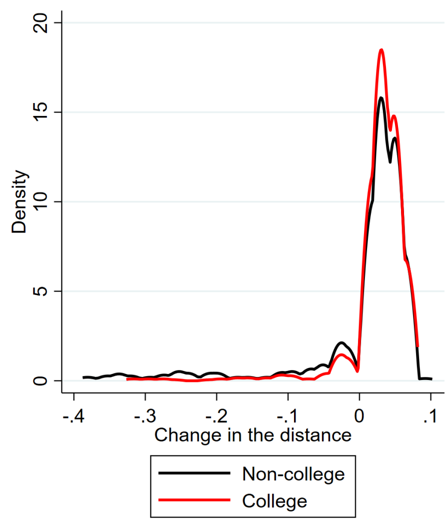
Source: Author's calculation from ASEC 1968-2019 and O*NET.

Figure 12: Kernel Density of Switching

(a) Technology switching by age

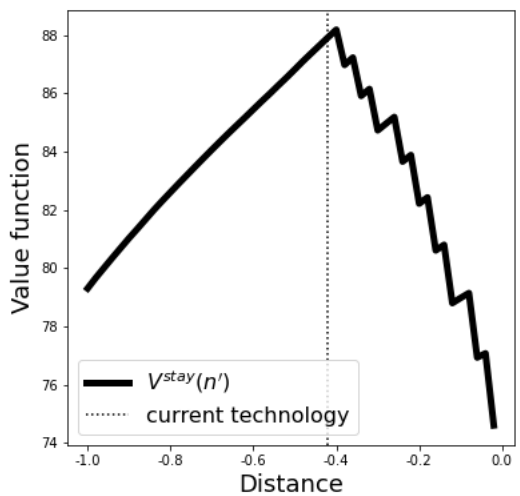


(b) Technology switching by education



Note: The figures show kernel density estimations of switching probabilities conditional on workers who are in the 3rd quintile group of the distance in the previous period and decide to switch to other technologies. A positive change in distance implies technology upgrading. Panel (a) shows the density for all workers by age. Panel (b) shows the density for workers at age 25 by education.

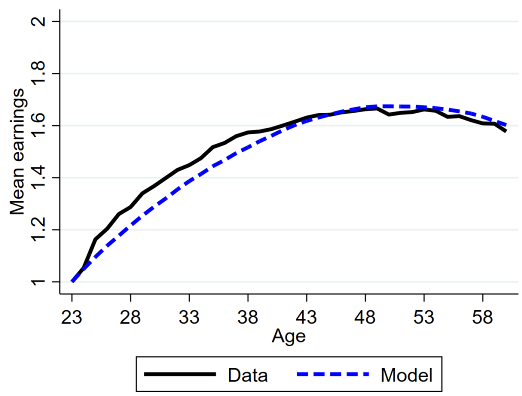
Figure 13: Value Function Plot



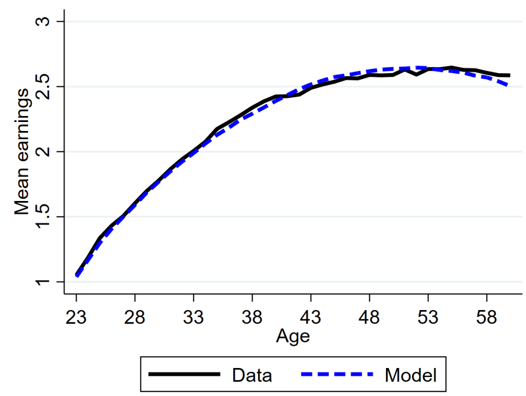
Note: The figure shows $V_C^{stay}(a, \tilde{h}(n', n, h), n', z_{n'}, j)$ as a function of n' at age 25 with all state variables evaluated at the median level. I also hold productivity shocks constant for all technologies. The vertical line stands for current technology position n.

Figure 14: Life-cycle Earnings Profiles

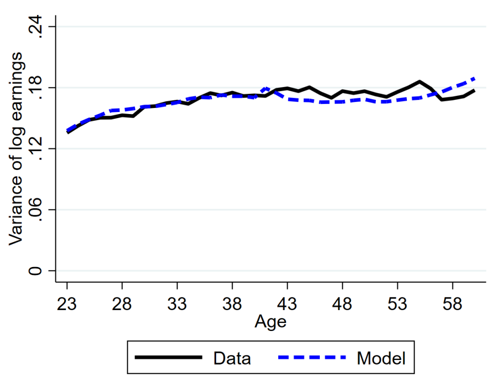
(a) Mean earnings: Non-college



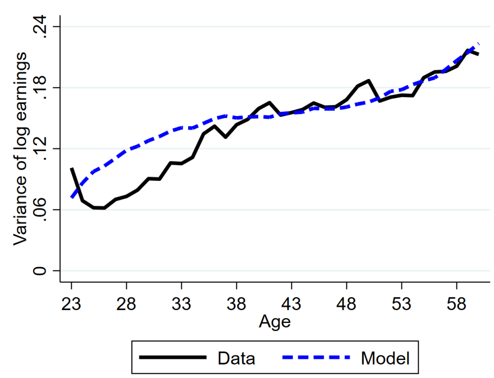
(b) Mean earnings: College



(c) Earnings inequality: Non-college



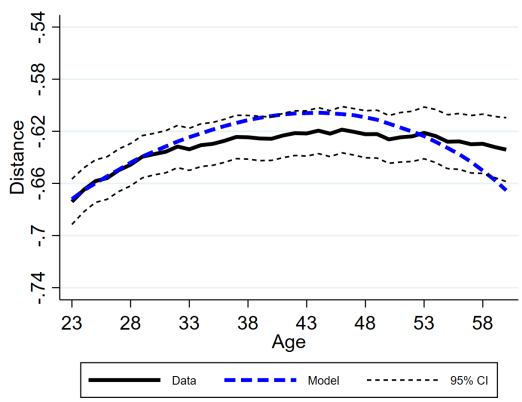
(d) Earnings inequality: College



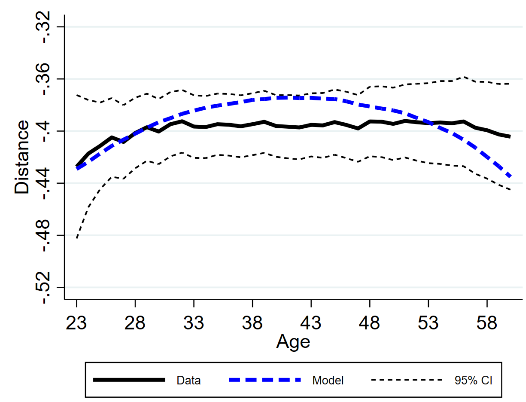
Note: Panel (a) shows the age profile of mean earnings for non-college workers and panel (b) is for college workers. The mean earnings of non-college workers at age 23 is normalized to 1 for comparison purposes. Panel (c) shows the age profile of variance of log earnings for non-college workers and panel (d) is for college workers.

Figure 15: Technology Usage Profile

(a) Mean distance: non-college



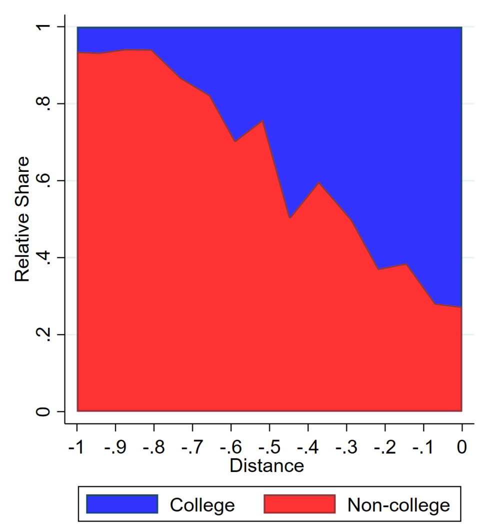
(b) Mean distance: college



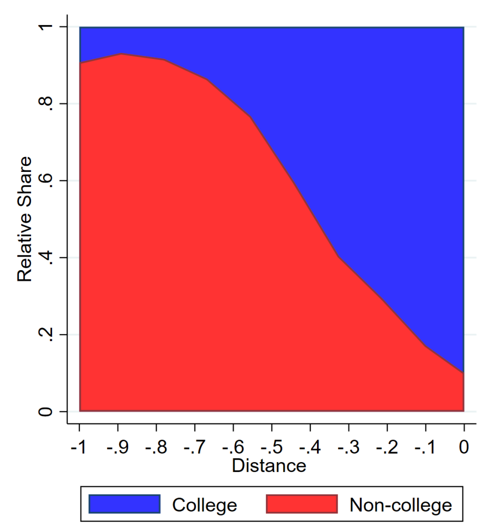
Note: Panel (a) shows the age profile of the mean distance for non-college workers and panel (b) shows the age profile of the mean distance for college workers.

Figure 16: Relative Share of Non-college Workers (untargeted)

(a) Data

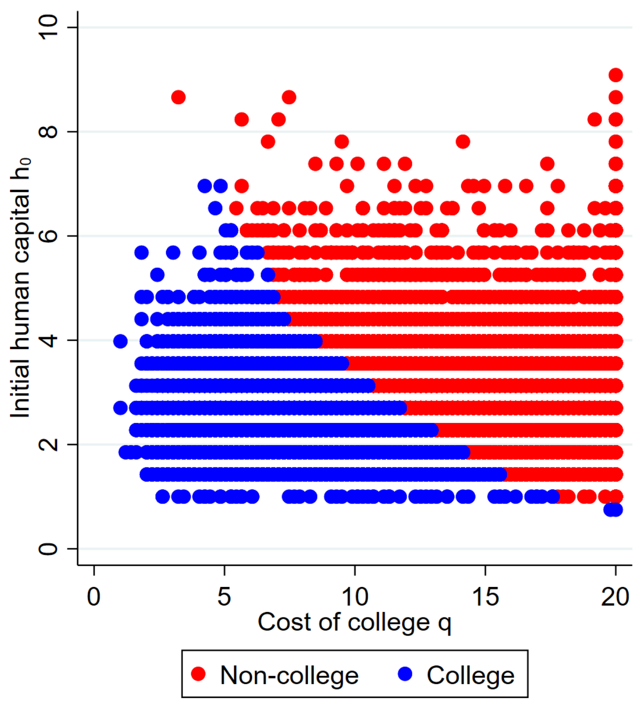


(b) Model



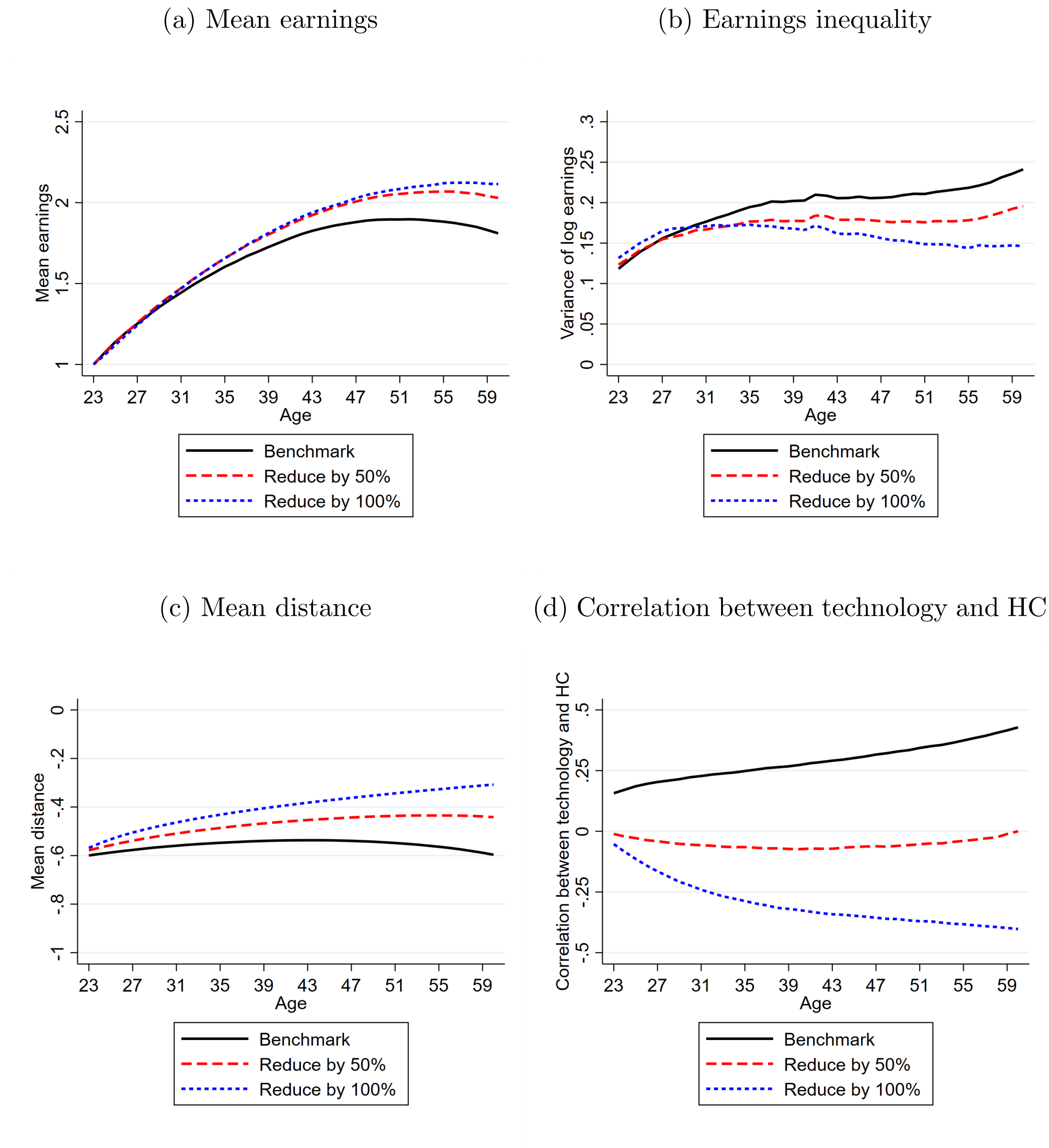
Note: This figure shows the relative share of non-college workers over the technology distribution. Specifically, I divide the all technologies into 15 bins with equal width and calculate the relative share of non-college workers in each bin.

Figure 17: College Decisions in the Benchmark Economy



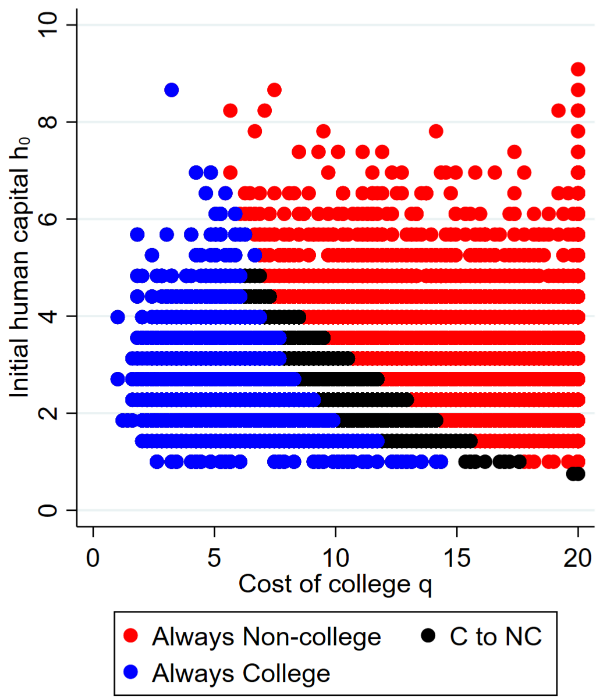
Note: This figure shows the college decision based on the joint distribution of initial human capital (y-axis) and cost of college education (x-axis). Blue dots denote people who attend college.

Figure 18: Experiments with the Catch-up Channel



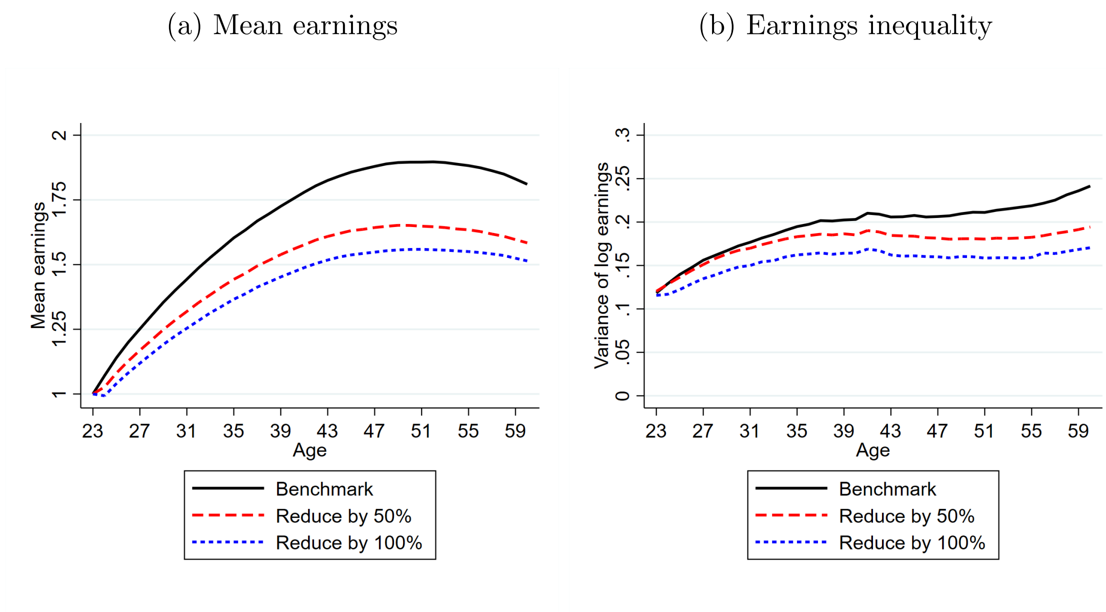
Note: The figure presents how life-cycle profiles change when reducing the catch-up channel. To reduce the catch-up channel by 50%, I set ϕ_0 to be 50% of the original level. ϕ_0 is set to 0 to completely shut down the catch-up channel. For comparison purpose, the mean earnings at age 23 are normalized to 1 in all scenarios in panel (a).

Figure 19: College Decisions After Shutting Down the Catch-up Channel



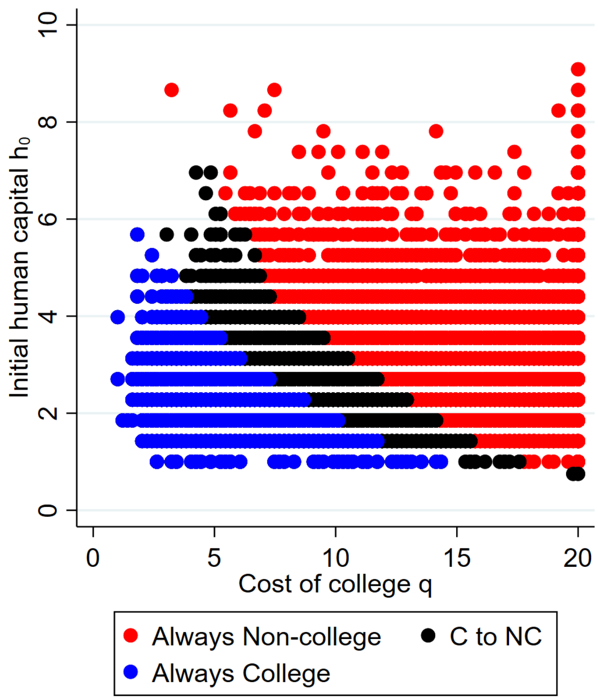
Note: Always college stands for people who go to college in both cases. C to NC are people who go to the college in the benchmark case ($\phi_0 = 3.1$) but decide to skip college after shutting down the catch-up channel ($\phi_0 = 0$).

Figure 20: Experiments with the Direct Channel



Note: The figure presents how life-cycle profiles change when reducing the direct channel. To reduce the catch-up channel by 50%, I set η to be 50% of the original level. η is set to 0 to completely shut down the catch-up channel. This implies that all technologies have the same productivity as the frontier technology. For comparison purpose, the mean earnings at age 23 are normalized to 1 in all scenarios in panel (a).

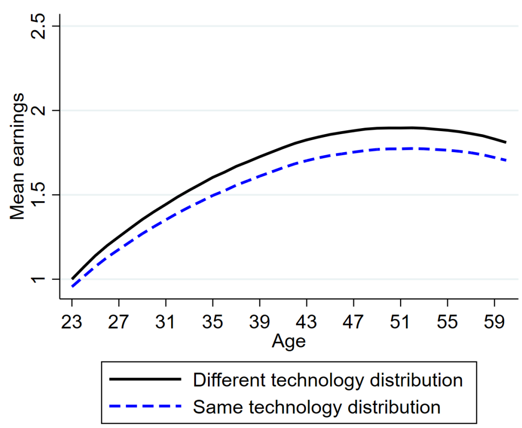
Figure 21: College Decisions After Shutting Down the Direct Channel



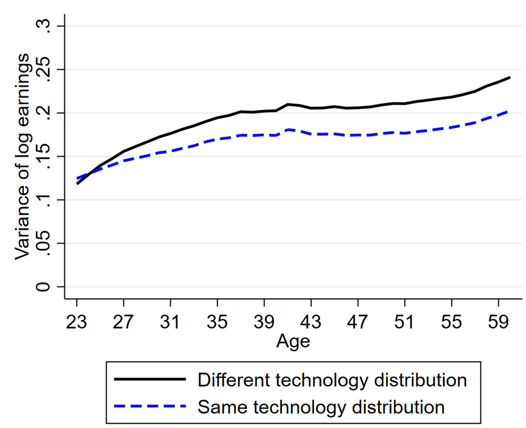
Note: Always college stands for people who go to college in both cases. C to NC are people who go to the college in the benchmark case ($\eta = 111$) but decide to skip college after shutting down the direct channel ($\eta = 0$).

Figure 22: Elimination of the Initial Advantage

(a) Mean earnings

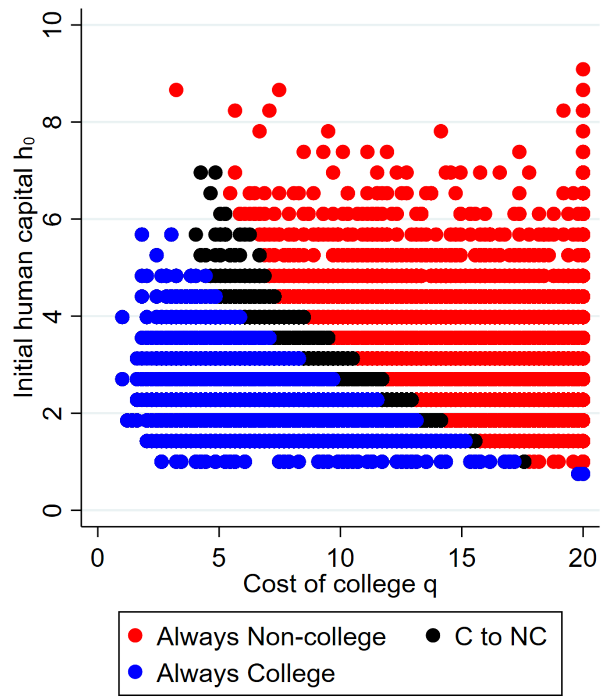


(b) Earnings inequality



Note: The figure presents life-cycle profiles when both educational groups draw initial technology choice from the same distribution (as non-college workers). In panel (a), the mean earnings at age 23 is normalized to 1 in the benchmark economy.

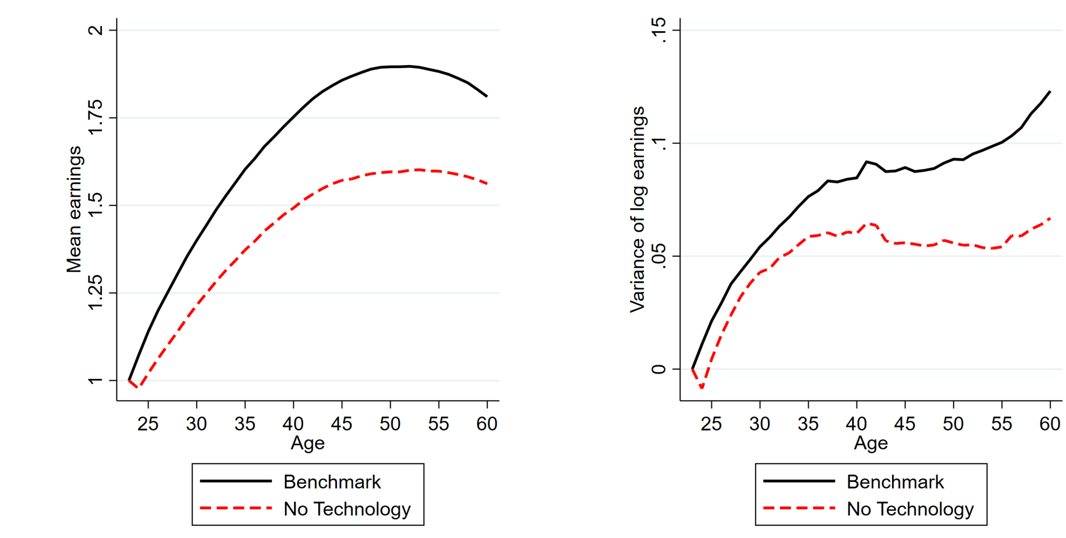
Figure 23: College Decisions when Eliminating the Initial Advantage



Note: Always college stands for people who go to college in both cases. C to NC are people who go to the college in the benchmark case but decide to skip college when the initial advantage is eliminated.

Figure 24: Remove Technology Usage

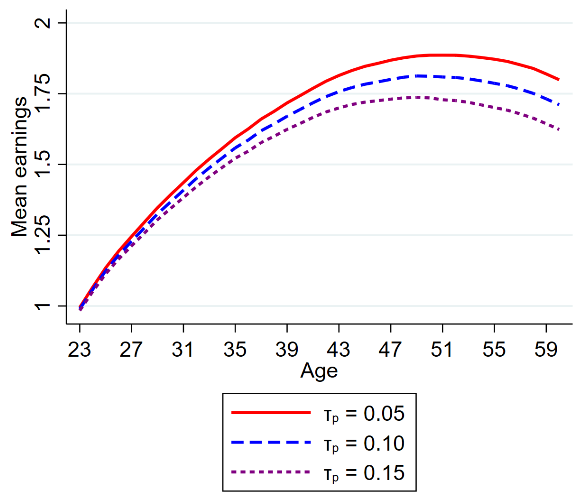
(a) Mean earnings (b) Earnings inequality



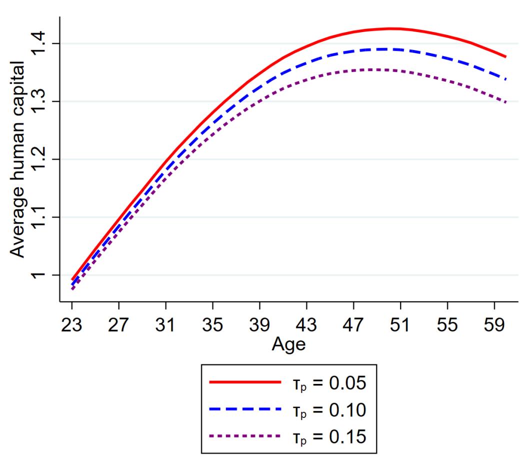
Note: The figure presents life-cycle profiles after removing technology usage from the benchmark model. The mean earnings at age 23 is normalized to 1 in panel (a) and the level of earning inequality at age 23 is normalized to 0 in panel (b).

Figure 25: Earnings Profiles under Progressive Taxes

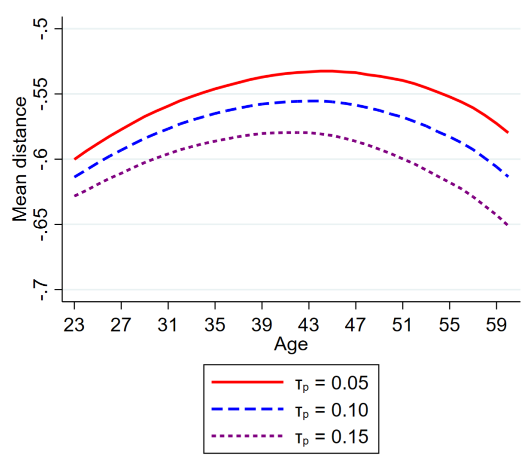
(a) Mean earnings



(b) Mean human capital

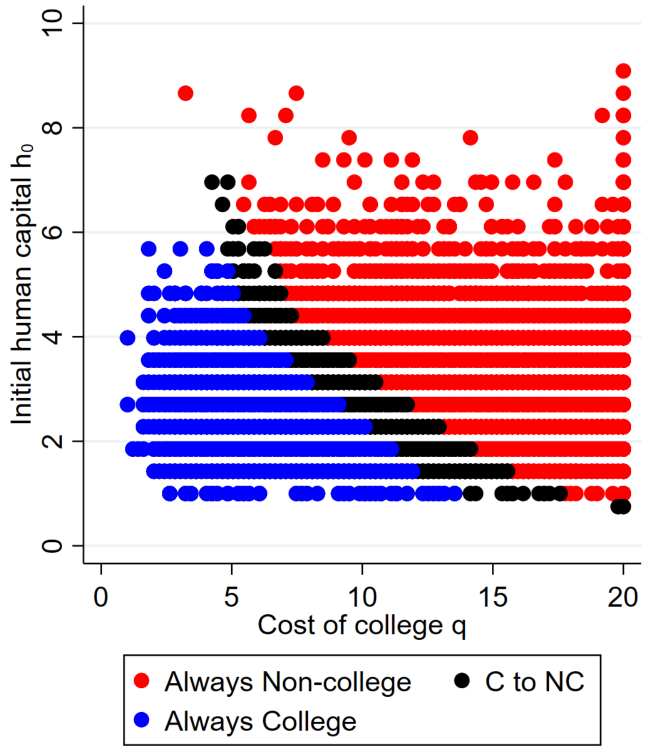


(c) Mean distance



Note: Panel (a) shows the average earnings profile over the life-cycle and panel (b) shows the average human capital profile. Both values at age 23 are normalized to 1 when $\tau_p = 0.05$. Panel (c) presents the mean distance profile, i.e., the average technology usage profile. A higher τ_p implies a more progressive tax schedule.

Figure 26: College Decisions under Progressive Taxes



Note: The figure shows college decisions conditional on the combination of college cost q and initial human capital h_0. I consider a tax reform where the economy switches from a U.S. progressivity level ($\tau_p = 0.05$) to a European level ($\tau_p = 0.15$). Always college means people who go to college in both cases. C to NC are people who go to college when tax is proportional but decide to skip the college stage when a tax on labor earnings is progressive.

REFERENCES

Acemoglu, D. and P. Restrepo, "Robots and jobs: Evidence from us labor markets", Journal of Political Economy **128**, 6, 2188–2244 (2020a).

Acemoglu, D. and P. Restrepo, "Unpacking skill bias: Automation and new tasks", in "AEA Papers and Proceedings", vol. 110, pp. 356–61 (2020b).

Acemoglu, D. *et al.*, "What does human capital do? a review of goldin and katz's the race between education and technology", Journal of Economic Literature **50**, 2, 426–63 (2012).

Akerman, A., I. Gaarder and M. Mogstad, "The skill complementarity of broadband internet", The Quarterly Journal of Economics **130**, 4, 1781–1824 (2015).

Autor, D. H., "How long has this been going on? a discussion of 'recent flattening in the higher/education wage premium: Polarization, skill downgrading, or both?'by robert g. valletta", in "NBER Conference on Research and Income in Wealth "Measuring and Accounting for Innovation in the 21st Century," Washington, DC (March 10–11, 2017)", (2017).

Autor, D. H., L. F. Katz and M. S. Kearney, "Trends in us wage inequality: Revising the revisionists", The Review of economics and statistics **90**, 2, 300–323 (2008).

Autor, D. H., F. Levy and R. J. Murnane, "The skill content of recent technological change: An empirical exploration", The Quarterly Journal of Economics **118**, 4, 1279–1333 (2003).

Badel, A., M. Huggett and W. Luo, "Taxing top earners: a human capital perspective", The Economic Journal **130**, 629, 1200–1225 (2020).

Beaudry, P. and D. A. Green, "Cohort patterns in canadian earnings: assessing the role of skill premia in inequality trends", Canadian Journal of Economics/Revue canadienne d'économique **33**, 4, 907–936 (2000).

Becker, G. S., "Investment in human capital: A theoretical analysis", Journal of Political Economy **70**, 5, Part 2, 9–49 (1962).

Ben-Porath, Y., "The production of human capital and the life cycle of earnings", Journal of political economy **75**, 4, Part 1, 352–365 (1967).

Benabou, R., "Tax and education policy in a heterogeneous-agent economy: What levels of redistribution maximize growth and efficiency?", Econometrica **70**, 2, 481–517 (2002).

Berger, M. C., "The effect of cohort size on earnings growth: A reexamination of the evidence", Journal of Political Economy **93**, 3, 561–573 (1985).

Blandin, A., "Learning by doing and ben-porath: Life-cycle predictions and policy implications", Journal of Economic Dynamics and Control **90**, 220–235 (2018).

Bloom, N., R. Sadun and J. Van Reenen, “Americans do it better: Us multinationals and the productivity miracle”, American Economic Review **102**, 1, 167–201 (2012).

Bowlus, A., E. Bozkurt, L. Lochner and C. Robinson, “Wages and employment: The canonical model revisited”, Tech. rep., National Bureau of Economic Research (2017).

Bowlus, A. J. and C. Robinson, “Human capital prices, productivity, and growth”, American Economic Review **102**, 7, 3483–3515 (2012).

Burstein, A., E. Morales and J. Vogel, “Changes in between-group inequality: computers, occupations, and international trade”, American Economic Journal: Macroeconomics **11**, 2, 348–400 (2019).

Casanova, M., “Revisiting the hump-shaped wage profile”, Unpublished mimeo, UCLA (2013).

Chari, V. V. and H. Hopenhayn, “Vintage human capital, growth, and the diffusion of new technology”, Journal of Political Economy **99**, 6, 1142–1165 (1991).

Ciccone, A. and G. Peri, “Long-run substitutability between more and less educated workers: evidence from us states, 1950–1990”, Review of Economics and statistics **87**, 4, 652–663 (2005).

Cubas, G. and P. Silos, “Career choice and the risk premium in the labor market”, Review of Economic Dynamics **26**, 1–18 (2017).

David, H. and D. Dorn, “The growth of low-skill service jobs and the polarization of the us labor market”, American Economic Review **103**, 5, 1553–97 (2013).

Deaton, A. and C. Paxson, “Intertemporal choice and inequality”, Journal of political economy **102**, 3, 437–467 (1994).

Dillon, E. W., “Risk and return trade-offs in lifetime earnings”, Journal of Labor Economics **36**, 4, 981–1021 (2018).

Erosa, A. and T. Koreshkova, “Progressive taxation in a dynastic model of human capital”, Journal of Monetary Economics **54**, 3, 667–685 (2007).

Esfahani, M., “Inequality over the life-cycle: Us vs europe”, Available at SSRN 3529362 (2020).

Feldstein, M. S., “The effects of taxation on risk taking”, Journal of Political Economy **77**, 5, 755–764 (1969).

Gallipoli, G. and C. A. Makridis, “Structural transformation and the rise of information technology”, Journal of Monetary Economics **97**, 91–110 (2018).

Galor, O. and O. Moav, “Ability-biased technological transition, wage inequality, and economic growth”, The Quarterly Journal of Economics **115**, 2, 469–497 (2000).

Goldin, C. D. and L. F. Katz, *The race between education and technology* (harvard university press, 2009).

Gouskova, E., "Parameter estimates comparison of earnings functions in the psid and cps data, 1976–2007", Economics Letters **122**, 2, 353–357 (2014).

Greenwood, J. and M. Yorukoglu, "1974", in "Carnegie-Rochester conference series on public policy", vol. 46, pp. 49–95 (Elsevier, 1997).

Guner, N., R. Kaygusuz and G. Ventura, "Income taxation of us households: Facts and parametric estimates", Review of Economic Dynamics **17**, 4, 559–581 (2014).

Guner, N., R. Kaygusuz and G. Ventura, "Rethinking the welfare state", (2021).

Guvenen, F., "Learning your earning: Are labor income shocks really very persistent?", American Economic Review **97**, 3, 687–712 (2007).

Guvenen, F., "An empirical investigation of labor income processes", Review of Economic Dynamics **12**, 1, 58–79 (2009).

Guvenen, F., F. Karahan, S. Ozkan and J. Song, "What do data on millions of us workers reveal about life-cycle earnings risk?", Tech. rep., National Bureau of Economic Research (2015).

Guvenen, F., B. Kuruscu and S. Ozkan, "Taxation of human capital and wage inequality: A cross-country analysis", Review of Economic Studies **81**, 2, 818–850 (2014).

Heathcote, J., K. Storesletten and G. L. Violante, "Presidential address 2019: How should tax progressivity respond to rising income inequality?", Journal of the European Economic Association **18**, 6, 2715–2754 (2020).

Heckman, J. J., L. Lochner and C. Taber, "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 (1998).

Hudomiet, P. and R. J. Willis, "Computerization, obsolescence, and the length of working life", Tech. rep., National Bureau of Economic Research (2021).

Huggett, M., G. Ventura and A. Yaron, "Sources of lifetime inequality", American Economic Review **101**, 7, 2923–54 (2011).

Jeong, H., Y. Kim and I. Manovskii, "The price of experience", American Economic Review **105**, 2, 784–815 (2015).

Jovanovic, B. and Y. Nyarko, "Learning by doing and the choice of technology", Econometrica **64**, 6, 1299–1310 (1996).

Jung, P. and M. Kuhn, "Earnings losses and labor mobility over the life cycle", Journal of the European Economic Association **17**, 3, 678–724 (2019).

Kambourov, G. and I. Manovskii, “Accounting for the changing life-cycle profile of earnings”, Unpublished manuscript, University of Toronto (2009a).

Kambourov, G. and I. Manovskii, “Occupational mobility and wage inequality”, The Review of Economic Studies **76**, 2, 731–759 (2009b).

Kambourov, G. and I. Manovskii, “Occupational specificity of human capital”, International Economic Review **50**, 1, 63–115 (2009c).

Katz, L. F. and K. M. Murphy, “Changes in relative wages, 1963–1987: supply and demand factors”, The quarterly journal of economics **107**, 1, 35–78 (1992).

Keller, E., “The slowdown in american educational attainment”, Journal of Economic Dynamics and control **46**, 252–270 (2014).

Kong, Y.-C., B. Ravikumar and G. Vandenbroucke, “Explaining cross-cohort differences in life-cycle earnings”, European Economic Review **107**, 157–184 (2018).

Kredler, M., “Vintage human capital and learning curves”, Journal of Economic Dynamics and Control **40**, 154–178 (2014).

Krueger, D. and A. Ludwig, “On the optimal provision of social insurance: Progressive taxation versus education subsidies in general equilibrium”, Journal of Monetary Economics **77**, 72–98 (2016).

Lagakos, D., B. Moll, T. Porzio, N. Qian and T. Schoellman, “Life cycle wage growth across countries”, Journal of Political Economy **126**, 2, 797–849 (2018).

Lemieux, T., “Postsecondary education and increasing wage inequality”, American Economic Review **96**, 2, 195–199 (2006).

Liu, K., “Wage risk and the value of job mobility in early employment careers”, Journal of Labor Economics **37**, 1, 139–185 (2019).

Magnac, T., N. Pistolesi and S. Roux, “Post-schooling human capital investments and the life cycle of earnings”, Journal of Political Economy **126**, 3, 1219–1249 (2018).

McKenzie, D. J., “Disentangling age, cohort and time effects in the additive model”, Oxford bulletin of economics and statistics **68**, 4, 473–495 (2006).

Mincer, J., “Human capital responses to technological change in the labor market”, (1989).

Parente, S. L., “Technology adoption, learning-by-doing, and economic growth”, Journal of economic theory **63**, 2, 346–369 (1994).

Restuccia, D. and G. Vandenbroucke, “The evolution of education: A macroeconomic analysis”, International Economic Review **54**, 3, 915–936 (2013).

Riddell, W. C. and X. Song, “The role of education in technology use and adoption: Evidence from the canadian workplace and employee survey”, ILR Review **70**, 5, 1219–1253 (2017).

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

Schultz, T. W., “The value of the ability to deal with disequilibria”, Journal of Economic Literature **13**, 3, 827–846 (1975).

Stiroh, K. J., “Information technology and the us productivity revival: what do the industry data say?”, American Economic Review **92**, 5, 1559–1576 (2002).

Storesletten, K., C. I. Telmer and A. Yaron, “Consumption and risk sharing over the life cycle”, Journal of monetary Economics **51**, 3, 609–633 (2004).

Valletta, R. G., “Recent flattening in the higher education wage premium: Polarization, skill downgrading, or both?”, Tech. rep., National Bureau of Economic Research (2016).

Violante, G. L., “Technological acceleration, skill transferability, and the rise in residual inequality”, The Quarterly Journal of Economics **117**, 1, 297–338 (2002).

Welch, F., “Effects of cohort size on earnings: The baby boom babies’ financial bust”, Journal of Political Economy **87**, 5, Part 2, S65–S97 (1979).

APPENDIX A

LIFE-CYCLE SKILL PREMIUMS ACROSS COHORTS

CPS ASEC Data

The data source is from IPUMS CPS dataset with concentration on the Annual Social and Economic Supplement (ASEC). The harmonization of the data largely follows Lemieux (2006). The analysis only focuses on wage and salary workers age 25-55.

The measurement of earnings is hourly wages, which is calculated by annual wage and salary income divided by total hours worked last year. The ASEC data only has information on the number of weeks worked last year. I use hours worked last week as an approximation of weekly working hours last year. The total hours worked last year are given by the product of total weeks worked last year and hours worked last week. The variable of the number of weeks worked last year is intervalled before 1975. So I replace this value with the average number of weeks in each interval calculated after 1975.

I adjust hourly wages by PCE deflator and drop all observations whose hourly wage is below \$ 1 or above \$ 100 (in 1979 dollars). I also multiply earnings by 1.4 for top-coded earners.

For educational category, I use the harmonized variable which combines two other variables that measure educational attainment in different ways. However, this variable is not available in 1963 so I start my analysis from 1964.

Regression Details

Based on BLS's definition, I divide 50 states and the D.C into four divisions: Northeast, South, Midwest and West. The race is divided into four categories: white, black, hispanic and others. The classificiation of marital status is dichotomous: married or not.

Underitendification of Human Capital

In my model, the observed hourly wage is a product of the rental rate and human capital:

$$w_{L,t,j} = W_{L,t} \cdot h_{L,t,j} \quad \text{and} \quad w_{H,t,j} = W_{H,t} \cdot h_{H,t,j}$$

Suppose I decompose the skill premium $\ln \frac{w_{H,t,j}}{w_{L,t,j}}$ into two time series $\ln \frac{W_{H,t}}{W_{L,t}}$ and $\ln \frac{h_{H,t,j}}{h_{L,t,j}}$, then I have three time series available: $w_{H,t,j}$, $w_{L,t,j}$, and $\ln \frac{h_{H,t,j}}{h_{L,t,j}}$ (or equivalently $\ln \frac{W_{H,t}}{W_{L,t}}$). However, there are four unknowns ($W_{L,t}$, $h_{L,t,j}$, $W_{H,t}$, $h_{H,t,j}$) to be determined so the model is underidentified. Therefore I have to normalize one of these four unknowns.

In this paper, I normalize the human capital of low-skill workers to be 1. This normalization means that I attribute all changes in relative human capital $\ln \frac{h_{H,t,j}}{h_{L,t,j}}$ to high-skill workers.

Cohort Grouping

I group 16 cohorts based on the order of birth and index them by g in the following way:

$$g = \begin{cases} 1 & \text{if} \quad \text{birth year} \in [1914, 1929) \\ 2 & \text{if} \quad \text{birth year} \in [1929, 1939) \\ 3 & \text{if} \quad \text{birth year} \in [1939, 1949) \\ 4 & \text{if} \quad \text{birth year} \in [1949, 1959) \\ 5 & \text{if} \quad \text{birth year} \in [1959, 1969) \\ 6 & \text{if} \quad \text{birth year} \in [1969, 1979) \\ 7 & \text{if} \quad \text{birth year} \in [1979, 1989] \end{cases}$$

Robustness Check: Granular Bins

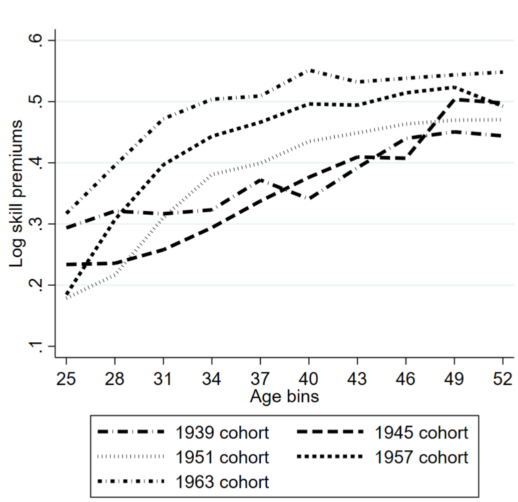
In Figure 27, I present life-cycle profiles of the skill premium using a 3-year bins. The left panel shows the estimated skill premium from the raw data and the right panel is the smoothed figure using LOWESS with a 0.5 bandwidth. A more granular grouping would generate noisy life-cycle profiles but patterns are greatly not affected by that after smoothing.

The change in age patterns that I mentioned in the paper still holds after smoothing. The 1939 and 1945 cohort have increasing skill premium profiles while for the subsequent cohorts the skill premium stops growing after age 43.

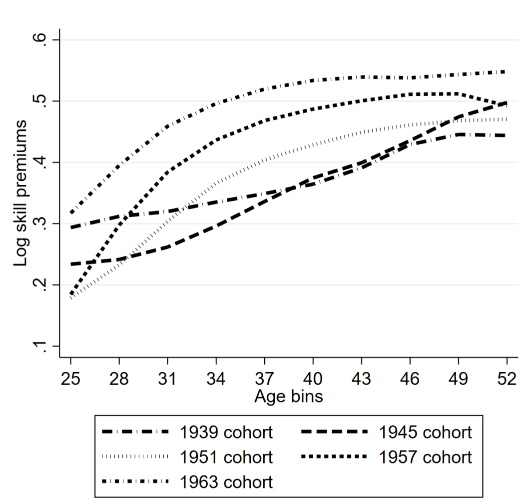
I formally document the growth pattern in Table 13. Similarly, the growth in the second phase still becomes smaller for recent cohort. Besides, the life-cycle growth increases before the 1951 cohort. All these patterns are similar to my analysis in the paper.

Figure 27: Life-cycle Profiles of Skill Premiums (3-Year bin width)

(a)



(b)



Note: The left panel shows the life-cycle profiles of the skill premium using 3-year bin. The right panel is the smoothed result using LOWESS with 0.5 bandwidth.

Table 13: Life-cycle Skill Premium Growth Patterns (3-Year bin width)

Cohort	Life-cycle growth	First phase	Second phase	$\frac{\text{First phase's growth}}{\text{Life-cycle growth}}$
1939	0.150	0.071	0.079	50.7%
1964	0.264	0.141	0.123	53.4%
1945	0.292	0.250	0.042	85.6%
1951	0.308	0.303	0.005	98.3%
1957	0.231	0.217	0.014	93.9%

Note: The life cycle growth is the difference in the skill premium between age 25 and 52. The first phase's growth is the difference between age 25 and 40. The second phase's growth is the difference between age 40 and 52.

APPENDIX B

AN EMPIRICAL INVESTIGATION OF TECHNOLOGY USAGE ON EARNINGS

The Construction of the Distance to the Frontier

I use a combination of inputs from the O*NET data set to construct the index to measure technology usage at the individual level. The O*NET data set provides detailed information on the importance of knowledge, tasks and skills for each occupation. In particular, a random sample of workers chooses the description that best fits their daily work in one specific aspect (for example programming skills). The answers are on a scale from 1 ("not important") to 6 ("extremely important"). The index of importance for that occupation is the average responses from the sample of workers.

I extract indices of the following characteristics: knowledge about computers and electronics, activities interacting with computers, programming skills, systems evaluation skills, quality control analysis skills, operations analysis skills, activities with updating and using relevant knowledge, technology design skills, activities analyzing data and information, activities processing information, knowledge with engineering and technology, and activities managing material resources.

I sum all the values from the above characteristics and normalize the sum to the interval $[-1, 0]$. The normalized index is denoted as the *distance to the frontier*. By construction, it measures how intensively workers use information technology at their daily work. The occupation that uses information technology most intensively is considered to be the frontier technology and its distance to the frontier is 0.

APPENDIX C

TECHNOLOGY USAGE AND LIFE-CYCLE EARNINGS

Stationary Equilibrium

Definition: A stationary equilibrium is a collection of college decision $s(h_0, q)$ and joint initial distribution $\Lambda(h_0, q)$, individual choice $\{c_j(\Theta), a_j(\Theta), n_j(\Theta), e_j(\Theta)\}_{j=23}^{J_R}$ at the working stage with state $\Theta = (a, n, h, z; s)$, individual choice $\{a_j(a_{j-1}, s)\}_{j=J_R+1}^{J}$ at the retirement stage, government policies $\{\tau_{ss}, b_C^{ss}, b_{NC}^{ss}, \tau, G\}$ and the sequence of population shares $\{\mu_j\}_{j=23}^{J}$ such that:

1. Individuals' decisions solve the optimization problems discussed in Section 3.

2. Government budget constraint is balanced:

$$\sum_{j=23}^{J} \mu_j \int E[T(w_j(\Theta), a_j(\Theta))] d\Lambda = G$$

3. The social security budget is balanced:

$$\tau_{ss} \sum_{j=23}^{J_R} \mu_j \int E[w_j(\Theta)] d\Lambda = \sum_{j=J_R+1}^{J} \mu_j [\omega b_C^{ss} + (1-\omega) b_{NC}^{ss}]$$

where $\omega = \int \mathbb{1}\{s(h_0, q) = C\} d\Lambda$ is the fraction of college workers.

Technology Switching Moments

I utilize the panel property of the ASEC data set to construct moments related to switching probabilities and compare them with the simulated model. I define "staying" as the absolute change in the distance is less than 0.02, which is the minimum step that one can move in the model. Any change that exceeds 0.02 is considered to be a technology upgrade and the definition of a downgrade is similar.

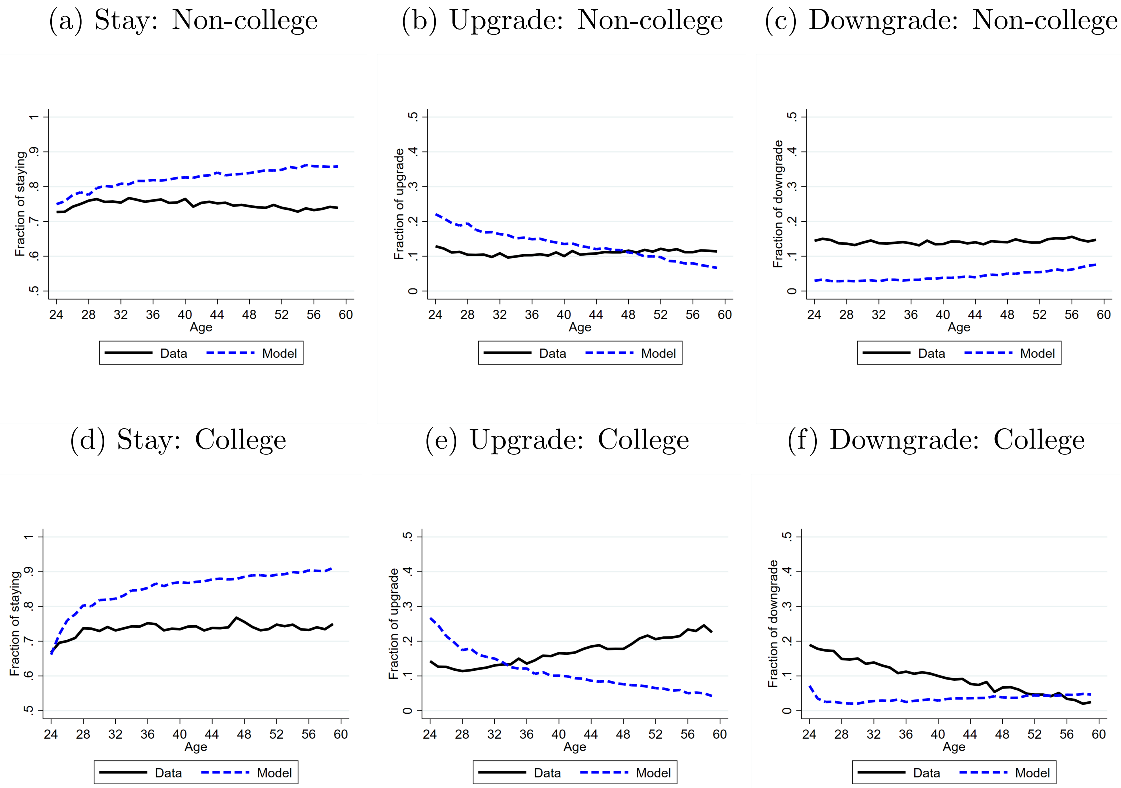
One potential issue in this exercise is the inconsistency in the time period. The CPS outgoing rotation group (ORG), which allows me to keep track of workers over time, starts in 1977. However, my analysis of technology usage takes information from 1968. Therefore it not is guaranteed that a well-parameterized model could match the switching moments well.

Indeed, as shown in Figure 28, my model overstates the probability of staying relative to the data as shown in panel (a) and (d). For instance, at age 60, around 90% of college workers stay with the same technology in the model while this fraction is only 70% in the data. Also, the model understates the probability of downgrading as shown in panel (c) and (f).

Life-cycle Profiles Conditional on Educational Group

In this part I present the life-cycle earnings profiles in the counterfactual experiments conditional on educational group. In Figure 29, I present the conditional life-cycle earnings profiles when shutting down the catch-up channel. The mean earnings growth increases for educational groups and the life-cycle inequality decreseas

Figure 28: Age Profiles of the Probabilities of Switching/Staying



Note: I plot the age profiles of the fractions of workers choose technology upgrade and downgrade. Technology upgrade is defined as the change in the distance greater than 0.02. Similarly, technology downgrade is defined as the change in the distance less than -0.02.
Source: Author's calculation from CPS ASEC/ORG 1978-2019 and O*NET.

for both educational groups. Therefore the change at the aggregate level is not only driven by the compositional effect.

However, the change in mean earnings for college workers is not monotone in the sense that the mean growth is larger when the catch-up cost is reduced by 50%. When there is not catch-up cost, college workers will upgrade technology more frequently so they suffer more human capital loss. So the growth in mean earnings slightly declines but the absolute level of earnings increases.

Similarly, Figure 30 shows the conditional life-cycle earnings profiles when shutting down the direct channel. The mean earnings growth and life-cycle inequality both decreases for each educational group, which also indicates that the change at the aggregate level is not only driven by the compositional effect.

Figure 31 and Figure 32 show the conditional life-cycle earnings profiles under taxation experiment. A more progressive depresses mean earnings growth and distorts the incentive for human capital accumulation and technology upgrading for both college and non-college workers.

Figure 29: Experiments with the Catch-up Channel

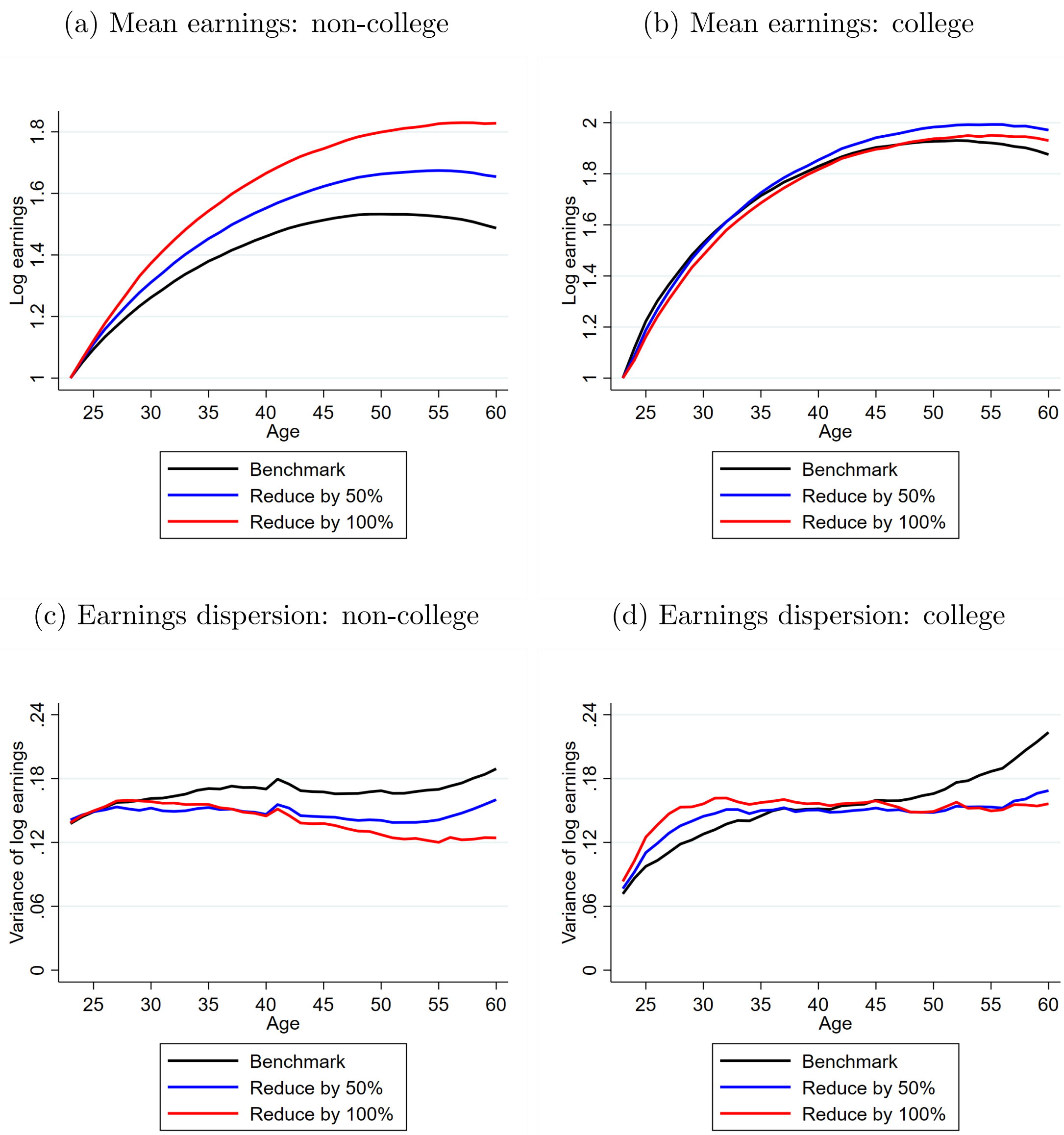
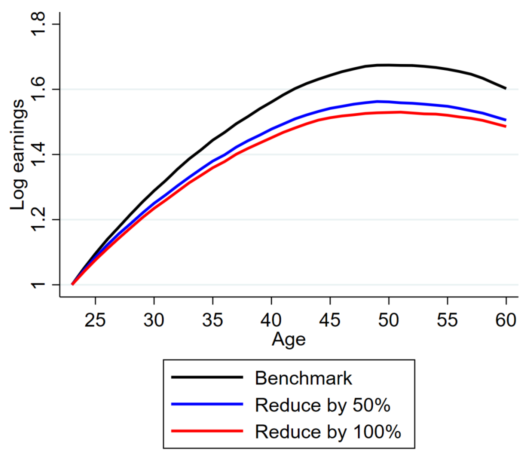
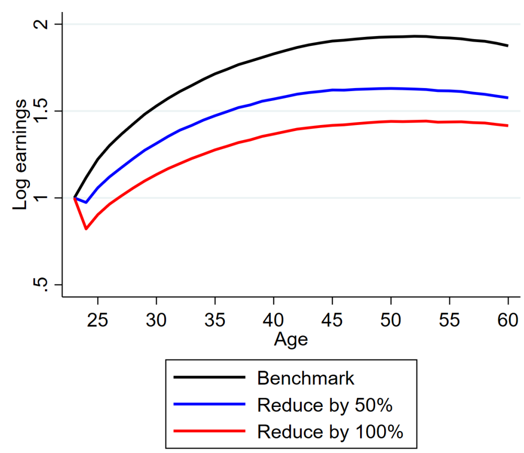


Figure 30: Experiments with the Direct Channel

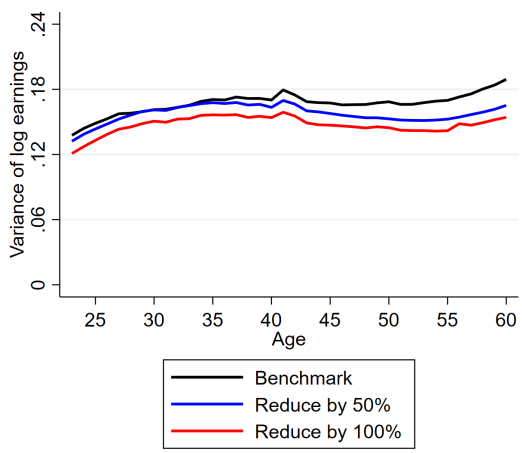
(a) Mean earnings: non-college



(b) Mean earnings: college



(c) Earnings dispersion: non-college



(d) Earnings dispersion: college

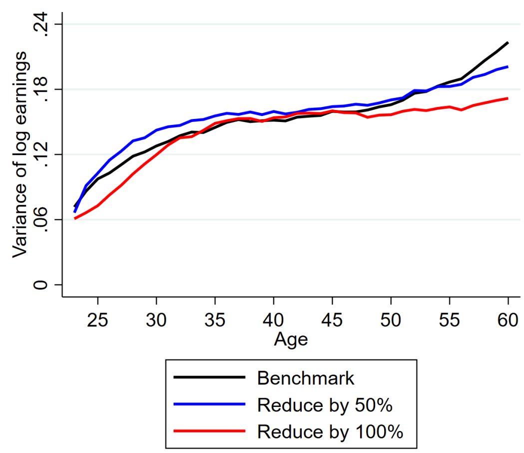
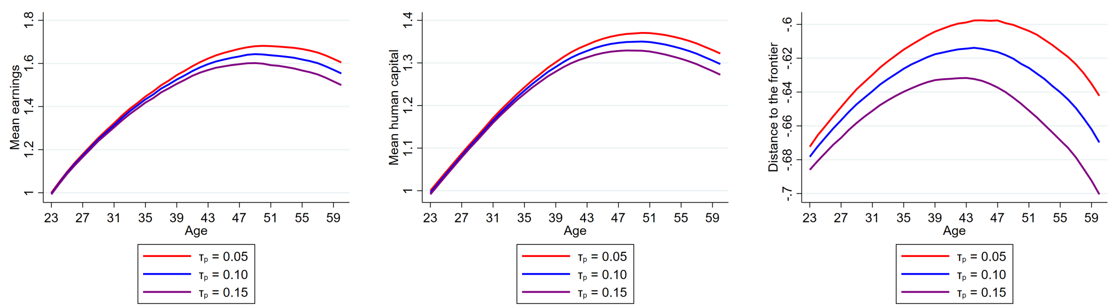


Figure 31: Earnings Profiles under Progressive Taxes (non-college workers)

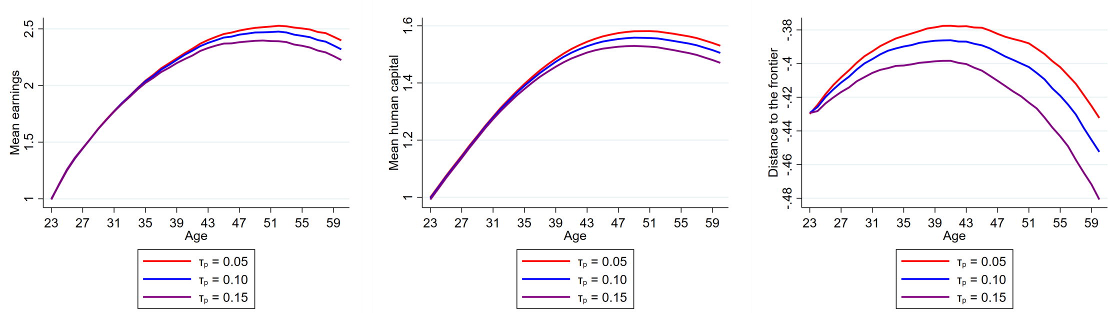
(a) Mean earnings (b) Mean human capital (c) Mean distance



Note: Panel (a) shows the average earnings profile over the life-cycle and panel (b) shows the average human capital profile. Both values at age 23 are normalized to 1 when $\tau_p = 0.05$. Panel (c) presents the mean distance profile, i.e., the average technology usage profile. A higher τ_p implies a more progressive tax schedule.

Figure 32: Earnings Profiles under Progressive Taxes (college workers)

(a) Mean earnings (b) Mean human capital (c) Mean distance



Note: Panel (a) shows the average earnings profile over the life-cycle and panel (b) shows the average human capital profile. Both values at age 23 are normalized to 1 when $\tau_p = 0.05$. Panel (c) presents the mean distance profile, i.e., the average technology usage profile. A higher τ_p implies a more progressive tax schedule.