# Life-cycle Skill Premiums across Cohorts

Siyu Shi*

Arizona State University

August 27, 2022

**Abstract**

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, which is the log ratio of wage rates between high-skill workers and low-skill workers.

**Keywords:** Human Capital, Life-cycle, Skill Premium, Skill-biased Technological Change

*I would like to thank Bart Hobijn, Alexander Bick and Gustavo Ventura for their continuous help and guidance. All errors are my own.

# 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 (2009) 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 (2009) 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

[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.

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 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.

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 2 presents empirical evidence about the life-cycle pattern of the skill premium by cohorts. In Section 3, I introduce the

[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).

human capital model and use its implication to decompose the skill premium. Section 4 discusses the results. In Section 5, I conclude and discuss potential implications from my result.

# 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.

## 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

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

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 $t$:

$$\ln w_{i,t,j} = \gamma_0 + \omega_{t,j} C_{i,t,j} + X'_{i,t,j}\gamma_1 \tag{1}$$

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]

## 2.2 Construction of Life-cycle Profiles

Though the CPS is repeated cross-sectional data, I use the method from Kambourov and Manovskii (2009) 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

[4] See more details in Appendix A.

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

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.

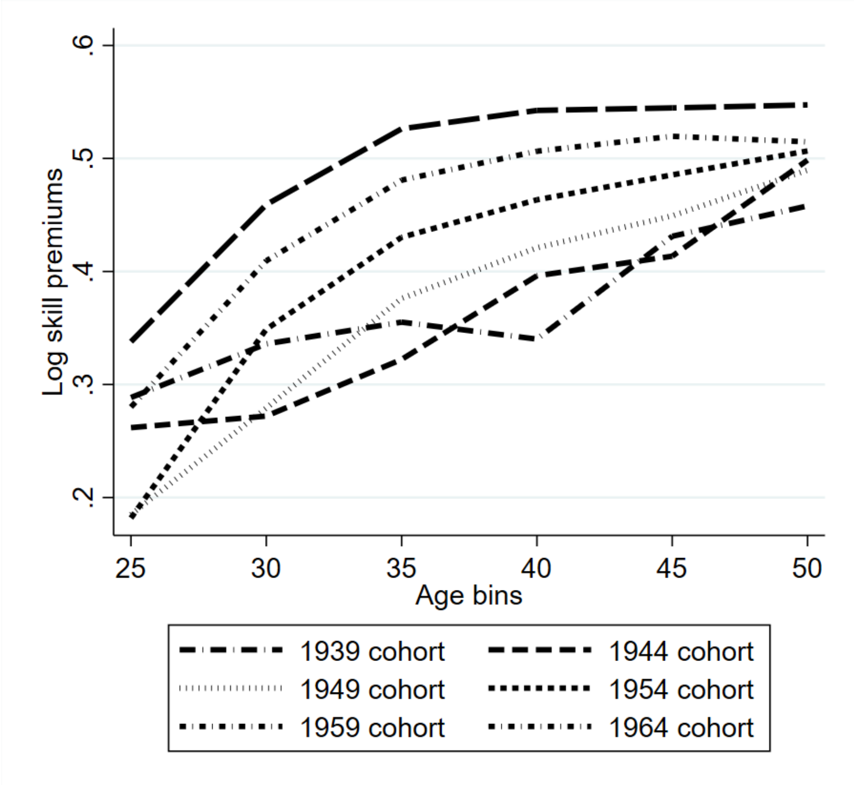
## 2.3 Life-cycle Profiles Vary Across Cohort

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 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,

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.

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

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.

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.

## 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.

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

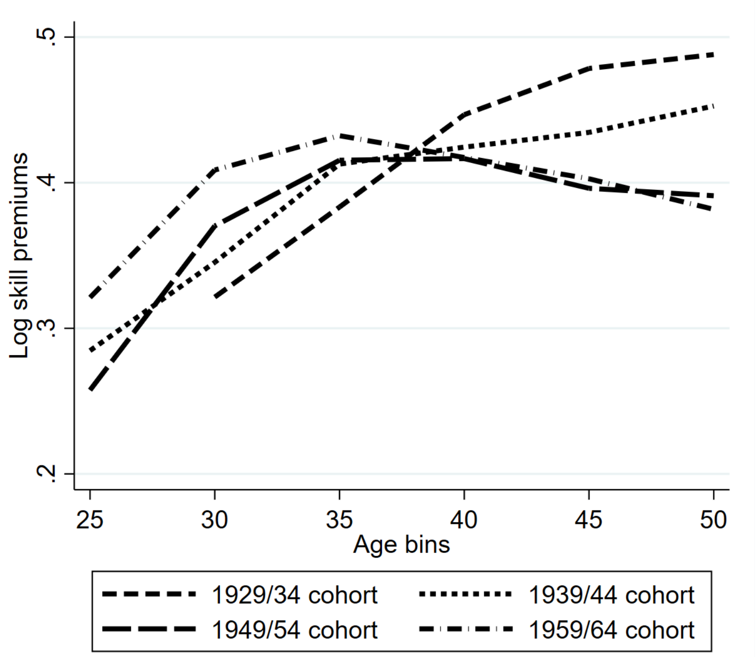
A well-know underidentification problem (see Deaton and Paxson (1994)) arises since a person's age added to the cohort year (birth year) gives the current year which means there is an exact linear relationship between time, age, and cohort effects. To avoid the problem of colinearity, I combine several 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 C.

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:

$$\omega_{t,j,g} = s_t^{time} + s_j^{age} + s_g^{cohort} + s_j^{age} \cdot s_g^{cohort} + \varepsilon_{t,j} \tag{2}$$

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.

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 $g$.

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.

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.

# 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).

## 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:

$$Y_t = [(A_{H,t}H_t)^{\frac{\sigma-1}{\sigma}} + (A_{L,t}L_t)^{\frac{\sigma-1}{\sigma}}]^{\frac{\sigma}{\sigma-1}} \tag{3}$$

where $\sigma$ 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:

$$W_{L,t} = Y_t^{\frac{1}{\sigma-1}} A_{L,t}^{\frac{\sigma-1}{\sigma}} L_t^{-\frac{1}{\sigma}} \quad \text{and} \quad W_{H,t} = Y_t^{\frac{1}{\sigma-1}} A_{H,t}^{\frac{\sigma-1}{\sigma}} H_t^{-\frac{1}{\sigma}} \tag{4}$$

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

$$\ln \frac{W_{H,t}}{W_{L,t}} = \frac{\sigma-1}{\sigma} \ln(\frac{A_{H,t}}{A_{L,t}}) - \frac{1}{\sigma} \ln(\frac{H_t}{L_t}) \tag{5}$$

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 $\frac{1}{\sigma}$.

## 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 $\lambda_i$. 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

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

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.

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

$$\max_{\{x_{j,i},c_{j,i}^{H}\}_{j=1}^{J}} \quad E\left[\sum_{j=1}^{J} \beta^{j-1}\left(\ln c_{j,i}^{H} - \phi_i \frac{x_{j,i}^2}{2}\right)\right] \tag{6}$$

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

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 $j+i$:

$$c_{j,i}^{H} = w_{j,i} = W_{H,j+i} h_{j,i} \tag{7}$$

The following equation describes the technology of human capital production where $\rho_i$ is the cohort-specific rate of return of investment, and $\delta_i$ is the depreciation of human capital.

$$\ln h_{j+1,i} = \ln h_{j,i} + \rho_i x_{j,i} - \delta_i \tag{8}$$

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

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

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 $\phi_i \frac{x_{j,i}^2}{2}$. The marginal disutility of investment is zero when $x_{j,i} = 0$. The constant rate of return of investment $\rho_i$ 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 $\rho_i$, 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 $\delta_i$ 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.

### 3.2.2 Optimal Investment Decisions

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

$$V_{j,i}(h_{j,i}) = \max_{x_{j,i} \in [0,\infty]} \ln W_{H,j+i} + \ln h_{j,i} - \phi_i \frac{x_{j,i}^2}{2} + \beta E[V_{j+1,i}(h_{j+1,i})] \tag{9}$$

[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 I replace $\ln w_{H,j+i}$ with $\ln W_{H,j+i} + \ln h_{j,i}$ using equation (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:

$$x_{j,i}\phi_i = \beta E[\frac{\partial V_{j+1,i}}{\partial \ln h_{j+1,i}}]\frac{\partial lnh_{j+1,i}}{\partial x_{j,i}} = \beta E[\frac{\partial V_{j+1,i}}{\partial \ln h_{j+1,i}}]\rho_i \tag{10}$$

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 (8).

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

$$\frac{\partial V_{j,i}}{\partial \ln h_{j,i}} = 1 + \beta E[\frac{\partial V_{j+1,i}}{\partial \ln h_{j+1,i}}\frac{\partial \ln h_{j+1,i}}{\partial \ln h_{j,i}}] = 1 + \beta E[\frac{\partial V_{j+1,i}}{\partial \ln h_{j+1,i}}] \tag{11}$$

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

$$\frac{\partial V_{j,i}}{\partial \ln h_{j,i}} = 1 + \beta + ... + \beta^{J-j} = \frac{1-\beta^{J-j}}{1-\beta} \tag{12}$$

Taking this equation to equation (10) yields the optimal investment decisions:

$$x_{j,i} = \frac{\rho_i\beta}{\phi_i}\frac{1-\beta^{J-j}}{1-\beta} \quad \text{for all} \quad 1 \leq j \leq J-1 \tag{13}$$

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

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

older, which is shown in equation (12).

It is also intuitive that the investment is decreasing in cost parameter $\phi_i$ and increasing in discount factor $\beta$ and the rate of return $\rho_i$. 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]

### 3.2.3 Life-cycle Profiles of Human Capital

Using optimal investment decisions from equation (13), I can derive the life-cycle profile of log human capital as follows:

$$\ln h_{j,i} = \ln h_{1,i} + \left(\frac{\rho_i \beta}{\phi_i(1-\beta)} - \delta_i\right) \cdot (j-1) + \frac{\rho_i \beta^{J+1}}{\phi_i(1-\beta)^2} \cdot \left(1-\beta^{-(j-1)}\right) \tag{14}$$

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 (13). The speed of growth depends positively on the rate of return $\rho_i$ and negatively on investment cost $\phi_i$, 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

[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.

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

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.

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

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

1. Given wage rates and the initial condition, individuals from cohort $i$ choose $\{\{x_{j,i}\}_{j=1}^J, \{c^L_{j,i}\}_{j=1}^J, \{c^H_{j,i}\}_{j=1}^J\}$ 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 (8) given $\{\{x_{j,i}\}_{j=1}^J\}_{i=-\infty}^{\infty}$.

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.

### 3.4 Decomposition of the Skill Premium

From equation (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 (5), the skill premium of age group $j$ at period $t$ can be expressed as follows:

$$\begin{aligned} \omega_{t,j} \equiv \ln \frac{w_{H,t,j}}{w_{L,t,j}} &= \ln h_{j,t-j} + \ln \frac{W_{H,t}}{W_{L,t}} \\ &= \underbrace{\ln h_{j,t-j}}_{\text{human capital}} + \underbrace{\frac{\sigma-1}{\sigma} \ln(\frac{A_{H,t}}{A_{L,t}}) - \frac{1}{\sigma} \ln(\frac{H_t}{L_t})}_{\text{relative skill price}} \end{aligned} \tag{15}$$

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 (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 $\rho_i$, investment cost $\phi_i$ and the depreciation $\delta_i$. Therefore the life-cycle profile of human capital is a mixture of age effects and cohort effects.

# 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.

## 4.1 Identification

Equation (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 (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:

$$\omega_{t,j} = \beta_{0,i} + \beta_{1,i} \cdot (j-1) + \beta_{2,i} \cdot \left(1 - \beta^{-(j-1)}\right) + \alpha_t + \varepsilon_{t,j} \tag{16}$$

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

[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 $\beta = 0.98^5$.

### 4.1.1 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.

### 4.1.2 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 $\beta_1$ and $\beta_2$ 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 $i$.

#### 4.1.3 Relative skill price

To unpack the relative skill price $\alpha_t$, 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 $\eta$ is the log growth rate. So the relative skill price can be expressed as follows:

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

Given the estimated skill price $\hat{\alpha}_t$ from regression (16), I can further estimate the following equation:

$$\hat{\alpha}_t = \beta_3 \cdot t + \beta_4 \ln \frac{H_t}{L_t} + u_t \tag{17}$$

where $\beta_3 = \frac{\sigma-1}{\sigma}\eta$, $\beta_4 = -\frac{1}{\sigma}$, and $u_t$ represents unobservable shocks. The coefficient $\beta_3$ shows the growth rate in the skill premium resulting from skill-biased technological change. The reciprocal of the absolute value of $\beta_4$ 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 (16), human capital can be backed out as

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

[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.

The relative supply is then given by

$$\frac{H_t}{L_t} = \frac{\sum_{j=1}^{J}(\lambda_{t-j} \cdot \hat{h}_{j,t-j})}{\sum_{j=1}^{J}(1-\lambda_{t-j})} \tag{18}$$

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

Equation (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 (18).

## 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

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

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.

Table 3: Cohort-specific structural parameters

| Group | Cohort | Net return $\frac{\rho_i}{\phi_i}$ | Depreciation $\delta_i$ |
|---|---|---|---|
| 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

Figure 3: Life-cycle Profiles of log Human Capital

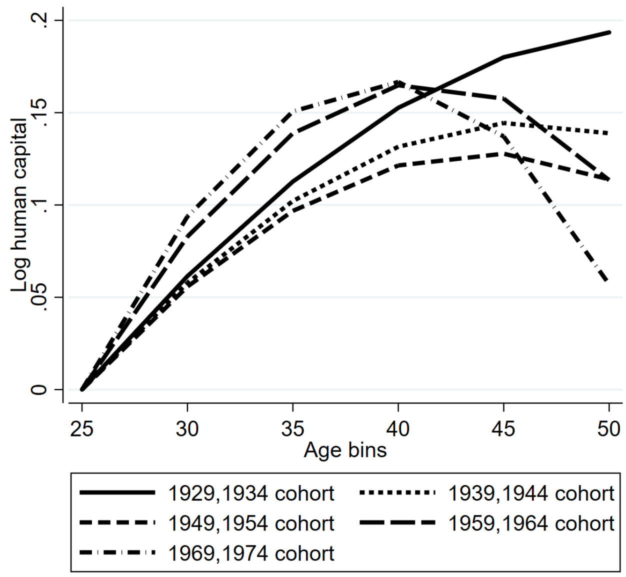



Figure 3 presents the life-cycle profiles of human capital based on the estimation. As discussed above, the net return determines how fast (high-skill) workers could accumulate their human capital mainly in the early stage of the life cycle. For example, the net return from the 1959 cohort is almost 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.

## 4.3 Relative Skill Price

Figure 4 shows the estimated relative skill price $\alpha_t$ 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.

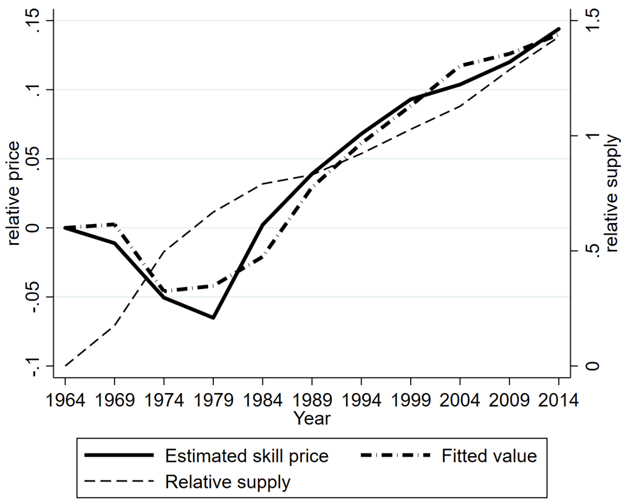
To better understand the evolution of the relative skill price, I fit equation (17) to obtain the following estimate[16]:

$$\begin{aligned} \alpha_t &= 0.064 \cdot t - 0.351 \cdot \ln \frac{H_t}{L_t} \\ &\quad\ (0.002) \quad\ (0.013) \end{aligned} \tag{19}$$

This first coefficient implies that the growth of relative skill price driven by demand

[15] The downward trend before 1980 is also documented by Acemoglu et al. (2012).
[16] Standard errors are shown in parentheses

Figure 4: Relative Skill Price




Note: The solid line represents $\alpha_t$ estimated from equation (16) and the dashdot line represents fitted values from equation (19). The relative supply is constructed based on equation (18) and is normalized to 0 in 1964.

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.

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. Sur-

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

prisingly, 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.

## 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

[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).

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.

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.

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.

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.

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.

# 5 Conclusion and Discussion

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 (2020) 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.

# References

ACEMOGLU, D. AND P. RESTREPO (2020): "Unpacking Skill Bias: Automation and New Tasks," in *AEA Papers and Proceedings*, vol. 110, 356–61.

ACEMOGLU, D. ET AL. (2012): "What does human capital do? A review of Goldin and Katz's The race between education and technology," *Journal of Economic Literature*, 50, 426–63.

AUTOR, D. H. (2017): "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).*

AUTOR, D. H., L. F. KATZ, AND M. S. KEARNEY (2008): "Trends in US wage inequality: Revising the revisionists," *The Review of economics and statistics*, 90, 300–323.

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

BEN-PORATH, Y. (1967): "The production of human capital and the life cycle of earnings," *Journal of political economy*, 75, 352–365.

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

BOWLUS, A., E. BOZKURT, L. LOCHNER, AND C. ROBINSON (2017): "Wages and employment: The canonical model revisited," Tech. rep., National Bureau of Economic Research.

BOWLUS, A. J. AND C. ROBINSON (2012): "Human capital prices, productivity, and growth," *American Economic Review*, 102, 3483–3515.

CASANOVA, M. (2013): "Revisiting the hump-shaped wage profile," *Unpublished mimeo, UCLA*.

CICCONE, A. AND G. PERI (2005): "Long-run substitutability between more and less educated workers: evidence from US states, 1950–1990," *Review of Economics and statistics*, 87, 652–663.

DEATON, A. AND C. PAXSON (1994): "Intertemporal choice and inequality," *Journal of political economy*, 102, 437–467.

GOLDIN, C. D. AND L. F. KATZ (2009): *The race between education and technology*, harvard university press.

GOUSKOVA, E. (2014): "Parameter estimates comparison of earnings functions in the PSID and CPS data, 1976–2007," *Economics Letters*, 122, 353–357.

GUVENEN, F., F. KARAHAN, S. OZKAN, AND J. SONG (2015): "What do data on millions of US workers reveal about life-cycle earnings risk?" Tech. rep., National Bureau of Economic Research.

HECKMAN, J. J., L. LOCHNER, AND C. TABER (1998): "Explaining rising wage inequality: Explorations with a dynamic general equilibrium model of labor earnings with heterogeneous agents," *Review of economic dynamics*, 1, 1–58.

HUGGETT, M., G. VENTURA, AND A. YARON (2011): "Sources of lifetime inequality," *American Economic Review*, 101, 2923–54.

JEONG, H., Y. KIM, AND I. MANOVSKII (2015): "The price of experience," *American Economic Review*, 105, 784–815.

KAMBOUROV, G. AND I. MANOVSKII (2009): "Accounting for the changing life-cycle profile of earnings," *Unpublished manuscript, University of Toronto*.

KATZ, L. F. AND K. M. MURPHY (1992): "Changes in relative wages, 1963–1987: supply and demand factors," *The quarterly journal of economics*, 107, 35–78.

KELLER, E. (2014): "The slowdown in American educational attainment," *Journal of Economic Dynamics and control*, 46, 252–270.

KONG, Y.-C., B. RAVIKUMAR, AND G. VANDENBROUCKE (2018): "Explaining cross-cohort differences in life-cycle earnings," *European Economic Review*, 107, 157–184.

LAGAKOS, D., B. MOLL, T. PORZIO, N. QIAN, AND T. SCHOELLMAN (2018): "Life cycle wage growth across countries," *Journal of Political Economy*, 126, 797–849.

LEMIEUX, T. (2006): "Postsecondary education and increasing wage inequality," *American Economic Review*, 96, 195–199.

MAGNAC, T., N. PISTOLESI, AND S. ROUX (2018): "Post-schooling human capital investments and the life cycle of earnings," *Journal of Political Economy*, 126, 1219–1249.

MCKENZIE, D. J. (2006): "Disentangling age, cohort and time effects in the additive model," *Oxford bulletin of economics and statistics*, 68, 473–495.

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

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

WELCH, F. (1979): "Effects of cohort size on earnings: The baby boom babies' financial bust," *Journal of Political Economy*, 87, S65–S97.

# Appendix A 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.

# Appendix B 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.

# Appendix C Cohort Grouping

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

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

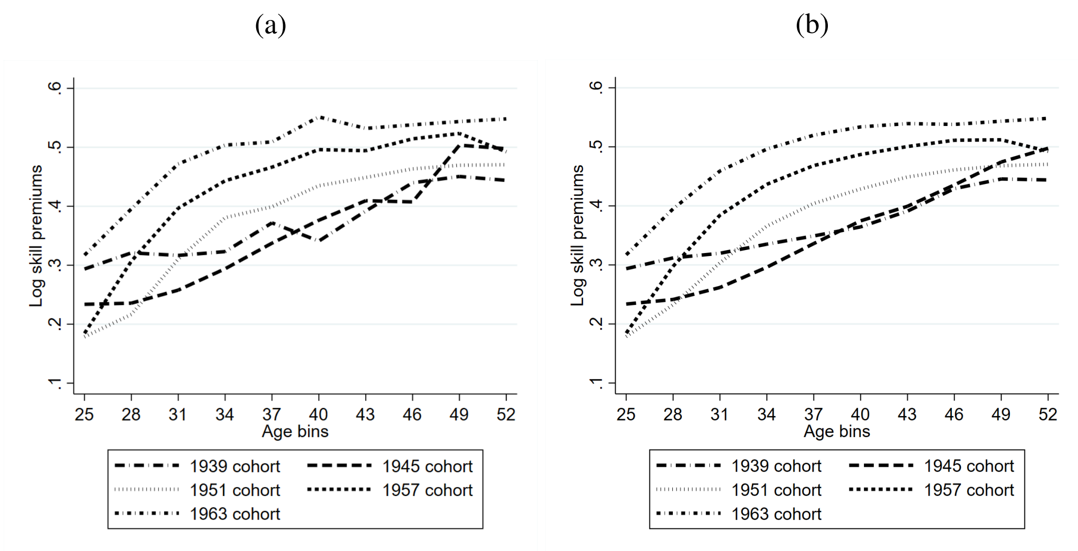
# Appendix D Robustness Check: Granular Bins

In Figure 5, 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 6. 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 5: Life-cycle Profiles of Skill Premiums (3-Year bin width)




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 6: 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.