"Golden Ages": A Tale of Two Labor Markets*

Hanming Fang† Xincheng Qiu‡

December 14, 2020
(*click here for latest version*)

Abstract

We document stark differences in the labor market outcomes between the U.S. and China, the largest two economies in the world, during the past 30 years. (1) The peak age in cross-sectional age-earnings profiles stays constant at around 45-50 years old in the U.S. but decreases sharply from 55 to 35 years old in China. (2) Age-specific earnings grow drastically in China but almost stagnate in the U.S. (3) The cross-sectional and life-cycle age-earnings profiles look remarkably similar in the U.S. but differ substantially in China. To address these facts, we provide a unified decomposition framework to infer life-cycle human capital accumulation, inter-cohort productivity growth, and human capital price changes over time, from repeated cross-sectional earnings data. We apply the framework to revisit several important and classical applications in macroeconomics and labor economics.

Keywords: Age-Earnings Profiles, Human Capital, Life Cycle
JEL Codes: E24, E25, J24, J31, O47

*We would like to thank Xiaodong Zhu for useful comments and suggestions. We would also like to thank the seminar participants at Fudan, Peking University, Penn, Wharton. All remaining errors are ours.

†Department of Economics, University of Pennsylvania, Ronald O. Perelman Center for Political Science and Economics, 133 S. 36th Street, Philadelphia, PA, United States, 19104; and the NBER. Email: hanming.fang@econ.upenn.edu.

‡Department of Economics, University of Pennsylvania, Ronald O. Perelman Center for Political Science and Economics, 133 S. 36th Street, Philadelphia, PA, United States, 19104. Email: qiux@sas.upenn.edu.

1 Introduction

The cross-sectional age-earnings profile is one of the most empirically examined objects in labor economics, dating back at least to Mincer (1974). A large and mature literature has confirmed a robust regularity of hump-shaped age-earnings profiles: earnings are low for young workers who have just entered the labor market, then rise with age, but at some point stop growing, and eventually decline after reaching the peak earnings age. In this paper, we dub such a peak age as "*golden age*", defined as the age group that achieves the highest average earnings in a cross section. For instance, the "golden age" in the United States has stayed at around 50 years old, meaning that 50-year-old workers have the highest average earnings among all age groups.

This paper starts with a systematic comparison of the age-earnings profiles between U.S. and China. We summarize the striking differences during the last thirty years, as the following three stylized facts:

1. The cross-sectional "golden age" stays stable at around 45-50 years old in the U.S. but continuously decreases from 55 to 35 years old in China.

2. Age-specific earnings almost stagnate in the U.S. but grow drastically in China.

3. The cross-sectional and life-cycle age-earnings profiles look remarkably similar in the U.S. but differ substantially in China.

In this paper, we seek to uncover the causes for the above differences between the two labor markets. To this end, we provide a framework to decompose the repeated cross-sectional age-earnings data into *experience*, *cohort,* and *time* effects, where:

- the experience effects capture how an individual's earning capacity grows with experience over his life-cycle;

- the cohort effects capture inter-cohort productivity growth, or, the relative human capital level of a cohort of workers at the time when they enter the labor market;

- the time effects capture the human capital rental prices at a given time, which of course, may change over time.

As is well-known (and we will show below), without further restrictions, these three factors cannot be separately identified due to perfect collinearity. David Lagakos, Benjamin Moll, Tommaso Porzio, Nancy Qian and Todd Schoellman (2018) (hereafter, LMPQS) present a state-of-the-art treatment of the experience-cohort-time identification issue. The identifying assumption we adopt in this paper is that there is no experience effect in late career, as implied by the standard human capital investment theory (Ben-Porath, 1967), where there is no human capital accumulation at the end of working life. In fact, this assumption is also consistent with several other prominent models of wage dynamics, such as search theories with on-the-job search (Burdett and Mortensen, 1998) and job matching models with learning (Jovanovic, 1979).[1] This identification idea is exploited originally by Heckman, Lochner and Taber (1998) (hereafter, HLT), and more recently also by, Huggett, Ventura and Yaron (2011), Bowlus and Robinson (2012), and Lagakos et al. (2018). We show that under this identifying assumption, we can use repeated cross-sectional age-earnings profiles to separately identify experience, cohort, and time effects, which in turn allows us to simultaneously account for the three stylized facts regarding the differences in the evolutions of the U.S. and China's labor markets in the last thirty years.

First, the "golden age" in a cross-sectional age-earnings profile is essentially determined by the race between life-cycle human capital growth (the experience effect) and the inter-cohort productivity growth (the cohort effect). When the experience effect dominates, the "golden age" tends to be old. When the inter-cohort productivity growth prevails, the "golden age" tends to be young. We find that, in China the rapid inter-cohort productivity growth wins the race against the experience effect, and as a result, the golden ages in China has experienced a gradual decline in the last thirty years. In contrast, in the U.S. the two forces form a stalemate, resulting in very stable golden ages.

Second, we find that the rental price to human capital (the time effect) is increasing much faster over the last thirty years in China than the U.S. Moreover, China experienced much higher inter-cohort human capital growth (the cohort effect) than the U.S. Both contribute to the much faster growth in age-specific earnings in China.

Lastly, we find that the cohort effect and the time effect are negligible in the U.S. compared to the experience effects. (In fact, we find that the real price to human capital is declining by about 1% per year in the U.S.) That is, the experience effect is the main driving force of both cross-sectional and life-cycle age-earnings profiles for U.S. As a result, they are close to experience effects and hence look similar to each other. In China, however, both the cohort and the time effects are substantial in the last thirty years. Thus, stationarity is lost and the life-cycle earnings profiles are very different from cross-sectional age-earnings profiles.

We also use our decomposition to fine tune some important accounting exercises in macroeconomics and labor economics. First, by teasing out human capital prices changes

[1] Rubinstein and Weiss (2006) provides an excellent review on these three classes of models of investment, search, and learning that explain life-cycle wage growth.

obtained from the time effects, our decomposition delievers a notion of effective human capital quantities, which constitutes both experience and cohort effects. This allows us to conduct a growth accounting exercise that properly accounts for the evolution of human capital. Adjusting for the changes of effective human capital, we obtain a series of estimated total factor productivity (TFP) growth lower than previous estimates. This idea has been conducted by Bowlus and Robinson (2012). Second, we further decompose human capital rental prices into the marginal product of human capital, which is affected by both TFP and capital intensity, and a wedge which we interpret as a proxy measure of labor market frictions. Our results show that most of the decreases in U.S. human capital rental prices are due to the increase in the wedge; while in China, half of the rise in the human capital price can be accounted for by increasing capital intensity, and one-third by the growth of its TFP. This provides a novel way of estimating wage markdown. Third, we also implement the same decomposition separately for college and high school workers. There we obtain an estimated series for skill-biased technological change where relative human capital quantities between high-skilled and low-skilled workers are allowed to change over time.

Related Literature. There is a long tradition in the economics literature that documents the age-earnings profiles. This line of research is so huge that we do not attempt to provide a thorough survey here. The age-earnings profile is an important input for many topics, including human capital, incentive contracts, labor supply, retirement, consumption and saving. More recent work includes Rupert and Zanella (2015) who document the life-cycle profiles separately for different cohorts, and find that the tracking between hour and wage profiles over the life cycle is broken for cohorts born between 1937 and 1946. Specifically, for cohorts born between 1937 and 1946, hours start to decline after age 55, while wages are still increasing until very late into one's 60s. Their finding calls for modifications to the standard life cycle model where hours profile tracks the wage profile closely. Throughout this paper, we focus on annual earnings and do not speak to implications on hours worked, mainly due to the lack of data on hours worked in UHS. Casanova (2013) focuses on wage and earnings profiles for older workers over age 50 facing retirement decisions. She finds that once the employment status of full-time or part-time is controlled, wage does not decline as one ages. There is indeed a one-time drop in wage when one transitions from full-time work to part-time work. The same pattern also holds for hours worked and hence earnings. The observed downward sloping part in a typical age-wage profile or age-earning profile is thus a result of increasing composition of part-time workers. Without proper information on hours worked in UHS, we cannot directly address the partial retirement issue. Instead, we use median regression techniques, which could minimize the potential bias under the assumption that a median worker is working full time.

The remainder of the paper is structured as follows. Section 3 lays out the theoretical framework and discusses issues on identification. Section 4 describes the main results from the decomposition and the applications. Section 5 extends the benchmark framework and presents additional results. Section 6 concludes and discusses future research directions.

2 Facts

2.1 Cross-Sectional Age-Earnings Profiles and "Golden Ages"

We use the March Current Population Survey (CPS) Annual Social and Economic (ASEC) Supplement extracted from IPUMS (Flood et al., 2018) as the primary dataset for the United States. CPS is the official source to produce many labor market statistics. The choice of sample period is to facilitate comparison with China, for which we only have access to data from 1986 to 2012.[2]

Figure 1a depicts the cross-sectional age-earnings profiles for male workers in the U.S. Each curve represents a cross section that pools five or four adjacent years. In the construction of each curve, we first perform a nonparametric kernel regression of annual labor earnings on age separately for each cross section, where the Epanechnikov kernel function and rule-of-thumb bandwidth estimator are applied, and then display the smoothed values with the 95% confidence intervals. To avoid potential effects of extreme values, 5% outliers are dropped (earnings in the top 2.5% and bottom 2.5% in each year). Individuals are weighted by the person-level ASEC weight. Figure 1a reveals that, (1) first, the "golden age" in the U.S. is relatively stable at around 50 years old during the past three decades; (2) second, the U.S. has witnessed little growth in age-specific mean real earnings.

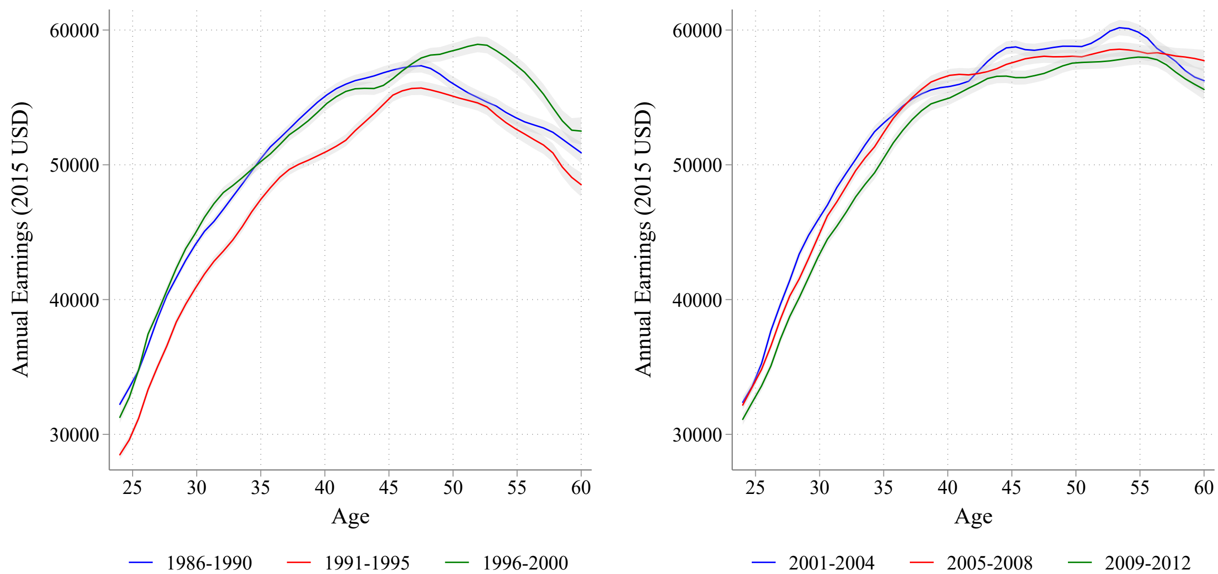
To study China's labor market, we use the Urban Household Survey (UHS) administered by the National Bureau of Statistics (NBS). UHS is the only nationally representative microdata covering consecutive years since the late 1980s. Although UHS is representative only for urban China, it is the most comparable survey for China to CPS.

In Figure 1b, we plot the cross-sectional age-earnings profile for Chinese male workers, using the same procedure as discussed before. There are several striking contrasts

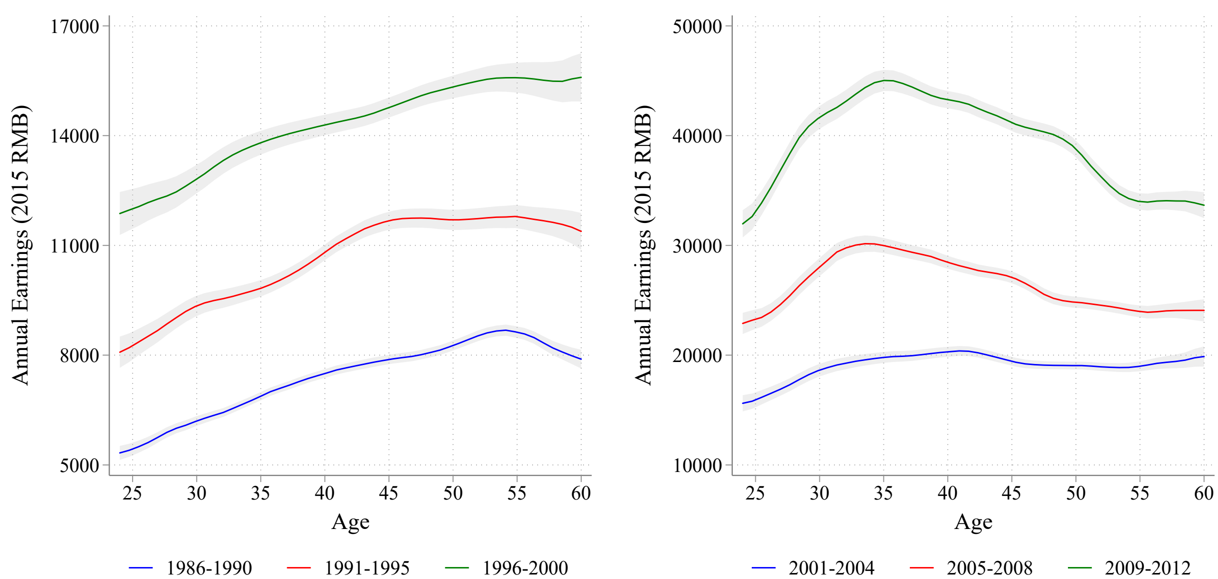
[2]Throughout this paper, a year refers to the year to which the income variable corresponds.

Figure 1: Evolution of Cross-Sectional Age-Earnings Profiles

(a) U.S.



(b) China



Notes: The top panel plots the cross-sectional age-earnings profiles of U.S. male workers, using March CPS from 1986 to 2012. The bottom panel plots the cross-sectional age-earnings profiles of Chinese Urban male workers, using UHS from 1986 to 2012. Each curve represents a cross section that pools adjacent years. The solid lines are kernel smoothed values and the gray shaded areas are the 95% confidence intervals. Note that the vertical scale of the left and right subgraphs in the bottom panel differs.

between Figures 1a and 1b. (1) First, Chinese workers have experienced dramatic increase in real earnings in the past 30 years for all age groups. It is reflected in the large vertical shifts of the age-earnings profiles for later cross sections. The earnings of Chinese urban male workers increased by nearly 6 folds. This is in marked contrast to the earnings stagnation in the United States. (2) Second, while the shape of the cross-sectional age-earnings profiles and hence the corresponding "golden ages" have stayed more or less constant in the U.S., the "golden age" of China is continuously evolving to younger ages. Prior to 2000, the age-earnings profiles of China had a familiar hump-shape with the "golden age" at around 55, although there already seems to be some signs of a declining "golden age" in 1996-2000. Between 2001 and 2004, the age-earnings profile is almost flat and humps at around age 40-45. After 2005, the "golden age" is 35 years old.[3]

To sum up, Figure 2 plots the evolution of the cross-sectional "golden ages" in the U.S. and China during 1986-2012. For each country and each year, we run a kernel regression of log earnings on age to predict age-specific earnings, and therefore obtain an estimated golden age in that year as the age achieving the maximal predicted earnings. Furthermore, we fit a linear time trend of the estimated golden age for each country. Figure 2 show clearly that in the U.S., the golden age has stayed constant at around 48 years old in the past thirty years, while in China there exhibits a clear downward trend in the golden ages from 1986-2012, decreasing from more than 55 years old to 35 years old.

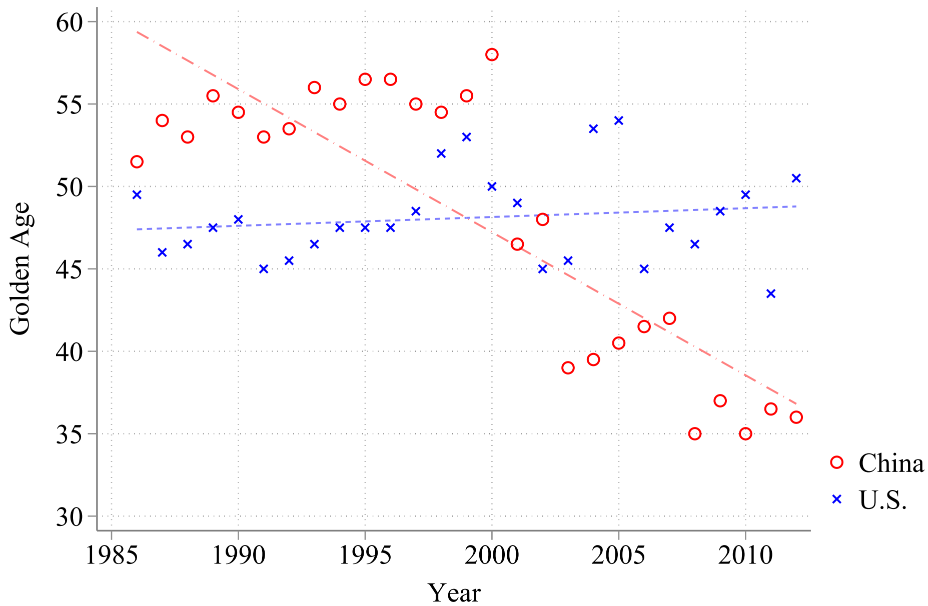
We make sure that the facts are a robust feature of the two labor markets. First, as pointed out before, by its design, UHS only covers urban households.[4] The stark difference in the evolution of cross-sectional age-earnings profiles between the U.S. and China, however, is not merely a result of the sample restriction on urban workers in UHS. In figure B.1, we restrict attention to CPS households that live in a metropolitan area, which is the closest notion to urban households in UHS. There is virtually no difference in the shape of age-earning profiles, although the level of earnings is on average higher for workers in metropolitan areas than those who are not in metropolitan areas, as one may expect.

Second, due to our limited access to the UHS microdata, we do not have all provinces covered consecutively in our sample. Because the main goal of this study is to investi-

[3] Cai et al. (2014) plot the income profiles in 2002 and 2007 using data from the Chinese Household Income Project (CHIP), and also notice an earlier arrival of peak earning age.

[4] Prior to 2002, UHS only covers households with local urban *hukou*. Although UHS started to include households without local urban *hukou* since 2002, the coverage is so low that non-local-*hukou* residents are under represented. See, for example, the discussion in Ge and Yang (2014).

Figure 2: Evolution of Cross-Sectional "Golden Age" in the U.S. and China



Notes: Figure 2 shows the evolution of the cross-section "golden age" in the U.S. and China. The blue cross marker denotes the estimated golden age in the U.S. and the red circle marker denotes the estimated golden age in China. The blue short-dashed line and the red dash-dot lines are the respective linear time trend in the evolution of the golden age in each country.

gate how the labor market evolves over time, it is crucial to provide a comprehensive set of evidence than spans a long period of time. So we do not drop any time periods in our main analysis, but verifies that our analysis is not affected by the regional coverage. We have a random subset of the UHS sample households with a representative coverage of provinces (see Table C.1). The only provinces that are included continuously throughout all the 27 years from 1986 to 2012 are Liaoning, Shanghai, Guangdong, Sichuan. Although there are only four such provinces, they constitute an arguably representative picture of the nation with a dispersed geographic coverage – the Northeast, East, South Central, Southwest, respectively. To mitigate the concern for representativeness, we replicate Figure 1b for a much larger set of 15 provinces covering all 6 regions[5] in Figure B.2 in the Appendix B, but the whole set could only be tracked from 1986 to 2009. The pattern barely changes. Prior to 2000, the cross-sectional age-earnings profiles have a familiar hump shape with a "golden age" of 50-55 years old. During early 2000s, the

[5]They are Beijing, Shanxi, Liaoning, Heilongjiang, Shanghai, Jiangsu, Anhui, Jiangxi, Shandong, Henan, Hubei, Guangdong, Sichuan, Yunnan, Gansu. They altogether span all 6 regions in China – North, Northeast, East, South Central, Southwest, and Northwest.

profiles are very flat after age 40. In 2007-2009, it already exhibits a very young "golden age" of 35-40 years old.

Finally, one natural question is whether the aforementioned pattern is a result of wages or hours worked. Though UHS does not collect information on hours worked for most years, we can partially answer this question for a sub-period from 2002-2006, where UHS does collect information on "total number of hours worked last month". A typical month contains about $30/\approx 4.286$ weeks, so we use this number to convert the monthly measure of hours worked to a weekly measure, in order to facilitate comparison with the variable "total number of hours usually worked per week over all jobs the year prior to the survey" in CPS. Figure B.3 shows that the age-hours profiles are almost on top of each other for these two labor markets after 25, although there is a disagreement for earlier ages between 18-25. This suggests that the patterns we document above is more likely to be about wages, rather than hours, at least for prime-age workers older than 25.

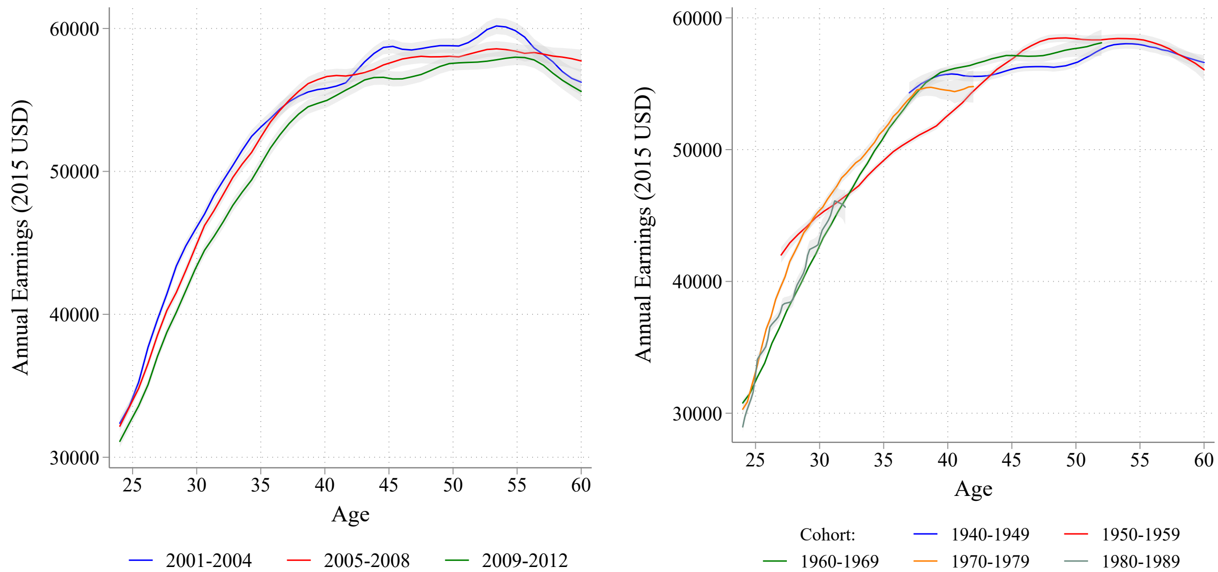
2.2 Cross-Sectional v.s. Life-Cycle Age-Earnings Profiles

Conceptually, a cross-sectional age-earnings profile, which summarizes earnings of workers of different ages at a given point of time, is a different notion to the life-cycle earnings profile, which tracks the earnings of a typical person over his life course. Thus one should not expect the cross-sectional age-earnings profiles to coincide with the life-cycle ones. In Figure 3, we reproduce the cross-sectional profiles from Figure 1 on the left, and plot the life-cycle earnings path of various birth cohorts on the right, with each curve representing a 10-year cohort bin. The top panel is for US and bottom for China.

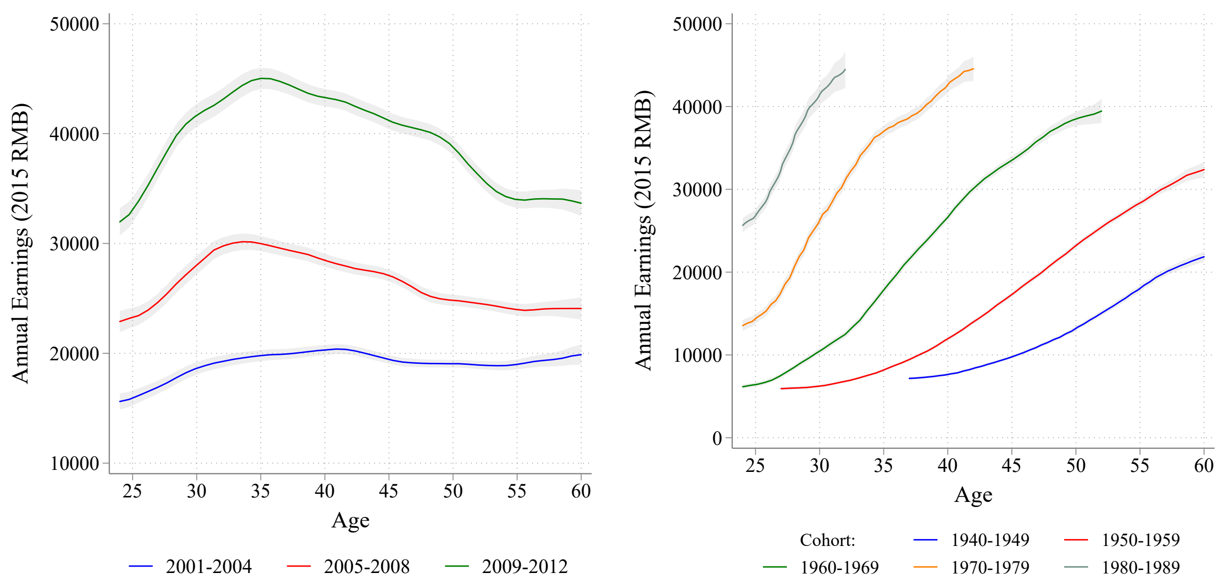
In the U.S. (Figure 3a), cohorts with year of birth expanding half a century share remarkably similar life-cycle earnings paths. Furthermore, life-cycle profiles on the right of Figure 3a resemble closely the cross-sectional profiles on the left (which is reproduced from the right panel of Figure 1a), in both its shape and level. In a stationary environment where the life-cycle profile does not vary across cohorts, the cross-sectional profiles and the life-cycle profiles essentially coincide with each other. In this economy, a 20-year-old worker who wants to predict his earnings 20 years later can simply take a look at the contemporary earnings of a 40-year-old worker. This provides a justification for voluminous previous works that use cross-sectional profiles as approximations to life-cycle patterns, although conceptually it is not immediately correct to read cross-sectional profiles as life-cycle patterns, in practice they are close enough to each other for the U.S. case. Put it in another word, stationarity is an reasonable assumption when studying the U.S. earnings profiles.

Figure 3: Cross-Sectional v.s. Life-Cycle Age-Earnings Profiles

(a) U.S.



(b) China



Notes: The top panel compares the cross-sectional and life-cycle age-earnings profiles for U.S. male workers, and the bottom panel for urban Chinese male workers. The left subgraph of each panel is the cross-sectional profiles reproduced from Figure 1a and 1b. The right subgraph of each panel shows the life-cycle age-earnings profiles, where each curve represents a 10-year cohort bin.

However, as shown in Figure 3b, the life-cycle patterns of different cohorts differ utterly for China. More recent cohorts enjoy both much higher earnings and steeper life-cycle earnings growth. These life-cycle profiles also demonstrates no resemblance at all to the cross-sectional profiles. In such a fast-growing economy, stationarity is not a valid approximation.

3 Framework

3.1 Setup

We consider a competitive interpretation of wages, where the observed level of wage reflects the product of the price of human capital and the quantity of human capital a worker supplies. Denote $W_{i,t}$ the wage of worker i at time t, $H_{i,t}$ the human capital owned by worker i at time t, and P_t the price of human capital at time t. We have

$$W_{i,t} = P_t \cdot H_{i,t}.$$

Note that the rental price of human capital is allowed to vary over time but restricted to be the same across individuals. This formulation assumes that the only heterogeneity among workers is in the quantity of human capital they possess, but not in the type of human capital. Put it differently, we are imposing a *scalar* representation of human capital. As a result, any potential complementarity between different types of skills is ruled out.[6] Write the above equation in logs,

$$w_{i,t} = p_t + h_{i,t}, \tag{1}$$

where lower case variables are in logs and upper case variables in levels.

A cohort of workers is labeled by the year when they enter the labor market. Consider a "representative" worker of cohort c at time t. Define the human capital supplied by the "representative" worker of cohort c at time t as the average human capital among all workers of cohort c at time t,

$$h_{c,t} = \mathbb{E}\left[h_{i,t} | i \in c, t\right].$$

By construction, the idiosyncratic component $\epsilon_{i,t} := h_{i,t} - h_{c,t}$ has a conditional mean of

[6]We extend this assumption later on.

zero (conditional on cohort c and time t). Therefore, we can rewrite equation (1) as

$$w_{i,t} = p_t + h_{c,t} + \epsilon_{i,t},$$

where $\mathbb{E}_i [\epsilon_{i,t} | i \in c, t] = 0$, for all c and t. The expectation is taken over individual workers i, for a given pair of c and t.

Since neither price nor quantity of human capital is observed, this specification leads to a non-identification problem. It is worth pointing out that a normalization alone does not solve the problem, because $\{p_t, h_{c,t}\}$ are not only observationally equivalent to $\{\lambda p_t, h_{c,t}/\lambda\}$ for any constant λ ("normalization"), but also observationally equivalent to $\{\lambda_t p_t, h_{c,t}/\lambda_t\}$ for any arbitrary series of $\{\lambda_t\}$ ("non-identification"). Consequently, without imposing further restrictions, we cannot tell how much of a wage change is due to human capital price changes and how much is due to human capital quantity changes.

We further decompose human capital into two component $h_{c,t} = s_c + r^c_{t-c}$, where $s_c := h_{c,c}$ is the level of human capital of cohort c when they enter the labor market at year c, and $r^c_k := h_{c,c+k} - s_c$ is the return to k years of experience for cohort c. This notation is without loss of generality. Obviously, $r^c_0 = h_{c,c} - s_c = 0$ by definition.[7] Using this notation, we can decompose (log) wages into time effects, cohort effects, and experience effects,

$$w_{i,t} = p_t + s_c + r^c_k + \epsilon_{i,t}, \tag{2}$$

with $\mathbb{E}_i [\epsilon_{i,t} | i \in c, t] = 0$, where (i) time effects p_t reflect the human capital prices, (ii) cohort effects s_c reveal cohort-specific human capital upon entry, and (iii) experience effects r^c_k are associated with the life-cycle human capital accumulation. Note that with $k = t - c$, we have perfect collinearity among year, cohort, and experience, which leads to non-identification.

A popular practice in the literature is to impose the returns to experience to be the same across cohorts, i.e., to restrict $r^c_k = r_k, \forall c$, which gives rise to a variant of equation (2):

$$w_{i,t} = p_t + s_c + r_k + \epsilon_{i,t}. \tag{2$'$}$$

The main benefit of restricting $\{r^c_k\}$ not to vary across cohorts is that we can get a complete estimated age profile even if every cohort is observed for only part of their life

[7]We do not model the labor market entry decision and abstract from the difference between age and experience. In other words, workers enter the labor market at the same age. Therefore, we use age and experience interchangably. Later on, we allow for difference between age and experience by introducing different levels of education. But we still assume that workers with the same level education enter the labor market at the same age. That is, conditional on education, we abstract away from any other potential difference between age and experience.

in the data.[8] We will first stick to this common practice as our baseline analysis, but discuss and relax the restriction later in Section

3.2 Cross-Sectional Age-Earnings Profiles and "Golden Ages"

Suppose one has constructed cross-sectional age-earnings profiles as we have done in Figure 1a and 1b. Each cross-sectional age-earnings profile for time t could be denoted $\{w(k;t)\}_{k=0}^{R}$, where k goes from 0 (entry) to R (retirement).[9] $w(k;t)$ is defined as the average (log) earnings of workers with experience k at time t

$$w(k;t) := \mathbb{E}_i\left[w_{i,t} | i \in c = t-k, t\right].$$

where the expectation is taken over individuals i for given time t and experience k (and hence cohort is also given by $c = t-k$). Due to the conditional mean zero property illustrated in the previous section, we could represent the cross-sectional age-earnings profiles as

$$w(k;t) = p(t) + s(t-k) + r(k).$$

We move the subscripts to inside the brackets in order to emphasize that human capital price p is a function of time t, cohort-specific human capital s is a function of cohort $c = t-k$, and returns to experience is a function of experience k (under the restriction of the homogeneous returns to experience across cohorts).

Assuming differentiability, the slope of the cross-sectional age-earnings profiles at time t is given by[10]

$$\frac{\partial}{\partial k} w(k;t) = \dot{r}(k) - \dot{s}(t-k), \tag{3}$$

which is positive if $\dot{r}(k) > \dot{s}(t-k)$ and negative if $\dot{r}(k) < \dot{s}(t-k)$. Note that both r and s are in logs, so the correct interpretation of $\dot{r}$ and $\dot{s}$ is the *rate* of life-cycle human capital growth and the *rate* of inter-cohort human capital growth, respectively. This observation immediately gives a characterization of the shape of a cross-sectional age-earnings profile:

[8] It is worth noting that this restriction is not solving the non-identification problem mentioned above. Even under this restriction, we still cannot isolate year effects, cohort effects, and experience effects without imposing additional assumptions, due to the perfect collinearity among year, cohort, and experience that $k = t - c$.

[9] Retirement decisions are abstracted and hence the retirement age is set exogenously throughout this paper.

[10] We present the result in continuous time for notational simplicity. The logic easily carries to a discrete time formulation as well, *mutatis mutandis*.

Proposition 1. *The cross-sectional age-earnings profile* $\{w(k;t)\}_{k=0}^{R}$ *is increasing (decreasing) in k when the rate of life-cycle human capital growth exceeds (falls behind) the rate of inter-cohort human capital growth.*

Though very straightforward, this proposition helps clarify the underlying forces determining the shape of cross-sectional age-earnings profiles. It states that the slope of a cross-sectional age-earnings profile is a result of the race between life-cycle human capital growth (experience effects) and inter-cohort human capital growth (cohort effects). If inter-cohort human capital growth is vast, then the older cohorts tend to earn less relative to more recent cohorts and hence then the cross-sectional age-earnings profiles tend to be decreasing. If life-cycle human capital growth dominates, then the older cohorts tend to have higher relative earnings and cross-sectional age-earnings profiles tend to be increasing.

Define the cross-sectional "golden age" at time t as

$$k^*(t) = \arg\max_{k\in[0,R]} w(k;t).$$

A characterization for the "golden age" follows immediately:

Corollary 1. *Suppose the cross-sectional age-earnings profile* $\{w(k;t)\}_{k=0}^{R}$ *is unimodal in k. The "golden age" at time t happens at experience* k^**, such that*

$$\dot{s}(t-k^*) = \dot{r}(k^*).$$

In other words, the cross-sectional "golden age" happens at the point where the speed of inter-cohort human capital growth exactly cancels out the rate of life-cycle human capital growth.

3.3 Cross-Sectional v.s. Life-Cycle Age-Earnings Profiles

Our notation also helps clarify the difference between cross-sectional and life-cycle profiles. Suppose one has constructed life-cycle age-earnings profiles as we have done in Figure 3. Denote the life-cycle age-earnings profile for cohort c by $\{\tilde{w}(k;c)\}_{k=0}^{R}$, where $\tilde{w}(k;c)$ is defined as the average (log) earnings of workers in cohort c with experience k

$$\tilde{w}(k;c) := \mathbb{E}_i\left[w_{i,t} | i \in c, t = c+k\right],$$

where the expectation is taken over individuals i for given cohort c and experience k (and hence time is also given by $t = c + k$). Due to the conditional mean zero property $\mathbb{E}_i [\epsilon_{i,t} | i \in c, t] = 0, \forall c, t$, we could represent the life-cycle age-earnings profiles as

$$\tilde{w}(k; c) = p(c+k) + s(c) + r(k).$$

The slope of the life-cycle age-earnings profiles for cohort c is given by

$$\frac{\partial}{\partial k} \tilde{w}(k; c) = \dot{r}(k) + \dot{p}(c+k). \tag{4}$$

Comparing equation (3) with equation (4) makes it clear how the cross-sectional profiles differ from life-cycle ones. For example, if both inter-cohort human capital growth and human capital price increase are fast in China (i.e., both $\dot{s}$ and $\dot{p}$ are large), then these two equations tell us that the cross-sectional profiles tend to be flat and the life-cycle profiles tend to be steep. If both inter-cohort human capital growth and human capital price changes are minor in US (i.e., both $\dot{s}$ and $\dot{p}$ are small), then we would expect the cross-sectional profiles to be close to life-cycle profiles. In fact, they should both approximate the path of returns to experience. Given the facts we have documented in Section 2, this tale serves a very promising description of what happened in the two labor markets in the past three decades.

3.4 Identification

Suppose one has access to a repeated cross-sectional dataset on earnings

$$\{w_{i,t}\}, \qquad i = 1, 2, \ldots, N; \quad t = 1, 2, \ldots, T,$$

where i refers to an individual observation and t to time. At each cross section t, the individual observations span a range of experience $k \in \{1, 2, \ldots, R\}$. Note that repeated cross-sections differ from panel data in that the pool of individuals can vary in different periods. For convenience, we reproduce equation (2′) here:

$$w_{i,t} = p_t + s_c + r_k + \epsilon_{i,t}. \tag{2′}$$

where p_t, s_c, r_k are vectors of time dummies, cohort dummies, and experience dummies with $k = t - c$. The residual satisfies conditional mean zero condition $\mathbb{E}_i [\epsilon_{i,t} | i \in c, t] = 0, \forall c, t$.

There are two issues that are worth mentioning. (i) First, *normalization* (or non-identification of levels). For each indicator vector, we have to omit one category as the baseline group. All estimates for the indicator vectors are relative terms to the baseline group. For example, in the main analysis, we set the baseline group to be cohort 1935, year 1985 and experience 0. The log earning of the baseline group is loaded on a constant term. (ii) Second, *non-identification* (of first differences). By definition, $k = t - c$ holds. Due to the perfect collinearity among them, cohort, experience, and time effects cannot be separately identified without further restrictions.[11]

We generally follow the approach by LMPQS, which in turn builds on the insights of Deaton (1997) and HLT. We review in detail the literature related to the age-cohort-time identification issue in Appendix D.1 but summarizes briefly the main message here. McKenzie (2006) points out that second differences (and higher order differences) of these effects can be identified without any assumption, while first differences of these effects can be identified with one restriction. Deaton (1997) views time effects as to capture cyclical fluctuations and picks an assumption that time effects are orthogonal to a linear trend and sum up to zero. In this way, all growth is due to cohort. LMPQS also considers an opposite extreme that all growth is due to time, and a specification in between with cohort and time each contribute half. HLT exploits predictions from economic theory along the lines of Ben-Porath (1967), where the optimality of human capital investment from a life-cycle problem implies that there should be little human capital accumulation towards the end of career. This provides a natural choice of restriction such that the experience is zero in the last few years of working life, which is also considered by LMPQS. Such identification idea has also been exploited by Bowlus and Robinson (2012) and Huggett, Ventura and Yaron (2011).

To see how this assumption facilitates identification of human capital prices (time effects), cohort-specific initial human capital (cohort effects), and human capital accumulation paths (experience effects), we discuss the intuition of the identification. Consider a cohort at a period of time such that this cohort is right at the end of their working life. For this cohort, the change in wages in that stage of life is coming solely from the change in human capital prices (time effects), as there is no human capital accumulation due to

[11] In practice, there might be cases where they are not perfectly collinear. For instance, some surveys provide information on the whole employment history. Then one would be able to construct the actual years of experience, by subtracting the nonemployment periods, instead of the potential years of actual experience that are typically imputed. Therefore, variation in the employment history can break the perfect collinearity such that individuals with the same labor market entry year may end up with different levels of experience at a given point of time. In this case, however, cohort, experience, and time are still typically highly correlated. We are facing an issue of near multicollinearity and the standard OLS estimator will generate imprecise estimates.

the identifying assumption, and there is no cohort effect since we are fixing a particular cohort. Following the same procedure repeatedly for many cohorts at the end of their working life will identify a time series of human capital prices. With the time effects at hand, we can readily identify cohort effects and experience effects. A concrete notation may illustrate the identification even more clearly. Suppose one assumes there is no human capital accumulation (say) from R to $R+1$ years old.[12] First, comparing the wages of R-year-old workers in year t and $(R+1)$-year-old workers in $t+1$ identifies the time effect from t to $t+1$ because (1) we are comparing the same cohort so by definition there is no cohort effect, and (2) by assumption there is no experience effect. Second, comparing the wages of a-year-old workers in t and $(a+1)$-year-old workers in $t+1$ provides information for the experience effect from a to $a+1$ because (1) we are again comparing the same cohort so there is no cohort effect, and (2) we have already backed out the year effect from t to $t+1$. Finally, further comparing the wages of a-year-old workers and $(a+1)$-year-old workers in t gives the cohort effect from cohort $c = t - a - 1$ to cohort $c+1$ because (1) they are in the same year so there is no time effect, and (2) we have already backed out the experience effect from a to $a+1$.

In general, the HLT identification approach requires one to pick her preferable values for a "flat region" of experience, for which there is no experience effect. It could also be extended to allow for a human capital depreciation rate. We acknowledge that either input is somewhat arbitrary and cannot be inferred internally from data (which is rooted in the non-identification problem discussed before). In particular, HLT assume a zero human capital depreciation rate, which is supported by Browning, Hansen and Heckman (1999). We follow this assumption, as many other papers studying life-cycle human capital accumulation do (e.g., Kuruscu, 2006). The choice for the flat spot region would also be *ad hoc*. We follow LMPQS by considering 40 years of experience and assuming there is no experience effects in the last ten years in the baseline specification. Bowlus and Robinson (2012) attempt to determine the flat regions more carefully and prefer the flat spot age ranges to be 50-59 for college graduates and 46-55 for high school graduates. Our choice of flat region is largely overlapped with theirs.

[12] The actual identifying assumption and algorithm is more sophisticated, but we provide the intuition in a nutshell here for transparency. See Appendix for detailed explanation on the iterative procedure in implementation.

4 Decomposition

Figure 4 performs an HLT decomposition of earnings among experience, cohort, and time effects. Specifically, we estimate the experience effects (relative to labor market entry) in 5-year bins, cohort effects (relative to the 1940-1944 birth cohorts) in 5-year bins, and year effects (relative to 1986) year by year. Figure B.4 reports the model fit, which shows that our simple decomposition framework turns out to match data pretty well.

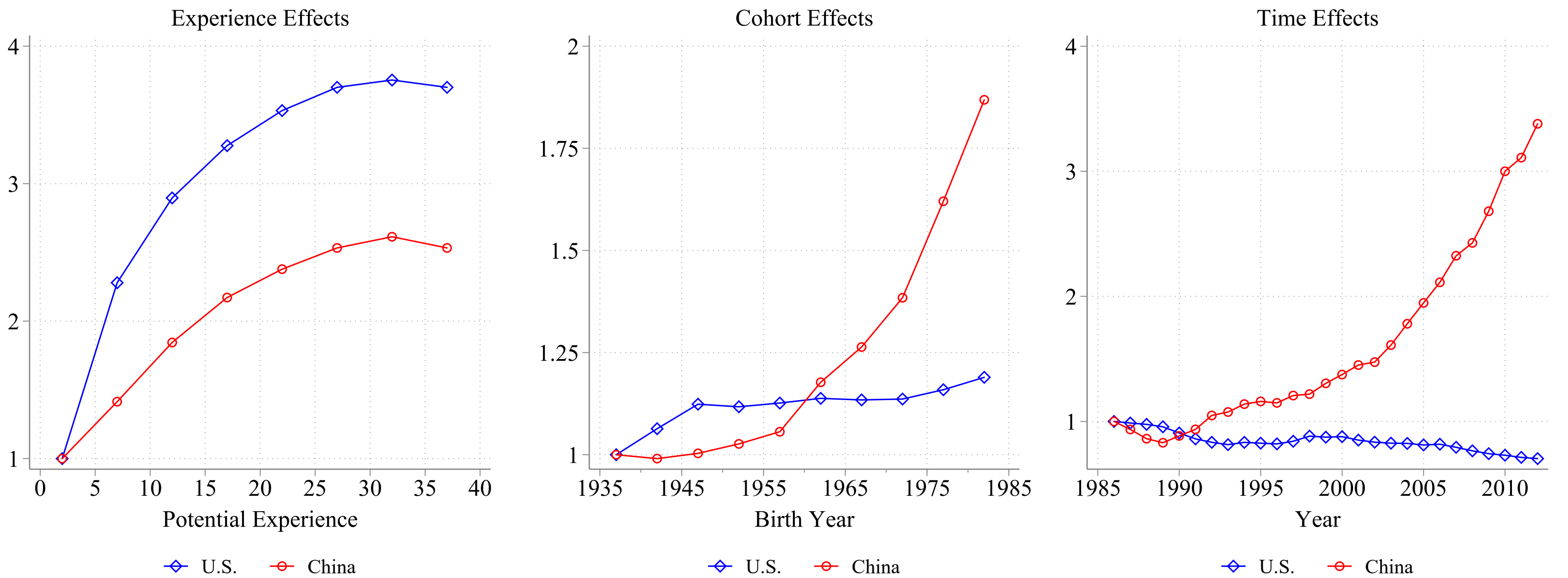
We will discuss each part in detail, but the main messages emerge very clearly: (1) Chinese workers have a 150% increase in earnings over the life course of 40 years working experience, while US workers have a 270% increase, which is nearly twice as higher. (2) There is only a 20% increase in cohort-specific productivity over 50 years of cohorts in US, most of which happened from cohort 1935 to cohort 1950. The inter-cohort productivity growth is almost 90%, most of which happened only since cohort 1960. (3) The time effect grows to more than three folds in China from 1986 to 2012, while it is negligible in the US. In anything, it declines at a rate of about 1% per year.

Before turning into the detailed discussion on interpretations and implications of the decomposition, we showcase that our decomposition result is robust to alternative specifications in Table 1. First, the pattern is by no means driven by regional differences of a particular set of locations. To show that, we control for state fixed effect for US and province fixed effect for China in Row 2. We also restrict attention to the 4 provinces that are covered in the UHS sample 2010-2012 in Row 3.

Second, we consider alternative definitions for potential experience. In the baseline, potential experience is imputed as experience $:= \min\{\text{age} - \text{edu} - 6, \text{age} - 18\}$ following LMPQS. That is, workers with more than 12 years of schooling are assumed to start schooling at 6 years old and enter the labor market after they finish schooling, and workers with fewer than 12 years of schooling are assumed to enter the labor market at 18 years old. We consider an alternative, and simpler definition for potential experience as experience $:= \text{age} - 20$ in Row 4. Since UHS provides information on the actual labor market entry year (when the respondent started the first job), we also consider experience as experience $:=$ current calendar year $-$ year of first job for China in Row 5.

Third, we examine whether our results are robust to alternative restrictions imposed by the HLT method. In Row 6, we consider an alternative flat region where there is no experience effect in the last five years. In Row 7, we assume there is a human capital depreciation rate of 1% per year in the last five years. In Row 8, we drop older samples and restrict attention to up to 35 years of experience, and assume a flat region in the

Figure 4: Decomposition



Notes: Figure 4 shows the decomposition results of experience, cohort, and time effects in US (blue diamond) and China (red circle), under the baseline specification.

Table 1: Experience, Cohort, Time Decomposition for U.S. and China

	Experience Effect (0-39)		Cohort Effect (1935-1984)		Time Effect (1986-2012)	
	U.S.	China	U.S.	China	U.S.	China
1. Baseline	3.70	2.53	1.19	1.87	0.70	3.38
2. State/province FE	3.71	2.53	1.19	1.78	0.71	2.96
3. Four provinces	/	2.37	/	1.79	/	3.27
4. Experience = Age − 20	3.24	2.55	1.20	1.84	0.85	3.56
5. Years since first job	/	2.31	/	1.71	/	3.92
6. Alternative flat region	4.10	3.18	1.36	2.52	0.65	2.82
7. Depreciation rate	2.87	2.22	0.86	1.57	0.86	3.76
8. 35 years of experience	3.46	2.10	1.03	1.38	0.76	4.15
9. Median regression	3.91	2.11	1.21	1.42	0.60	3.65
10. Controlling education	3.39	2.35	1.04	1.47	0.84	3.64
11. Hourly wage	1.84	/	1.03	/	0.80	/

Notes: This table reports various robustness results of the experience, cohort, time decomposition for US and China. The first row is the baseline result as discussed in the main text. Row 2 controls for state fixed effect for US and provincial fixed effect for China, and Row 3 focuses on the 4 provinces that are covered in the UHS 2010-2012 sample. Row 4 considers an alternative definition of potential experience as age minus 20, and Row 5 as years since the first jobm, which is only availble in UHS but not in CPS. Row 6-8 considers alternative input restrictions of the HLT method. Row 6 assumes no experience in the last 5 years, Row 7 assumes a human capital depreciation rate of 1% per year in the last 5 years, and Row 8 drops the sample with more than 35 years of experience. Row 9 performs a quantile regression at the median. Row 10 controls for years of schooling. Row 11 considers hourly wage for full-time workers in US.

last five years within that range. Although the magnitude of experience effects vary somewhat across specifications, as recognized by LMPQS, the general pattern we focus on is not affected by the specification, especially the comparison of the two labor markets.

Fourth, we consider in Row 9 the effects in terms of the median instead of the mean. Medians are less sensitive than means to outliers or missing values, both of which are quite common in earnings data. Median earnings are less likely to be affected by the evolving inequality in the top or bottom. Furthermore, average annual earnings, which we are forced to look at due to data limitation, is a combination of wages and hours. Medians also help in the sense that median men within each group may work more similar hours. Hence in Row 9, we perform a quantile regression estimating how conditional median earnings effects of experience, cohort, and time.

Fifth, our goal here is not to identify the "causal effect" on earnings but rather an accounting exercise. As a first step, we do not control for education. But we do separately consider college and high school groups in Section 5.1, which essentially allows college workers and high school workers to have heterogeneous types of skills. We provide in Row 10 as a robustness check the specification with years of schooling controlled. As expected, the cohort effect of China has decreased in this specification, since an important part of inter-cohort productivity growth is coming from the increasing overall level of education. That said, there is still a large increase in cohort effect even after education is controlled. This suggests after teasing out the compositional changes of education (between-group effects), there is still increasing quality of education (within-group effects). We will revisit the discussion of different education groups in Section 5.1.

Finally, since we do not have information on hours worked in UHS, we restrict attention mostly to earnings for US as well, for fair comparison. Nevertheless, we report in Row 11 the decomposition result using hourly wage for full-time male workers in CPS. The experience effects are much smaller than previous specifications using earnings, because workers increase hours a lot during the first few years since labor market entry (see Figure B.3 for direct evidence). That said, the cohort effect and time effect are barely changed.

4.1 Experience Effect: Life Cycle Earnings Capacity Growth

Consistent with the finding in Lagakos et al. (2018); Islam et al. (2018) that developed countries have higher returns to experience than developing countries, we find that the U.S. exhibits higher experience effects than China, as shown in the left panel of Figure

4. For an average American male worker, the human capital supplied at the end of his working life will be nearly 4-folds of his initial human capital supplied upon entry into the labor market. In China, the accumulated return to experience for the most experienced male workers is about 2.5 times the most inexperienced ones.

The magnitudes are not directly comparable to the result for US reported by LMPQS, however. The outcome variable they are concerning is hourly wage, constructed as labor earnings divided by the number of hours worked, while we are looking at annual earnings. As can be seen from Figure B.3, there is a large hours increase (or part-time to full-time transition) for the very young workers in the US. We provide an additional decomposition using hourly wage in Figure B.5 and Row 11 of Table 1. The result is consistent with the finding by LMPQS, which reassures us the validity of our decomposition.

In the classical life-cycle human capital accumulation literature, pioneered by Ben-Porath (1967); Mincer (1974), life-cycle earnings are interpreted as the amount of human capital supplied to the employers. In those models, earnings are increasing over the life cycle because (1) workers accumulate human capital to enlarge their human capital capacity, and (2) workers will invest less and hence contribute a larger fraction of their capacity to work when it approaches the end of their career.[13] An implicit assumption, when wage changes over the life cycle are interpreted as changes in the quantity of human capital supplied, is that the price of human capital is constant in different periods over the life cycle. Formally, only when assuming $P_t \equiv P, \forall t$, we have

$$\frac{W_{c,t_1}}{W_{c,t_2}} = \frac{P_{t_1} \cdot H_{c,t_1}}{P_{t_2} \cdot H_{c,t_2}} = \frac{H_{c,t_1}}{H_{c,t_2}}.$$

The considerable time effects estimated from our decompositon suggest that it cannot be an innocuous assumption for the case of fast-growing economies like China, although it is a rather good approximation for US.

Although we take a simple abstraction to model wages as from a competitive environment with perfect information, and hence interpret experience effects are life cycle human capital accumulation, it is worth pointing out that there are models consistent with the estimated experience effects. For instance, one could introduce search frictions

[13]In Ben-Porath (1967)'s framework, time devoted to working and learning are distinct concepts within the model, but a usual dataset cannot distinguish them. One has to take a stand on how much of the measured hours worked reflects time spent on working and investing. For example, Huggett, Ventura and Yaron (2011) assume that the measured hours worked is only work time and does not include training/learning time. One merit of focusing on annual earnings here is to avoid the measurement challenge of time allocation on working and training.

and allow for on-the-job search (e.g., Burdett and Mortensen, 1998). There the experience effects reflect workers climbing up the job ladder thanks to the arrival of new job offers. Alternatively, one could introduce information frictions in a job matching model (e.g., Jovanovic, 1979). There the experience effects reflect workers Bayesian learning the match quality.

How can we explain the steeper returns to experience in U.S. than in China (or more generally, in developed countries than in developing countries)? LMPQS concludes that evidence does not support long-term contracts as an important driver, but they do find human capital and search frictions are consistent the moments they look at. We propose another new, potential explanation for why experience effect is higher in U.S. than in China that relies on the multidimensionality of skills, where different skills may have different speeds of being accumulated. The investigation of this hypothesis is beyond the scope of the current paper, and we leave this direction for future research when there is suitable data to study heterogeneous distributions of multidimensional skills and skill requirements across countries.

4.2 Cohort Effect: Inter-Cohort Productivity Growth

Cohort effects capture the inter-cohort growth of initial human capital upon entry into the labor market. Since the life cycle human capital accumulation is imposed to be the same across cohorts in the baseline analysis, the same numbers also reflect the inter-cohort growth of human capital at any given age, or the life-time human capital. The middle panel of Figure 4 shows that China has experienced rapid human capital growth for subsequent cohorts. While U.S. workers' human capital increase by only about 20% in half a century of cohorts, the most recent cohort in China more than doubles the human capital as their counterparts 50 years ago. The cohort effects may come from that later cohorts receive more and/or higher-quality education, stay in better health conditions, or be equipped with the skills to perform more recent vintages of technologies.

Despite the rapidness of inter-cohort growth in China, the growth is unevenly shared among different cohorts. Most of the growth is reaped by workers born after 1960, while a whole generation prior to that witnessed very little human capital growth.

4.3 Time Effect: Human Capital Price Changes

We interpret the year effects in the right panel of Figure 4 as changes in the rental price to human capital. Human capital price in 2012 has increased to about 3.5 folds its level in

1986 in China, while there is little change in human capital prices in the U.S. If anything, the human capital price in the U.S. decreases by 30% from 1986 to 2012, which is around a 1% decline per year.

4.3.1 Growth Accounting

One application of the estimated human capital price change series is to fine tune a growth accounting. Consider an aggregate production function $Y_t = A_t K_t^{\alpha_t} H_t^{1-\alpha_t}$, where Y_t is the aggregate output, K_t the aggregate physical capital, H_t the aggregate human capital, A_t the total factor productivity (TFP), and α_t the share distribution parameter. Note that all elements are allowed to depend on time t. Denote lower case letters the corresponding variables in *per worker* terms, i.e., $x := X/L$, where $X = Y, K, H$ and L is the total number of workers. The output per worker can be expressed as $y_t = A_t k_t^{\alpha_t} h_t^{1-\alpha_t}$. Taking logs and differentials, we have

$$\mathrm{d}\ln y_t = \mathrm{d}\ln \tilde{A}_t + \alpha_t \,\mathrm{d}\ln k_t + (1-\alpha_t)\,\mathrm{d}\ln h_t,$$

where $\mathrm{d}\ln \tilde{A}_t := \mathrm{d}\ln A_t + (\ln k_t - \ln h_t)\,\mathrm{d}\alpha_t$.

We obtain four annual data series for each country: (1) real GDP Y_t, (2) capital stock K_t, (3) number of persons engaged L_t, and (4) share of labor compensation in GDP[14] s_t, all of which are from the Penn World Table 9.0 (Feenstra, Inklaar and Timmer, 2015) provided by the Federal Reserve Bank of St. Louis website.[15] We divide the real GDP Y_t and capital stock K_t by the number of workers L_t, to construct output per worker y_t and capital stock per worker k_t for each year t. We compute wage bills as the product of labor share and GDP $W_t = s_t Y_t$ and average the per worker counterpart $w_t = s_t y_t$. Using the price change series obtained from our previous decomposition[16], we could therefore get an estimated series for changes in human capital per worker

$$\mathrm{d}\ln h_t = \mathrm{d}\ln w_t - \mathrm{d}\ln P_t.$$

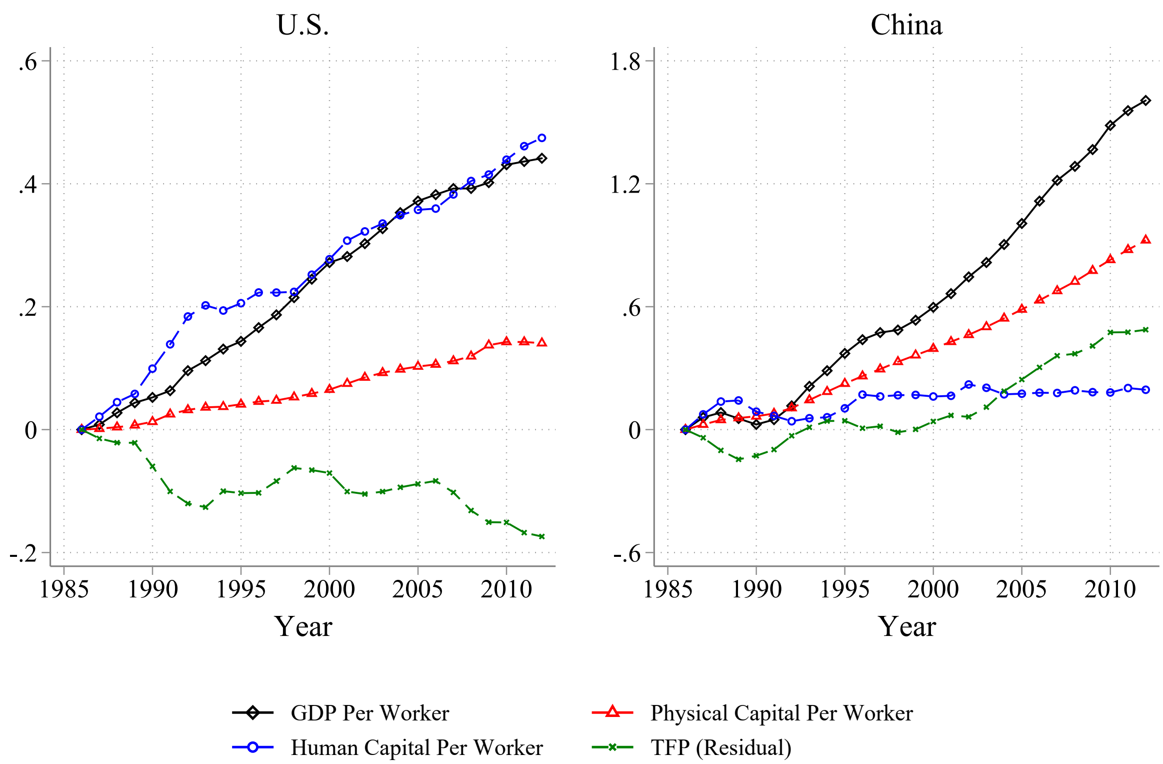
Under the competitive framework, the labor share is equal to $1-\alpha_t$, and hence we calibrate the distribution parameter α_t to $1 - s_t$. Taking stock, TFP changes could be measured as a residual. In fact, the exact term we obtain from the residual is $\mathrm{d}\ln \tilde{A}_t :=$

[14] The series on the share of labor compensation in GDP for China starts from 1992. We therefore are forced to impute the labor share between 1986 and 1991 to the same level of 1992.

[15] https://fred.stlouisfed.org/categories/33402

[16] For our estimated series of human capital price changes from male earnings data to apply to the growth accounting, one needs to assume that the human capital price *changes* (not necessarily *levels*) are the same for males and females.

Figure 5: Growth Accounting



Notes: This figures decomposes the growth in GDP per worker into contributions of physical capital per worker, human capital per worker, and TFP.

$\mathrm{d}\ln A_t + (\ln k_t - \ln h_t)\,\mathrm{d}\alpha_t$. However, our decomposition can only deliver changes but not levels. As a consequence, we cannot obtain the levels of h_t and are not able to distinguish $(\ln k_t - \ln h_t)\,\mathrm{d}\alpha_t$ from $\mathrm{d}\ln A_t$. In practice, as long as the annual labor share change $\mathrm{d}\alpha_t$ is small (it indeed is), this serves a reasonable approximation to TFP changes. Such approximation is commonly adopted in growth accounting (e.g., Fernald, 2014). We present the contribution of each source (including physical capital per worker, human capital per worker, and TFP) to the growth GDP per worker in Figure 5.

The growth accounting results in Figure 5 suggest little TFP growth in the U.S during the past 25 years, which is consistent with the productivity slowdown view (e.g., Fernald, 2015) in the same period. But our results give even negative estimates for TFP growth. This is understandable because the notion for human capital here properly incorporates inter-cohort human capital improvements and life-cycle human capital accumulation, and TFP is a model-based concept. In other words, part of the improvements in human capital was attributed to TFP growth in previous estimates that do not fully adjust for human capital growth. For China, TFP increases by around 50% during the same period, almost all of which are gained after the late 1990s and early 2000s, during many prominent economic reforms took place, such as SOE privatization, trade liberalization, and massive internal migration.[17] This is consistent with Zhu (2012), according to whom the average annual total factor productivity growth in nonagricultural sector is 2.17% and 0.27% for nonstate and state sectors during 1988-1998, but is 3.67% and 5.50% for nonstate and state sectors during 1998-2007. The numbers are not necessarily comparable as TFP are model-based notions, but we both find larger TFP growth arose after 2000.

Regarding the contributions of physical capital per worker, human capital per worker, and TFP to the GDP per worker growth, all growth in output per worker comes from factor growth, where physical capital accounts for about one quarter and human capital explain the remaining three quarters of net growth. The story is quite different in China, where physical capital is responsible for more than one half, and TFP for about one third.

There are several similar attempts to account for the human capital input in the aggregate production function. Hall and Jones (1999) presume that aggregate human capital is $H_i = \exp\{\phi(E_i)\} L_i$, where $\phi(E_i)$ reflects the efficiency units of labor with E years of education (relative to labor with no schooling). In practice, they estimate $\phi'(E)$ as the returns to schooling from a standard Mincerian wage regression. Bils and Klenow (2000) also exploit Mincerian returns, but extend the focus on education to include experience as well. In addition, they introduce interdependence on the human capital of older cohorts to capture the idea of impacts from teachers. Such an approach based on Mincerian rate of return to schooling to measure human capital typically finds cross-country differences in output per worker are largely driven by differences in TFP. Manuelli and Seshadri (2014), instead of relying on the Mincer equation, calibrate a model of human capital acquisition with early childhood development, schooling, and on-the-job training, and compute human capital stocks by evaluating the human capital production at the individual optimum under equilibrium prices. They find a larger role for human capital as well as a smaller role for TFP in explaining the cross-country differences in output per worker. The closest to our exercise here is Bowlus and Robinson (2012), which is also using the HLT identifying assumption to tease out human capital

[17]For example, state-owned enterprises (SOEs) were involved in waves of privatization and restructuring since mid-1990s under the slogan "grasp the large and let go of the small" to privatize small or medium-sized SOEs while retaining control of large ones. From 1995 to 2001, there were an estimated 34 million workers that were officially registered as laid off from the state sector. The privatization of SOEs contributes to the overall productivity of the economy (e.g., Chen et al., 2019). Another milestone event is that China joined the WTO in 2002. Trade liberalization is often believed as one of the key drivers of firms' productivity and hence China's rapid growth (see Brandt et al., 2017, for evidence)

price changes.

4.3.2 Understanding Human Capital Prices

A natural question is how we should understand human capital prices. Consider the same production function as before. In a competitive factor market, human capital price equals its marginal product

$$P_t = (1-\alpha_t)\, A_t\, (k_t/h_t)^{\alpha_t}.$$

Hence the observed changes in price to human capital are a combination of changes in TFP, factor share distribution parameter, and capital intensity (physical-capital-to-human-capital ratio). To account for why the U.S. and China exhibit opposite patterns in human capital prices, we decompose the human capital price changes into these three factors by taking logs and differentials:

$$\mathrm{d}\ln P_t = \mathrm{d}\ln \tilde{A}_t + \mathrm{d}\ln(1-\alpha_t) + \alpha_t\,\mathrm{d}\ln\left(\frac{k_t}{h_t}\right),$$

where $\mathrm{d}\ln \tilde{A}_t$ is defined in the same way as before $\mathrm{d}\ln \tilde{A}_t := \mathrm{d}\ln A_t + (\ln k_t - \ln h_t)\,\mathrm{d}\alpha_t$.

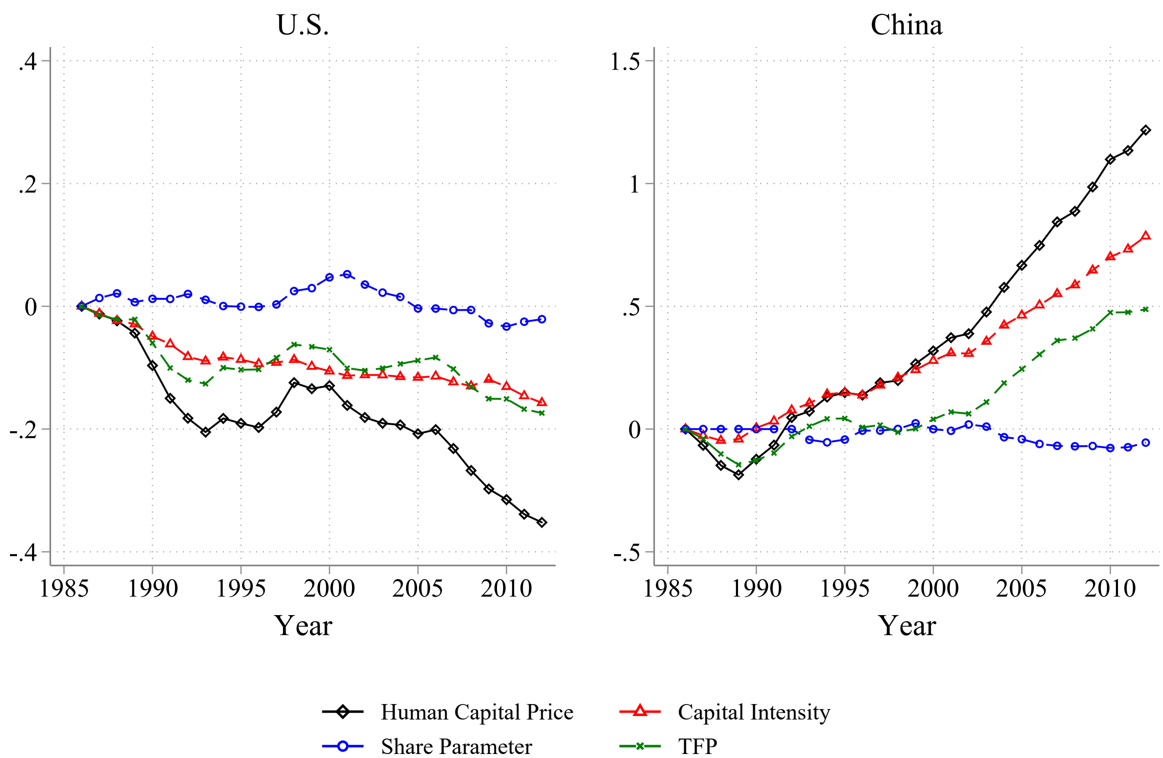
The decomposition is shown in Figure 6. In both countries, there is a declining labor share, as documented by Karabarbounis and Neiman (2013), mostly starting from 2000. However, its contribution to human capital price changes is minor in both countries. This is because the percentage change in labor share is rather small. About half of human capital price decrease in the US is due to TFP, and the other half due to decreasing capital intensity coming from a relatively fast improvement in human capital. In China, two-thirds of the rapid increase in human capital price is coming from increasing capital intensity, and one-third can be explained by improving TFP.

5 Discussions

5.1 Heterogeneous Human Capital by Education

In the previous baseline estimation, we assume that human capital is homogeneous so that every worker's skill can simply be represented by a single index indicating the level of efficiency units. We now extend the framework to allow for human capital to have different types. For example, college and high school graduates may possess different types of skills that are not perfect substitutes. It is straightforward to extend

Figure 6: Decompose Human Capital Price Changes



Notes: This figures decomposes changes in human capital price into contributions of TFP, share parameter in the production function, and capital intensity (physical-capital-to-human-capital ratio).

our framework to allow for heterogeneous human capital across education groups so that $W^e_{it} = P^e_t \cdot H^e_{it}$, where $e \in \{\text{cl}, \text{hs}\}$ refers to an education group ("cl" stands for college and "hs" high school). We then perform the decomposition as discussed before but now separately for college workers and high school workers. The only restriction is that for both college workers and high school workers, there is no additional skill accumulation from experience in the last two experience group. Since our imputation of potential experience assumes that college graduates start to gain experience from 22 years old and high school graduates start to gain experience from 18 years old, this is essentially assuming college graduates do not have additional returns to experience in 52-61 years old and high school graduates in 48-57 years old. This is largely consistent with Bowlus and Robinson (2012), who argued that a reasonble choice for the flat spot area is around 50-59 for college graduates and 46-55 for high school graduates. College and high school workers are allowed to have different paths of life-cycle human capital growth, different inter-cohort human capital growth, and different time series of human capital price changes. The results are presented in Figure 7.

First, within an education group, the returns to experience are still higher in the U.S. than in China. Within a country, the experience effects are larger for college workers than high school workers. This is consistent with findings documented by a long literature that earnings growth tends to be higher for workers with more education. The difference, however, is much smaller compared to the difference in cohort effects that we are turning to.

Second, China and U.S. exhibit very different patterns of cross-education comparisons in cohort effects. For the U.S., we find that the inter-cohort productivity growth is large and positive for college graduates, while it is even negative for high school graduates. This finding echoes the fanning out in inequality documented by Acemoglu and Autor (2011). In China, both education groups exhibit cohort-to-cohort improvement in human capital, but the inter-cohort growth is particularly high for college graduates. It is also interesting to note a decline of cohort-specific human capital that happened to 1980-1984 birth cohort college graduates. This is not surprising if one links to the institutional background this cohort experienced. The Chinese government expand college enrollment at a large scale in 1999. In the following years, the expansion was unprecedentedly massive.[18] As a much large fraction of this cohort can enroll in college, thanks to the higher education expansion, the selectivity of college decreases largely, and thus the average innate ability of college students for this cohort may be lower. It is very reassuring that our decomposition picks up this pattern.

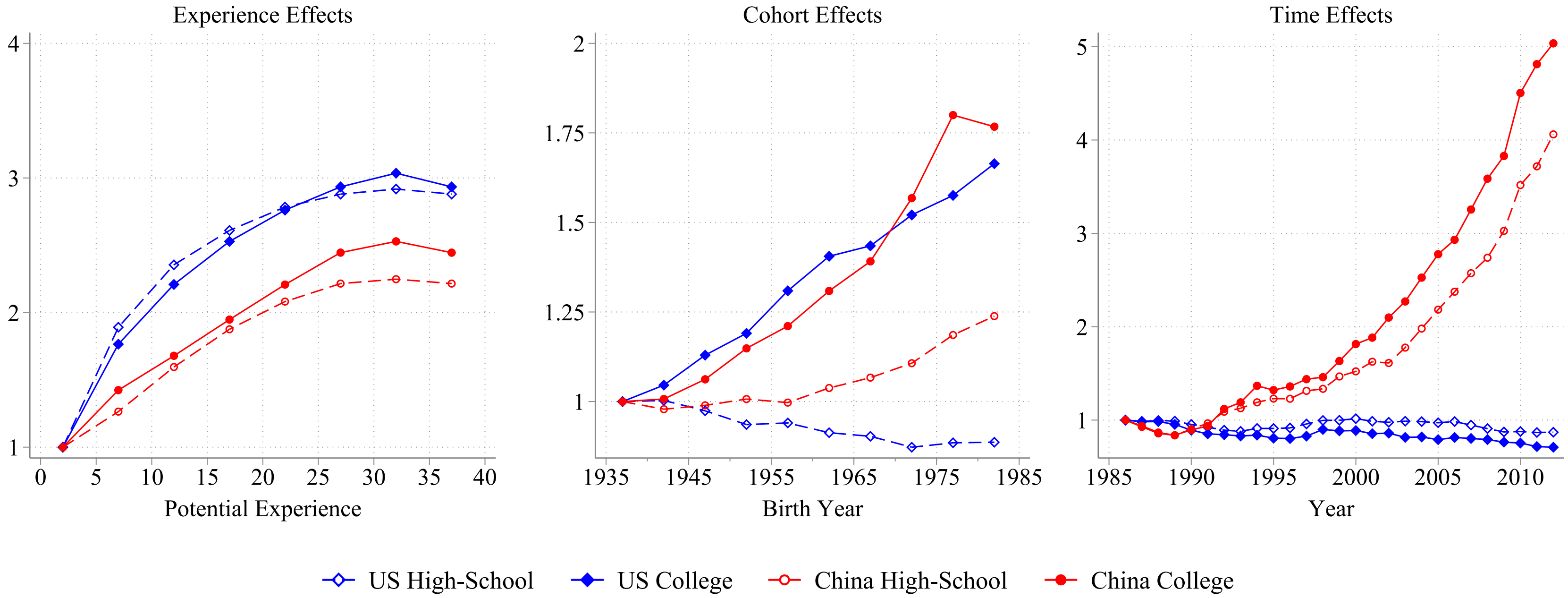
Finally, the time effects are similar for both education groups, especially in their pattern, although the level of changes differ a little bit. In China, the rental price to human capital increases rapidly, and the rental price to college human capital increases even faster than that to high school human capital.

5.1.1 Decomposing College Premium

The college premium literature typically interprets wage ratio of the college graduates and high school graduates as the relative price between high skill and low skill. Under this interpretation, the fact that an increase in the college wage premium is coming together with a remarkable increase in the supply of college workers in the U.S. becomes a puzzle and hence motivates the literature on the skill-biased technical change (see

[18]There were 1.08 million students admitted by colleges in 1998. The number doubled after only two years, with 2.21 million students admitted in 2000.

Figure 7: Decomposition for College and High School Male Workers



Notes: Figure 4 shows the decomposition results of experience, cohort, and time effects in US (blue diamond) and China (red circle).

Acemoglu and Autor, 2011, for an execellent overview). The implicit assumption, in the above argument, is that the relative amount of human capital between education groups is constant. Suppose the average wage of each education group $e \in \{\text{cl}, \text{hs}\}$ at time t is $w_t^e = p_t^e h_t^e$, where p_t^e is the rental price to human capital of type e at time t, and h_t^e is the average human capital for workers of education e. Note that

$$\frac{w_t^{\text{cl}}}{w_t^{\text{hs}}} = \frac{p_t^{\text{cl}}}{p_t^{\text{hs}}} \times \frac{h_t^{\text{cl}}}{h_t^{\text{hs}}} := \frac{p_t^{\text{cl}}}{p_t^{\text{hs}}} \times \xi_t$$

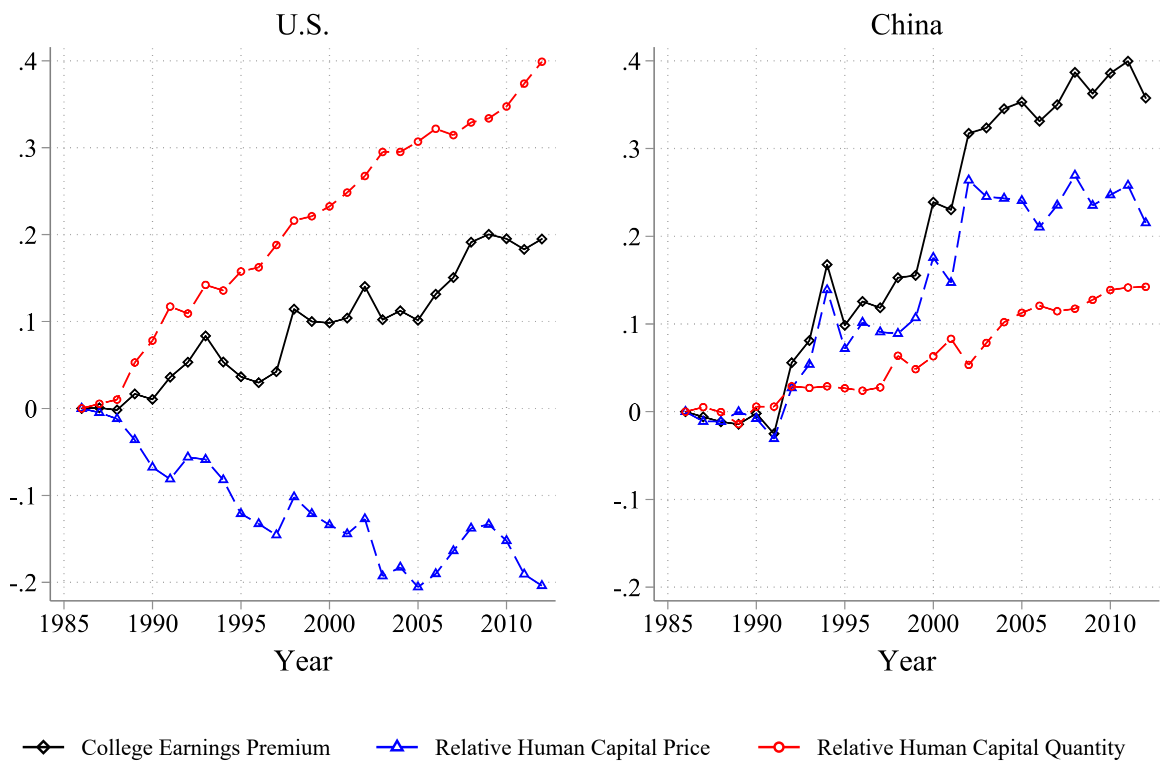
Only under the assumption of constant relative amount of human capital that $\xi_t = h_t^{\text{cl}}/h_t^{\text{hs}} \equiv \xi, \forall t$, do we have that changes in college premium over time completely reflect changes in the relative price of college human capital and high school human capital.

We carefully distinguish between the relative wage of college versus high school workers and the relative price to college versus high school human capital. Using our estimated relative price series, we decompose the evolution of average college premium into the relative changes in the prices of the two types of human capital and the relative changes in the quantity. The results are shown in Figure 8.

In the U.S., the relative price between college human capital and non-college human capital is actually declining. This means the rising college premium in the U.S. is mainly a result of the increase in the relative quantity between college human capital and non-college human capital. In fact, the relative human capital quantity of an average college worker has to increase even more than the college premium in order to offset the declining relative skill price. In China, the relative price increases, but more than the increase in relative human capital quantity.

The finding that most of the rise in the relative wage of college workers versus non-college workers in U.S. is accounted for by the relative human capital, rather than the relative skill price, is consistent with Bowlus and Robinson (2012). At first glance, this may seem a contradiction to the skill-biased technological changes. We thus further push the idea in Bowlus and Robinson (2012) to attempt to infer the extent of skill-biased technological changes in both countries. We find no contradiction between declining relative skill prices and the skill-biased technological changes. In fact, our decomposition results still reveal a large skill-biased technological change. The main idea is that even with skill-biased technological change, if relative quantity of college human capital grow so fast, then the relative price of college human capital could still go down.

Figure 8: Decomposing Changes in College Premium



Notes: This figure decomposes changes in college premium into changes in relative human capital price and changes in relative human capital quantity.

5.1.2 Skill-Biased Technical Change

In this section, we will re-estimate the magnitude of skill-biased technological change, taking into account the potential changes in the relative quantity of college versus high school human capital. Consider an aggregate production function that exhibits constant elasticity of substitution (CES) over college and high school human capital:

$$Y(t) = \left[\left(A_s(t) H_s(t) \right)^{\frac{\sigma-1}{\sigma}} + \left(A_u(t) H_u(t) \right)^{\frac{\sigma-1}{\sigma}} \right]^{\frac{\sigma}{\sigma-1}}, \tag{5}$$

where H_s and H_u are the aggregate human capital quantity of the two types of skills, A_s and A_u are the respective augmenting technology, and $\sigma > 0$ is the elasticity of substitution between these two types of human capital. Assume skills are paid by their marginal product, we have the relative price of the two types of skills being

$$\ln\left(\frac{p_s}{p_u}\right) = \frac{\sigma - 1}{\sigma}\ln\left(\frac{A_s}{A_u}\right) - \frac{1}{\sigma}\ln\left(\frac{h_s}{h_u}\right) - \frac{1}{\sigma}\ln\left(\frac{L_s}{L_u}\right),$$

where h_s and h_u are the efficiency units (human capital quantity) per worker of these two types, and L_s and L_u are the aggregate number of workers of these two types, such that the aggregate supply is $H_u = h_u L_u$ (and $H_s = h_s L_s$). The first term on the right-hand side captures the effect of skill-biased technological changes. The second term reflects changes in the relative quantity of human capital per worker. The last term is the simple labor supply effect.

We calibrate $\sigma = 1.4$, which is the benchmark value estimated by Katz and Murphy (1992), and obtain an estimated series for contributions of changes in A_s/A_u.[19] Since $\sigma > 1$, an increase in A_s/A_u (i.e., skill-biased technological change) will increase p_s/p_u, while an increase in either h_s/h_u or L_s/L_u (i.e., an increasing relative supply of skilled human capital) will decrease p_s/p_u. Our decomposition delivers changes in the relative price p_s/p_u and the relative human capital quantities per worker h_s/h_u. Since the relative labor supply L_s/L_u is observed, the skill-biased technical changes can thus be obtained as a residual.

The contributions of relative labor supply, relative human capital per worker, and skill-biased technological change to the evolution of relative human capital prices is shown in Figure 9. Both countries has witnessed an increasing fraction of college workers and continuing skill-biased technological changes during the same period.

We discuss the relation to previous estimates of skill-biased technological changes in Appendix D.2.

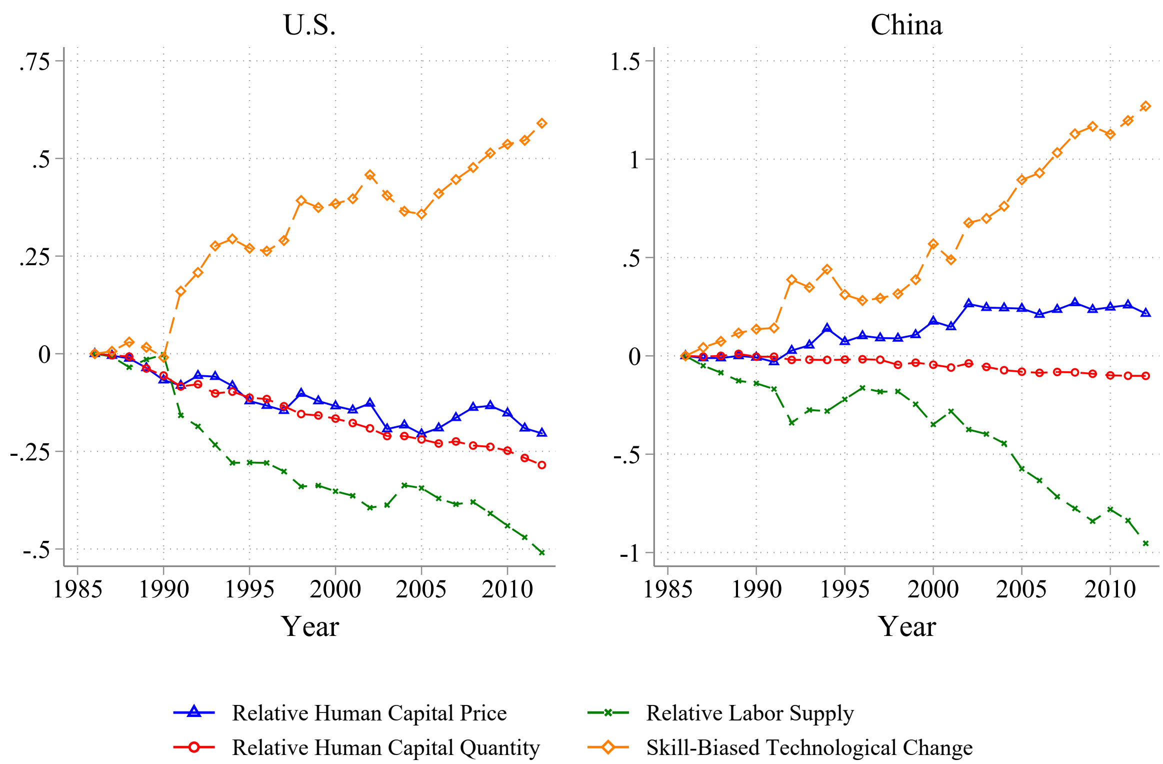
5.2 Heterogeneous Experience Effects Across Cohorts

In the previous decomposition exercise, the life-cycle human capital accumulation path is restricted to be the same across cohorts. Suppose the path of returns to experience varies across cohorts. That is, suppose the true model is

$$w_{i,t} = \text{cons} + p_t + s_c + r_k^c + \epsilon_{i,t},$$

[19] Acemoglu and Autor (2011) conclude that most estimates in the literature agreed on a value of σ to be between 1.4 and 2. We report the decomposition results under $\sigma = 2$ in Figure B.7 in the Appendix. Although the numbers change a bit, the overall pattern does preserve.

Figure 9: Decomposing Changes in Relative Human Capital Prices ($\sigma = 1.4$)



Notes: This figure decomposes changes in relative human capital prices into relative labor supply, relative human capital quantity per worker, and skill-biased technological change, under $\sigma = 1.4$ estimated by Katz and Murphy (1992).

where the experience effect r_k^c is allowed to vary by c, and still, $\mathbb{E}_i\left[\varepsilon_{i,t} | i \in c, t\right] = 0, \forall c, t$ by construction.

A natural algorithm for estimating the model with heterogeneous experience effects is to first estimate the model with homogeneous experience effects to obtain the time effects, and then tease out the time effects from the life-cycle profiles to obtain the cohort-specific experience effects. We will show that such a plug-in estimator turns out to be biased. Formally, this two-step algorithm is to first estimate a model with

$$w_{i,t} = p_t + s_c + r_k + \tilde{\varepsilon}_{i,t},$$

where $r_k := \mathbb{E}_c\left[r_k^c\right]$, and $\tilde{\varepsilon}_{i,t} := \varepsilon_{i,t} + r_k^c - r_k$, and then back out the heterogeneous paths of returns to experience by a simple plug-in estimator

$$\hat{r}_k^c = w_{c,t} - \widehat{\text{cons}} - \hat{p}_t - \hat{s}_c,$$

where $w_{c,t}$ is the average log earning across workers in cohort c at time t directly from data, and $\widehat{\text{cons}}, \hat{p}_t, \hat{s}_c$ are the estimates from the first step with the specification of homogeneous returns to experience. However, we show in Appendix A.1 that this two-step algorithm does not generate consistent estimates, if the returns to experience is heterogeneous by cohorts.

We therefore turn to different estimation algorithm, but based on the same identifying assumption. We follow Bowlus and Robinson (2012) to assume that the flat regions with no additional experience effects are 50-59 years old for college graduates and 46-55 years old for high school graduates. For each education group, we thus have 9 moments every year for the time effects. For instance, the 9 moments that determine the time effect at t for college workers are: (1) wage difference between a typical 50-year-old worker in year t and a typical 51-year-old worker in year $t+1$, (2) wage difference between a typical 51-year-old worker in year t and a typical 52-year-old worker in year $t+1$, ..., (9) wage difference between a typical 58-year-old worker in year t and a typical 59-year-old worker in year $t+1$. We put equal weights on each moment and identify a series of human capital price changes. In the second step, we exclude the human capital price changes from the raw life-cycle earnings profile of each cohort, to obtain the cohort-specific experience effects. Formally, using the correct estimates for price changes $\hat{p}_t$, we can back out the heterogeneous paths of life-cycle human capital accumulation for each cohort up to a normalization, by $\hat{h}_{c,t} = w_{c,t} - \hat{p}_t$.

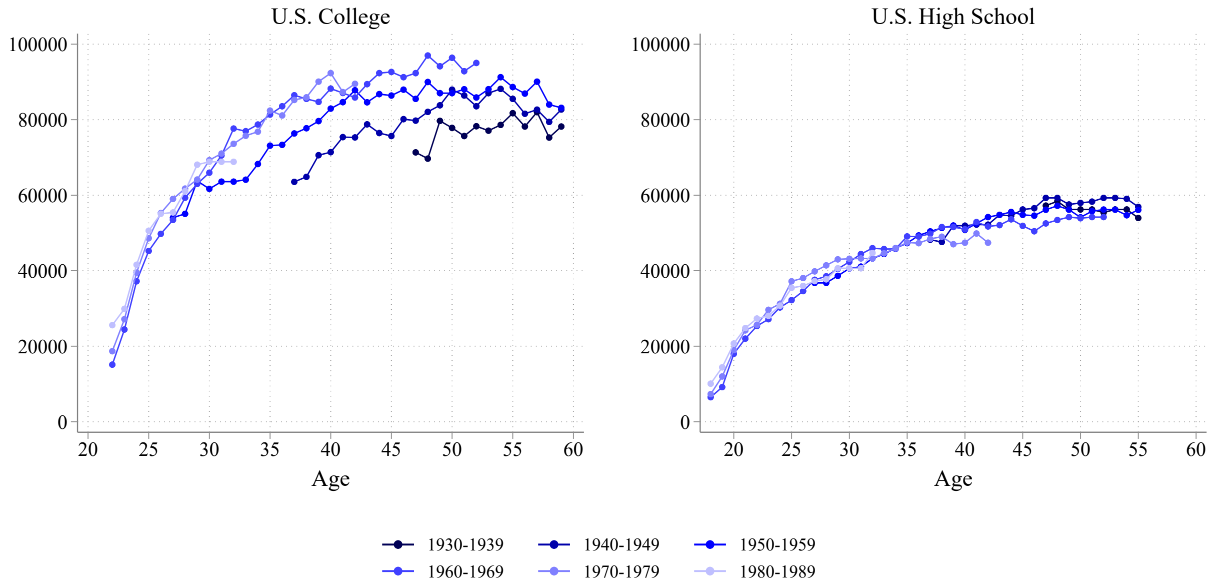
Figure 10 reports the life-cycle human capital accumulation by cohorts estimated, where darker lines are for older cohorts and lighter lines for younger cohorts. The life-cycle human capital accumulation paths for different cohorts highly overlap with each other, suggesting that the homogeneous returns to experience assumption turns out to be a reasonable one. If anything, the profiles are shifting upwards for college workers in both US and China, suggesting a large cohort effects for these two groups. This is consistent with our estimates in Figure 7. Although the levels shift upwards for these two groups, the shifts are mosly parallel, meaning that the life-cycle shapes for different cohorts are very close to each other. This reassures that our previous LMPQS specification assuming experience effects are homogeneous across cohorts does not suffer from the misspecification bias.

5.3 Looking Forward

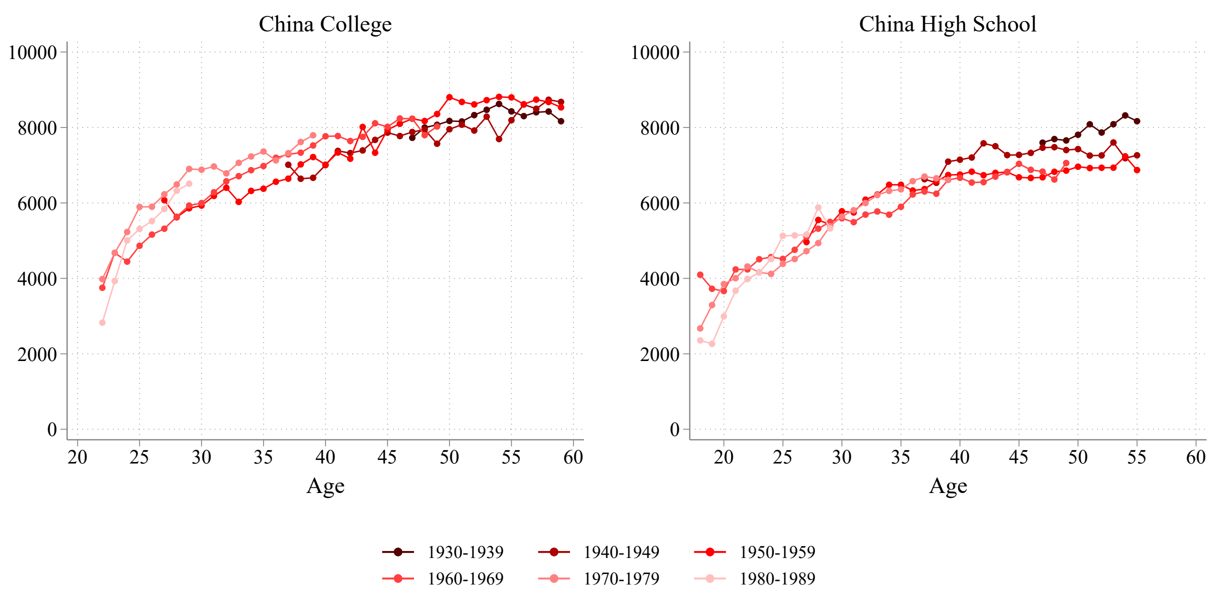
The fast growth in China is expected to slow down in the future. Between 1986 and 2012, the average inter-cohort human capital growth rate in China is 1.40% ($= 1.87^{1/45} - 1$)

Figure 10: Cohort-Specific Life-Cycle Paths of Human Capital Accumulation

(a) U.S.

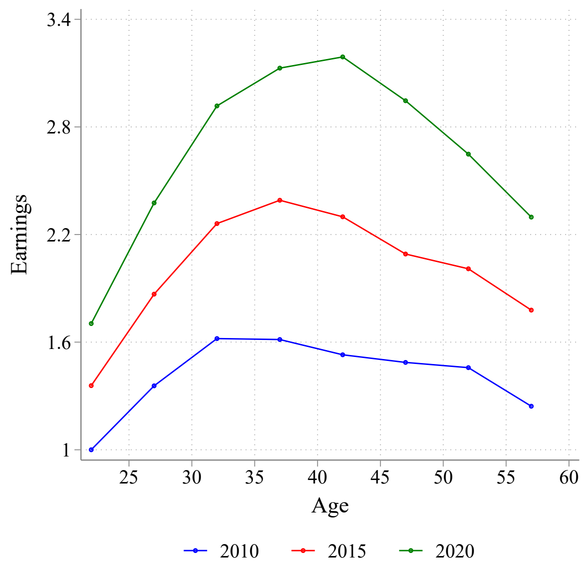


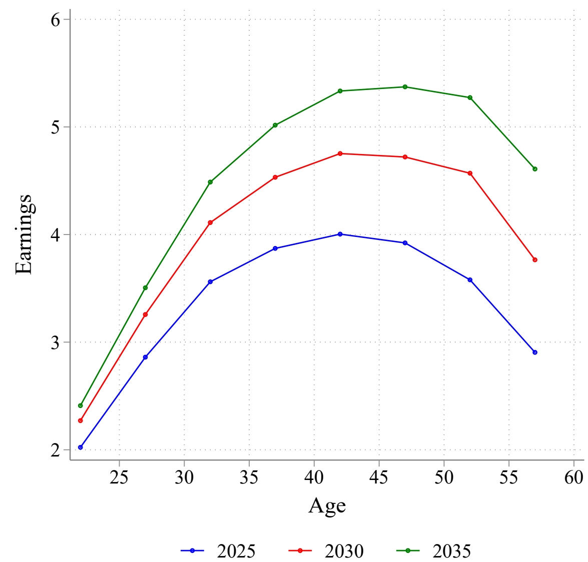
(b) China



Notes: Figure 10 reports cohort-specific life-cycle human capital accumulation, where darker lines are for older cohorts and lighter lines for younger cohorts.

Figure 11: A Hypothetical Scenario for China's Earnings Profiles in 30 years

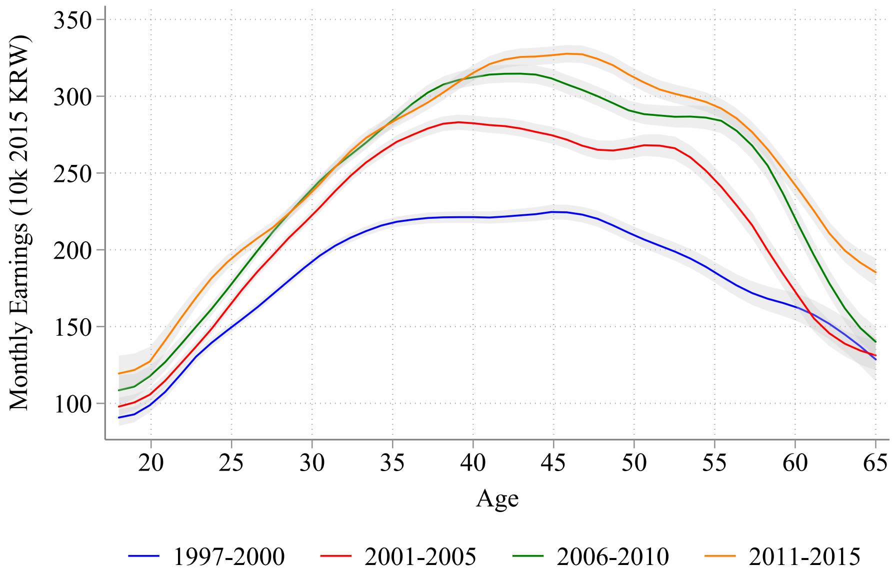




per year, and the average growth rate of human capital prices in China is 4.80% ($= 3.38^{1/26} - 1$) per year. Both are astonishing growth, while the two growth rates are 0.39% and -1.36% per year for US. This section performs an experiment that both the cohort effects and time effects still grow but start to uniformly decelerate until a stationary environment of zero growth in cohort and time effect (approximately the U.S. case) in 30 years, with the experience effects fixed at China's current estimated level.

In this scenario, the vertical gaps between two cross-sectional earnings profiles are shrinking, showing the slowdown in the time effects. Notably, the "golden age" is around 30-35 in 2010, but the "golden age" would be becoming older and to 45-50 years old in 2035. Recall Proposition 1 and its corollary that the position of the "golden age" is essentially a race between experience effects and cohort effects. The "golden age" becoming older and older is a result of the slowdown in cohort effects. This scenario for the future 30 years reveals an opposite pattern to what happened in China during the past 30 years, but a similar pattern to what happened in Korea during the past 10 years, using data from Korean Labor and Income Panel Study (KLIPS). Korea experienced its fastest growth during 1960s to 1990s. After that, it began to slowdown. Figure **??** depicts the decomposition for Korea, together with the previous decomposition for U.S. and China. It is worth noting that the cohort effects are particularly large from cohort 1945 to cohort 1960, but starts to decelerate afterwards. This is consistent with our explanation of the race between inter-cohort productivity growth and returns to experience.

Figure 12: Cross-Sectional Age-Earnings Profiles of Korean Male Workers



As inter-cohort productivity growth starts to give its way to experience in Korea, the "golden age" comes back to older ages, as in our hypothetical scenario in Figure 11.

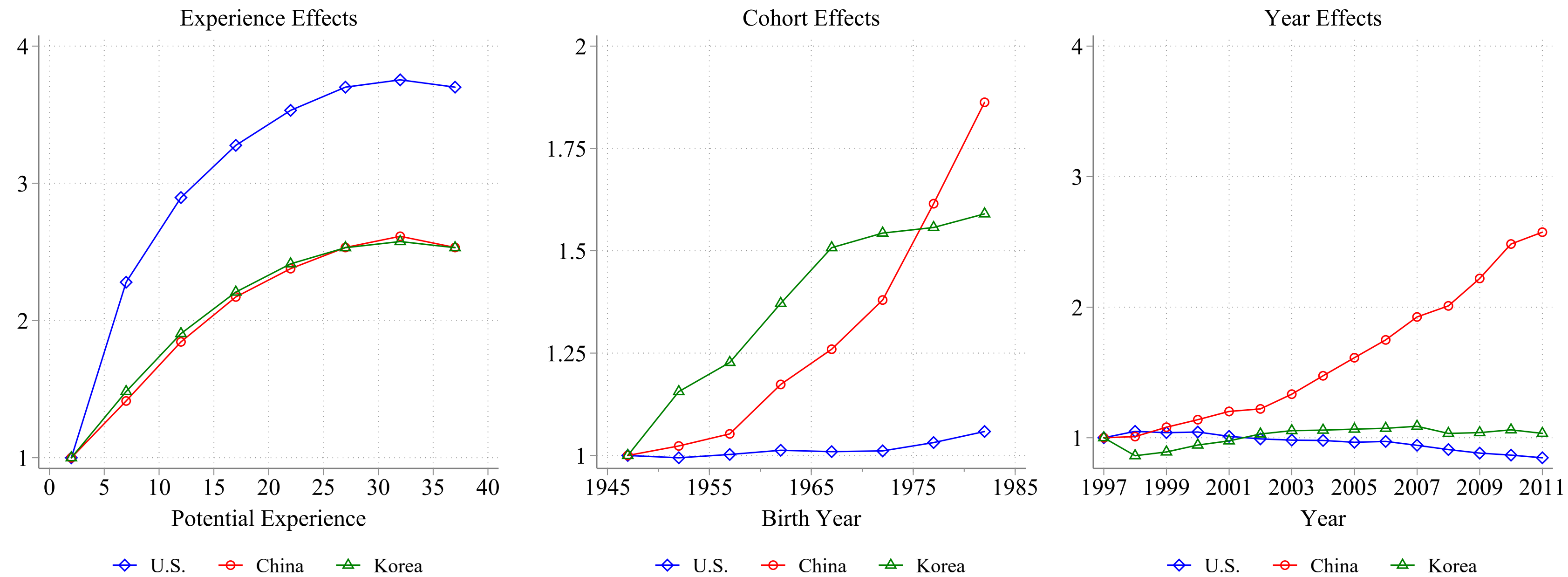
6 Conclusion

This paper documents three stark differences in the age-earnings profiles between the U.S. and China. First, the cross-sectional "golden age" stays stable at around 45-50 years old in the U.S. but continuously decreases from 50 to 35 years old in China during the past 30 years. Second, age-specific labor earnings grow dramatically in China but do not witness much growth in the U.S. Third, the cross-sectional and life-cycle age-earnings profiles look remarkably similar in the U.S. but differ substantially in China.

As the first step to explain these puzzles, we exploit a decomposition framework to address these facts. The decomposition suggests that China has experienced a much larger inter-cohort improvement in human capital and increase in the rental price to human capital, compared to the U.S. But the return to experience is higher in the U.S. We apply the decomposition result to adjust for the changes in human capital and obtain a series of estimated TFP growth that is lower than previous estimates.

In the baseline analysis, we assume workers are perfect substitutes. We discuss the possibility that different education groups are not perfect substitutes, but different cohorts are still restricted to be perfect substitutes. There might be some technical changes

Figure 13: Decomposition



Notes: This Figure shows the decomposition results of experience, cohort, and time effects in US (blue diamond), China (red circle) and Korea (green triangle), under the baseline specification.

that makes older cohorts obsolete, and therefore, cohorts are not perfect substitutes in the production function. In that case, cohort sizes will matter. But it is hard to distinguish technical changes that favors younger generations and inter-cohort human capital growth without better-suited data or richer structures. We leave this question for future research.

We believe another fruitful direction is to link the decomposition results to specific institutions and evaluate how much does each policy contribute to the three effects. We list some of our hypotheses here. The college expansion is the most related to the inter-cohort improvement in human capital. The accession to WTO and SOE reforms may well increase the overall efficiency of the economy and hence also the rental price to human capital. Throughout the paper, we focused on urban males. A natural question is how would structural change from agriculture to industry, increasing female labor participation, and internal migration (especially in China) be reflected in the decomposition.

References

Acemoglu, Daron. 2002. "Technical Change, Inequality, and the Labor Market." *Journal of Economic Literature*, 40(1): 7–72.

Acemoglu, Daron, and David Autor. 2011. "Skills, tasks and technologies: Implications for employment and earnings." In *Handbook of labor economics*. Vol. 4, 1043–1171. Elsevier.

Aguiar, Mark, and Erik Hurst. 2013. "Deconstructing Life Cycle Expenditure." *Journal of Political Economy*, 121(3): 437–492.

Autor, David H, Lawrence F Katz, and Alan B Krueger. 1998. "Computing inequality: have computers changed the labor market?" *The Quarterly journal of economics*, 113(4): 1169–1213.

Ben-Porath, Yoram. 1967. "The Production of Human Capital and the Life Cycle of Earnings." *Journal of Political Economy*, 75(4, Part 1): 352–365.

Bils, Mark, and Peter J Klenow. 2000. "Does Schooling Cause Growth?" *American Economic Review*, 90(5): 1160–1183.

Bowlus, Audra J, and Chris Robinson. 2012. "Human Capital Prices, Productivity, and Growth." *American Economic Review*, 102(7): 3483–3515.

Brandt, Loren, Johannes Van Biesebroeck, Luhang Wang, and Yifan Zhang. 2017. "WTO Accession and Performance of Chinese Manufacturing Firms." *American Economic Review*, 107(9): 2784–2820.

Browning, Martin, Lars Peter Hansen, and James J Heckman. 1999. "Micro Data and General Equilibrium Models." *Handbook of Macroeconomics*, 1: 543–633.

Burdett, Kenneth, and Dale T Mortensen. 1998. "Wage Differentials, Employer Size, and Unemployment." *International Economic Review*, 257–273.

Cai, Yong, Feng Wang, Ding Li, Xiwei Wu, and Ke Shen. 2014. "China's Age of Abundance: When Might It Run Out?" *The Journal of the Economics of Ageing*, 4: 90–97.

Casanova, Maria. 2013. "Revisiting the Hump-Shaped Wage Profile." *Unpublished mimeo, UCLA*.

Chen, Yuyu, Mitsuru Igami, Masayuki Sawada, and Mo Xiao. 2019. "Privatization and Productivity in China." *Available at SSRN 2695933*.

Deaton, Angus. 1997. *The Analysis of Household Surveys: A Microeconometric Approach to Development Policy.* The World Bank.

Deaton, Angus, and Christina Paxson. 1994. "Intertemporal Choice and Inequality." *Journal of Political Economy*, 102(3): 437–467.

Feenstra, Robert C., Robert Inklaar, and Marcel P. Timmer. 2015. "The Next Generation of the Penn World Table." *American Economic Review*, 105(10): 3150–82.

Fernald, John. 2014. "A Quarterly, Utilization-Adjusted Series on Total Factor Productivity." Federal Reserve Bank of San Francisco.

Fernald, John G. 2015. "Productivity and Potential Output before, during, and after the Great Recession." *NBER Macroeconomics Annual*, 29(1): 1–51.

Flood, Sarah, Miriam King, Renae Rodgers, Steven Ruggles, and J. Robert Warren. 2018. "Integrated Public Use Microdata Series, Current Population Survey: Version 6.0 [dataset]." *Minneapolis, MN: IPUMS*.

Ge, Suqin, and Dennis Tao Yang. 2014. "Changes in China's Wage Structure." *Journal of the European Economic Association*, 12(2): 300–336.

Gourinchas, Pierre-Olivier, and Jonathan A Parker. 2002. "Consumption over the Life Cycle." *Econometrica*, 70(1): 47–89.

Hall, Robert E. 1968. "Technical Change and Capital from the Point of View of the Dual." *The Review of Economic Studies*, 35(1): 35–46.

Hall, Robert E, and Charles I Jones. 1999. "Why Do Some Countries Produce So Much More Output Per Worker Than Others?" *Quarterly Journal of Economics*, 83–116.

Heckman, James J, Lance Lochner, and Christopher 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): 1–58.

Huggett, Mark, Gustavo Ventura, and Amir Yaron. 2011. "Sources of Lifetime Inequality." *American Economic Review*, 101(7): 2923–54.

Islam, Asif, Remi Jedwab, Paul Romer, and Daniel Pereira. 2018. "The Sectoral and Spatial Allocation of Labor and Aggregate Returns to Experience."

Jovanovic, Boyan. 1979. "Job Matching and the Theory of Turnover." *Journal of political economy*, 87(5, Part 1): 972–990.

Kambourov, Gueorgui, and Iourii Manovskii. 2009. "Accounting for the Changing Life-Cycle Profile of Earnings."

Karabarbounis, Loukas, and Brent Neiman. 2013. "The Global Decline of the Labor Share." *The Quarterly Journal of Economics*, 129(1): 61–103.

Katz, L. F., and K. M. Murphy. 1992. "Changes in Relative Wages, 1963-1987: Supply and Demand Factors." *The Quarterly Journal of Economics*, 107(1): 35–78.

Kuruscu, Burhanettin. 2006. "Training and Lifetime Income." *American Economic Review*, 96(3): 832–846.

Lagakos, David, Benjamin Moll, Tommaso Porzio, Nancy Qian, and Todd Schoellman. 2018. "Life Cycle Wage Growth across Countries." *Journal of Political Economy*, 126(2): 797–849.

Manuelli, Rodolfo E., and Ananth Seshadri. 2014. "Human Capital and the Wealth of Nations." *American Economic Review*, 104(9): 2736–2762.

McKenzie, David J. 2006. "Disentangling Age, Cohort and Time Effects in the Additive Model." *Oxford Bulletin of Economics and Statistics*, 68(4): 473–495.

Mincer, Jacob A. 1974. "Schooling, Experience, and Earnings."

Rubinstein, Yona, and Yoram Weiss. 2006. "Post Schooling Wage Growth: Investment, Search and Learning." *Handbook of the Economics of Education*, 1: 1–67.

Rupert, Peter, and Giulio Zanella. 2015. "Revisiting Wage, Earnings, and Hours Profiles." *Journal of Monetary Economics*, 72: 114–130.

Schulhofer-Wohl, Sam. 2018. "The Age-Time-Cohort Problem and the Identification of Structural Parameters in Life-Cycle Models." *Quantitative Economics*, 9(2): 643–658.

Violante, Giovanni L. 2008. "Skill-Biased Technical Change." *The New Palgrave Dictionary of Economics*, 2.

Zhu, Xiaodong. 2012. "Understanding China's Growth: Past, Present, and Future." *Journal of Economic Perspectives*, 26(4): 103–24.

A Proofs

A.1 Lemma 1

We have the following lemma.

Lemma 1. $\mathbb{E}_i\left[\tilde{\varepsilon}_{i,c,t}\cdot\mathbb{I}_{\{t=t'\}}\right]\neq 0,\forall t'$; $\mathbb{E}_i\left[\tilde{\varepsilon}_{i,c,t}\cdot\mathbb{I}_{\{c=c'\}}\right]\neq 0,\forall c'$; $\mathbb{E}_i\left[\tilde{\varepsilon}_{i,c,t}\cdot\mathbb{I}_{\{t-c=k'\}}\right]=0,\forall k'$.

Lemma 1 tells us that the aforementioned naive plug-in estimator is not consistent, because $\mathbb{E}_i\left[\tilde{\varepsilon}_{i,c,t}\cdot\mathbb{I}_{\{c=c'\}}\right]\neq 0,\forall c'$. In general, at the presence of even one endogenous variable in a regression, all of the other coefficients will be inconsistent. But a special case is that when the endogenous variables are uncorrelated with other right-hand side variables, then the endogeneity will not contaminate the other coefficients. We formalize this result in Lemma 2.

Proof. (1) Fix any t'. Using the law of iterated expectation, we have

$$\mathbb{E}_i\left[\tilde{\varepsilon}_{i,c,t}\cdot\mathbb{I}_{\{t=t'\}}\right]=\mathbb{E}_{c,t}\left[\mathbb{E}_i\left[\tilde{\varepsilon}_{i,c,t}\cdot\mathbb{I}_{\{t=t'\}}|c,t\right]\right].$$

Observe that

$$\mathbb{E}_i\left[\tilde{\varepsilon}_{i,c,t}\cdot\mathbb{I}_{\{t=t'\}}|c,t\neq t'\right]=0,$$

and

$$\mathbb{E}_i\left[\tilde{\varepsilon}_{i,c,t}\cdot\mathbb{I}_{\{t=t'\}}|c,t=t'\right]=\mathbb{E}_i\left[\ln r^c_{t-c}-\ln r_{t-c}+\varepsilon_{i,c,t}|c,t=t'\right]=\ln r^c_{t'-c}-\ln r_{t'-c}.$$

Note that $\mathbb{E}_c\left\{\mathbb{E}_i\left[\tilde{\varepsilon}_{i,c,t}\cdot\mathbb{I}_{\{t=t'\}}|c,t=t'\right]\right\}=\mathbb{E}_c\left[\ln r^c_{t'-c}-\ln r_{t'-c}\right]\neq 0$ in general.

(2) Fix any c'. Using the law of iterated expectation, we have

$$\mathbb{E}_i\left[\tilde{\varepsilon}_{i,c,t}\cdot\mathbb{I}_{\{c=c'\}}\right]=\mathbb{E}_{c,t}\left[\mathbb{E}_i\left[\tilde{\varepsilon}_{i,c,t}\cdot\mathbb{I}_{\{c=c'\}}|c,t\right]\right].$$

Observe that

$$\mathbb{E}_i\left[\tilde{\varepsilon}_{i,c,t}\cdot\mathbb{I}_{\{c=c'\}}|c\neq c',t\right]=0,$$

and

$$\mathbb{E}_i\left[\tilde{\varepsilon}_{i,c,t}\cdot\mathbb{I}_{\{c=c'\}}|c=c',t\right]=\mathbb{E}_i\left[\ln r^c_{t-c}-\ln r_{t-c}+\varepsilon_{i,c,t}|c=c',k\right]=\ln r^{c'}_{t-c'}-\ln r_{t-c'}.$$

Note that $\mathbb{E}_t\left\{\mathbb{E}_i\left[\tilde{\varepsilon}_{i,c,t}\cdot\mathbb{I}_{\{c=c'\}}|c=c',t\right]\right\}=\mathbb{E}_t\left[\ln r^{c'}_{t-c'}-\ln r_{t-c'}\right]$ does not equal 0 in general. Therefore, $\mathbb{E}_i\left[\tilde{\varepsilon}_{i,c,t}\cdot\mathbb{I}_{\{c=c'\}}\right]\neq 0, \forall c'$ in general.

(3) Fix any k'. Using the law of iterated expectation, we have

$$\mathbb{E}_i\left[\tilde{\varepsilon}_{i,c,t}\cdot\mathbb{I}_{\{t-c=k'\}}\right]=\mathbb{E}_{c,t}\left[\mathbb{E}_i\left[\tilde{\varepsilon}_{i,c,t}\cdot\mathbb{I}_{\{t-c=k'\}}|c,t\right]\right].$$

Observe that

$$\mathbb{E}_i\left[\tilde{\varepsilon}_{i,c,t}\cdot\mathbb{I}_{\{t-c=k'\}}|t-c\neq k'\right]=0,$$

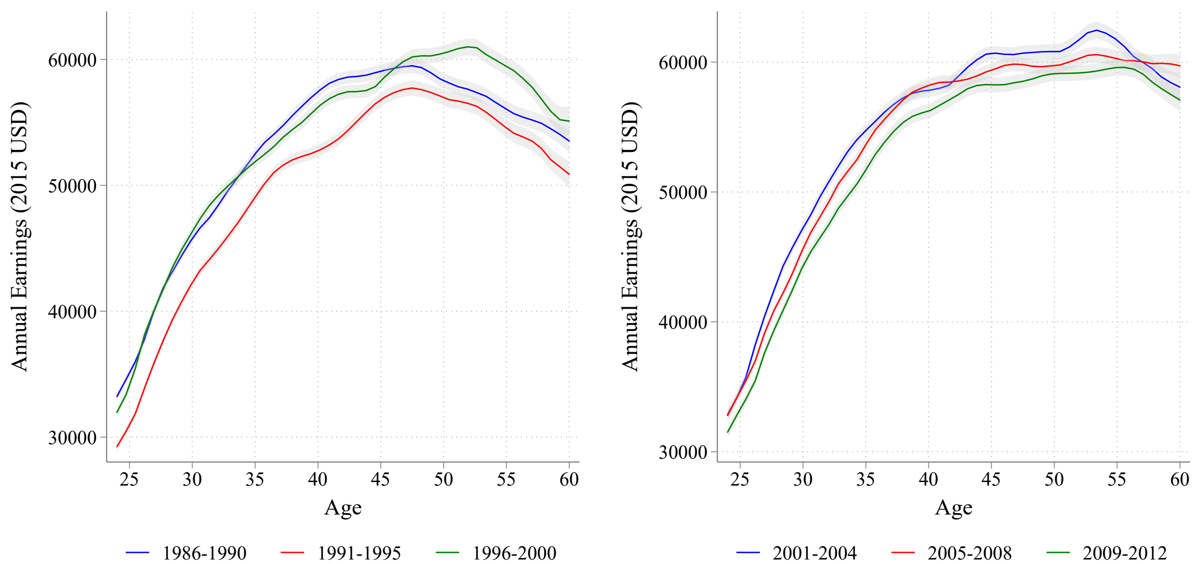
and

$$\mathbb{E}_i\left[\tilde{\varepsilon}_{i,c,t}\cdot\mathbb{I}_{\{t-c=k'\}}|t-c=k'\right]=\mathbb{E}_i\left[\ln r^c_{t-c}-\ln r_{t-c}+\varepsilon_{i,c,t}|t-c=k'\right]=\ln r^c_{k'}-\ln r_{k'}.$$

Note that $\mathbb{E}_c\left\{\mathbb{E}_i\left[\tilde{\varepsilon}_{i,c,t}\cdot\mathbb{I}_{\{t-c=k'\}}|t-c=k'\right]\right\}=\mathbb{E}_t\left[\ln r^c_{k'}-\ln r_{k'}\right]=0$. Thus $\mathbb{E}_i\left[\tilde{\varepsilon}_{i,c,t}\cdot\mathbb{I}_{\{t-c=k'\}}\right]=0, \forall k'$. □

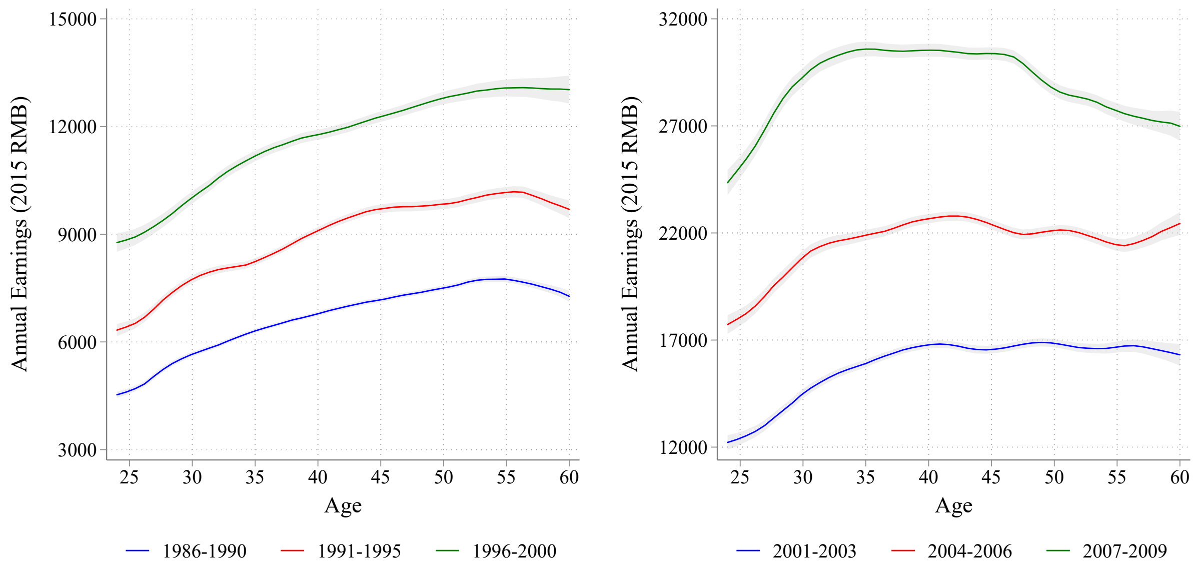
B Additional Figures

Figure B.1: Cross-Sectional Age-Earnings Profiles of U.S. Male Workers that Live in Metropolitan Areas



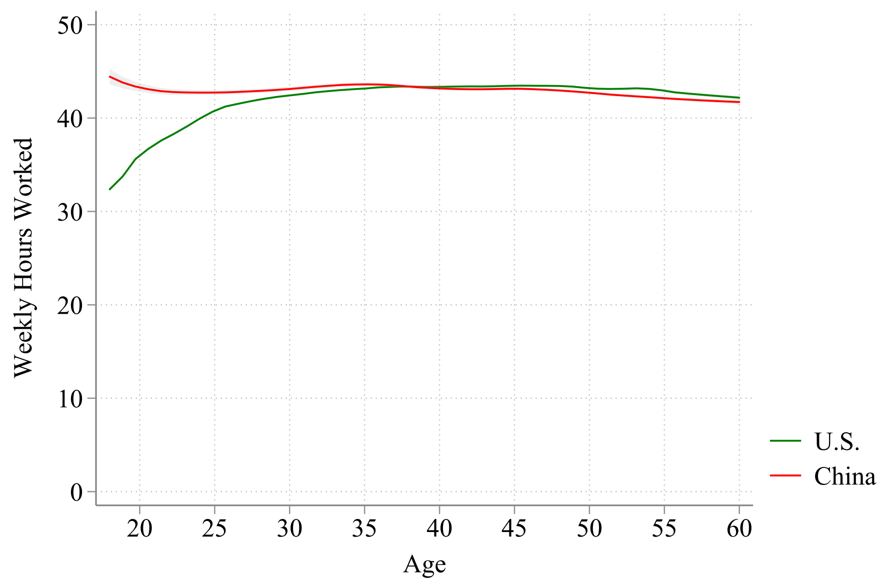
Notes: Figure B.1 plots the cross-sectional age-earnings profiles of U.S. male workers that live metropolitan areas, using March CPS from 1986 to 2012. Each curve represents a cross section that pools adjacent years. The solid lines are kernel smoothed values and the gray shaded areas are the 95% confidence intervals.

Figure B.2: Cross-Sectional Age-Earnings Profiles of Chinese Urban Male Workers in 15 Provinces Covering 1986-1991 and 2002-2009



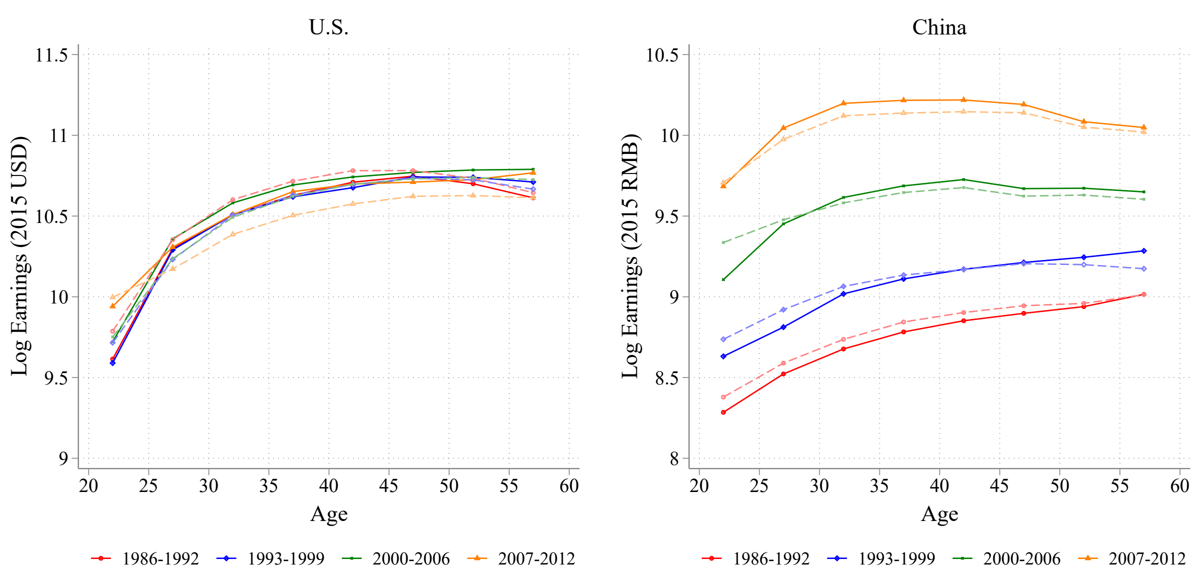
Notes: Figure B.2 plots the cross-sectional age-earnings profiles of Chinese Urban male workers in Beijing, Shanxi, Liaoning, Heilongjiang, Shanghai, Jiangsu, Anhui, Jiangxi, Shandong, Henan, Hubei, Guangdong, Sichuan, Yunnan, Gansu, covering 1986-1991 in the left panel and 2002-2009 in the right panel. Each curve represents a cross section that pools adjacent years. The solid lines are kernel smoothed values and the gray shaded areas are the 95% confidence intervals. Note that the vertical scale of the left and right panels differ.

Figure B.3: Cross-Sectional Age-Hours Profiles



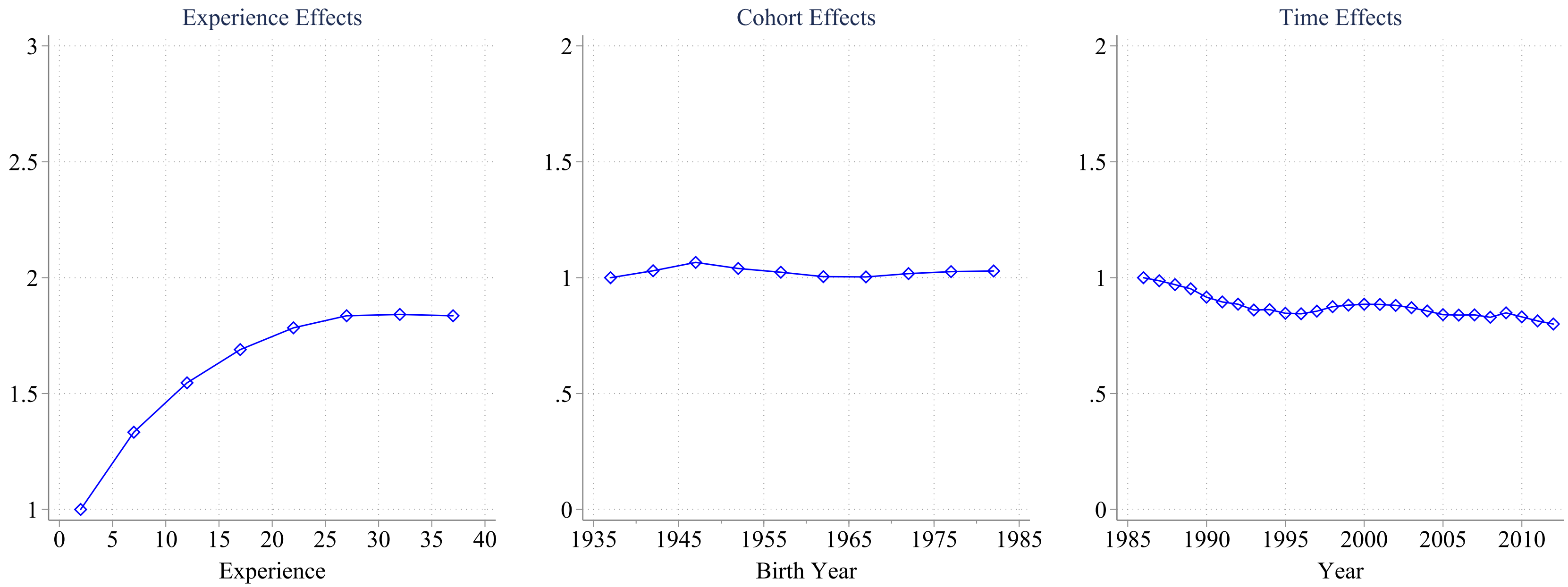
Notes: Figure B.3 plots the cross-sectional age-hours profiles of U.S. and Chinese male workers in 2002-2006. Hours worked per week is measured by the "total number of hours usually worked per week over all jobs the year prior to the survey" from CPS (for U.S.) and imputed as "total number of hours worked last month" divided by 4.286 from UHS (for China).

Figure B.4: Goodness of Fit



Notes: This figure shows the model fit. Solid lines are data and dashed lines are model predictions.

Figure B.5: Decomposition Using Hourly Wage for Full-Time Workers



Notes: This figure shows the decomposition results of experience, cohort, and time effects in US based on hourly wage for full-time workers.

Figure B.6: College Share by Birth Cohort

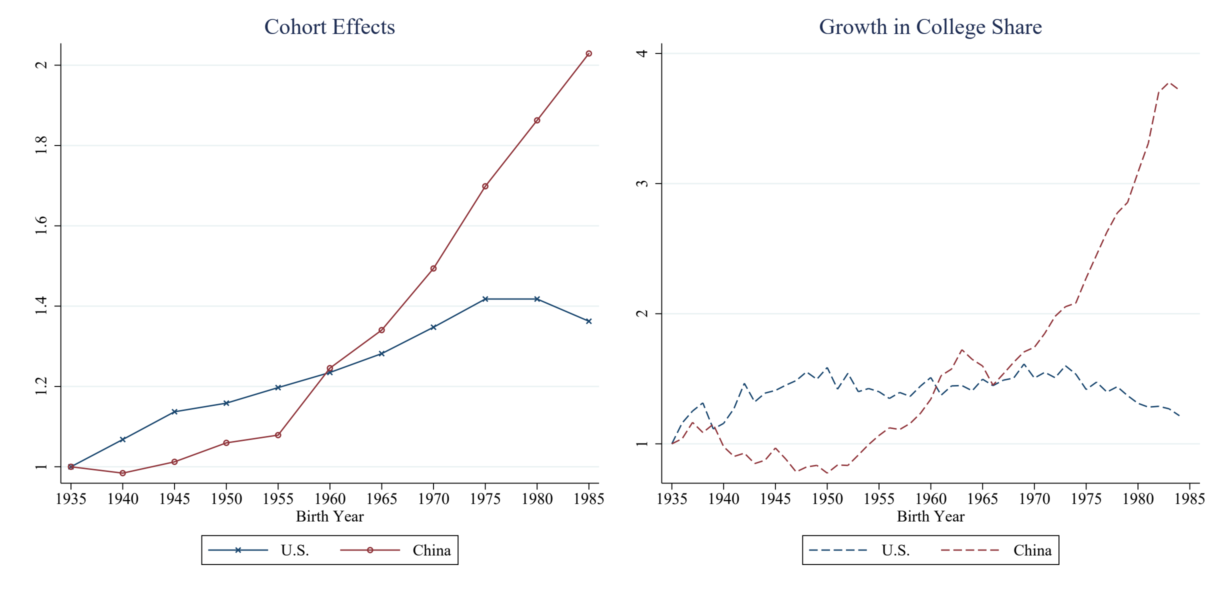
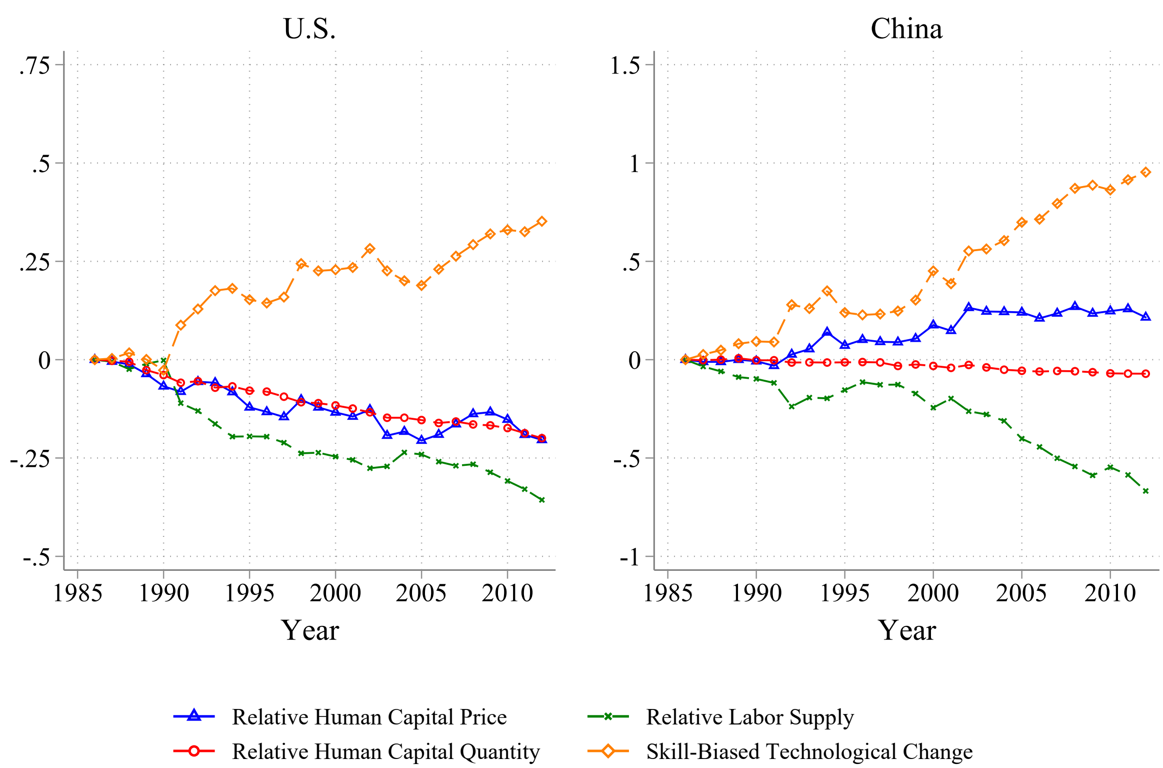


Figure B.7: Decomposing Changes in Relative Human Capital Prices ($\sigma = 2$)



Notes: This figure decomposes changes in relative human capital prices into relative labor supply, relative human capital quantity per worker, and skill-biased technological change, under $\sigma = 2$, the upper bound for σ estimated in the literature, according to Acemoglu and Autor (2011).

C Additional Tables

Table C.1: Sample Provinces in Our UHS Random Subsample

Provinces	Code	1986-2001	2002-2009	2010-2012
Beijing	11	X	X	
Shanxi	14	X	X	
Liaoning	21	X	X	X
Heilongjiang	23	X	X	
Shanghai	31	X	X	X
Jiangsu	32	X	X	
Zhejiang	33	X		
Anhui	34	X	X	
Jiangxi	36	X	X	
Shandong	37	X	X	
Henan	41	X	X	
Hubei	42	X	X	
Guangdong	44	X	X	X
Chongqing	50		X	
Sichuan	51	X	X	X
Yunnan	53	X	X	
Shaanxi	61	X		
Gansu	62	X	X	
Total		17	16	4

Notes: This table reports the regional coverage of our UHS random subsample.

D Discussion on Literature

D.1 Age-Cohort-Time Identification

D.1.1 McKenzie (2006)

Hall (1968) and McKenzie (2006) show that second differences (and higher order differences) of these effects can be identified without any assumption, while first differences of these effects can be identified with one restriction. To see this, suppose the variable of interest is a linearly additive model of cohort, time and age effects with

$$y_{c,t} = \alpha_c + \beta_k + \gamma_t + \varepsilon_{c,t},$$

where $k := t - c$. Consider cohort c_1 at time periods t_1 and $t_2 = t_1 + 1$ and take a first difference:

$$\Delta_t y_{c_1,t_2} \equiv (y_{c_1,t_2} - y_{c_1,t_1}) = \left(\beta_{k_2} - \beta_{k_1}\right) + (\gamma_{t_2} - \gamma_{t_1}) + \Delta_t \varepsilon_{c_1,t_2},$$

where $k_1 = t_1 - c_1$ and $k_2 = t_2 - c_1 = k_1 + 1$. Similarly, consider cohort $c_0 = c_1 - 1$ at the same time periods t_1 and t_2:

$$\Delta_t y_{c_0,t_2} \equiv (y_{c_0,t_2} - y_{c_0,t_1}) = \left(\beta_{k_3} - \beta_{k_2}\right) + (\gamma_{t_2} - \gamma_{t_1}) + \Delta_t \varepsilon_{c_0,t_2},$$

where $k_3 = t_2 - c_0$. Taking a second difference of the above two first differences we have

$$\Delta_c \Delta_t y_{c_0,t_2} \equiv (\Delta_t y_{c_0,t_2} - \Delta_t y_{c_1,t_2}) = \left(\beta_{k_3} - \beta_{k_2}\right) - \left(\beta_{k_2} - \beta_{k_1}\right) + \Delta_c \Delta_t \varepsilon_{c_0,t_2}.$$

Thus the change in the slope of the age-effect profile is identified. Second differences of time and cohort effects are also identified in the same fashion.

Furthermore, by normalizing one first difference, one can recover all remaining slopes. To illustrate, say, we normalize one first difference of experience effects. Then, we can obtain all other first differences of experience effects from the identified second differences. With first differences of experience effects at hand, we can identify first differences of time effects, using the fact that the time differences of the outcome variable for a given cohort are the sum of first differences of experience effects and first differences of time effects. Similarly, we can identify first differences of cohorts, too. Hence one normalization on a first difference suffices for identification of all first differences.

In addition, normalizing one level each of two effects, one can recover all levels. But throughout this paper, what we care most are the slopes, which is the relative effects up to a benchmark group, not the levels. Hence we load the level of the benchmark group to a constant term, and aim at identifying first differences (i.e., slopes).

D.1.2 Deaton (1997)

As argued in the previous section, one normalization suffices for identification. Many papers thus proceed in this way and adopt one normalization. The consumption literature, though studies a different topic, offers one popular approach to deal with the collinearity issue. Deaton and Paxson (1994) and later Deaton (1997), in the section "Decompositions by age, cohort, and year" (page 123) of his book "*The Analysis of Household Surveys*," view year dummies as a device to capture cyclical fluctuation, with the restriction that time effects are orthogonal to a linear time trend. Aguiar and Hurst (2013) is a recent example that follows the same practice to study life cycle expenditures.

Suppose again

$$y_{i,c,t} = \text{cons} + \alpha_c + \beta_k + \gamma_t + \varepsilon_{i,c,t}.$$

In matrix form, we have

$$y = C + A\alpha + B\beta + \Gamma\gamma + \varepsilon,$$

where each row is an observation, A, B, Γ are matrices of cohort dummies, experience dummies, and time dummies, respectively, and α, β, γ are vectors of cohort effects, experience effects, and time effects, respectively. Note that the collinearity across time, cohort, and age $t = c + k$ implies

$$\Gamma s_t = A s_c + B s_k,$$

where the s vectors are arithmetic sequences $\{0, 1, 2, 3, \dots\}$ of the length given by the number of columns of the matrix that premultiplies them. Replace the parameter vectors by

$$\tilde{\alpha} = \alpha + \kappa s_c, \quad \tilde{\beta} = \beta + \kappa s_k, \quad \tilde{\gamma} = \gamma - \kappa s_t.$$

Thus an arbitrary time-trend can be added to the age dummies and cohort dummies by subtracting it from the year dummies, which sheds light on the non-identification problem. Deaton assumes that the year effects capture cyclical fluctuations or business-cycle effects. Formally, in addition to $\sum_t \gamma_t = 0$ (which is an innocuous normalization as it only adjusts the constant term), he restricts that $s'\gamma = 0$ to capture the idea that time effects are orthogonal to a linear trend. Notice that the label of years is without loss of generality, for any chronological relabel of years will still satisfy this relation.

To implement Deaton's idea, one can regress y on a set of dummies for each cohort excluding (say) the first, a set of dummies for each age excluding (say) the first, and a set of $T-2$ year dummies defined as follows for $t = 3, .., T$,

$$d_t^* = d_t - [(t-1)d_2 - (t-2)d_1].$$

The coefficients of d_t^*'s thus give the third through final year coefficients. Then one can recover the first and second coefficients γ_1, γ_2 by solving the system of equations $\sum_t \gamma_t = 0$ and $s'\gamma = 0$.

This approach assumes that secular trends appear only in cohort effects and time effects simply reflect fluctuations. Alternatively, one could also take an opposite restriction that cohort dummies are orthogonal to the time trend. Some papers may investigate both, or a mixture of the two (e.g., LMPQS), to examine the sensitivity of their results to the identifying assumption.

Another related but different approach is even simpler – instead of imposing a normalization, it directly uses observable measures as proxies for time effects. For instance, in Gourinchas and Parker (2002) studying age-consumption profiles and Kambourov and Manovskii (2009) studying age-earnings profiles, they use unemployment rates to capture the time effects arising from booms and recessions.

D.1.3 Schulhofer-Wohl (2018)

Schulhofer-Wohl (2018) proposes an alternative method that does not require the somewhat arbitrary normalization, but virtually shifts focus from directly estimating the age effects to estimating the parameters in age effects implied by a structural model. That is, now the aim is to estimate θ in the following equation

$$y_{c,t} = \text{cons} + \alpha_c + \beta(k, \theta) + \gamma_t + \varepsilon_{c,t},$$

where $\beta(k, \theta)$ is derived from an economic model and θ is a vector of model fundamentals. To achieve identification, this approach requires the function $\beta(k, \theta)$ to be sufficiently nonlinear in k. Under this condition, θ can be estimated consistently via a minimum distance procedure. Essentially, the structural parameters are identified from second and higher derivatives of the age effects. This approach ultimately facilitates identification of structural parameters associated with age effects without first identifying the age effects.

D.1.4 Heckman, Lochner and Taber (1998)

Heckman, Lochner and Taber (1998) (HLT hereafter) deal with the non-identification issue using economic theory. The HLT identifying assumption is that there is no human capital accumulation at the end of working life. This assumption could be justified in a Ben-Porath (1967) framework, where zero on-the-job investment in that stage is the optimal choice. HLT's approach is a perfect combination of the previous two approaches. On the one hand, the identifying assumption is essentially a normalization on the first difference of experience effects. On the other hand, this restriction is coming from economic theory and can be derived in a structural model of human capital investment.

LMPQS adopt the HLT method as their preferred estimates for returns to experience and set the flat spot phase as from 25 years of experience to 35 years of experience. We generally follow LMPQS, which in turn combines the identification assumption proposed by HLT with the procedure laid out by Deaton (1997). The basic idea is to recast the regression of interest (**??**) as

$$\ln w_{i,c,t} = \text{cons} + \alpha_c + \beta_k + gt + \tilde{\gamma}_t + \varepsilon_{i,c,t},$$

where $\tilde{\gamma}$ satisfies $\sum_t \tilde{\gamma}_t = 0$ and $s'\tilde{\gamma} = 0$. That is, they rewrite an arbitrary time series γ_t as a linear trend gt plus fluctuations $\tilde{\gamma}_t$. The benefit of this manipulation is that once a value of g is guessed, we can run Deaton's procedure on the deflated wage $\ln \tilde{w}_{i,c,t} = \ln w_{i,c,t} - gt$ and get estimates for cohort, experience, time effects under this particular guess g. The time trend is thus pinned down by the HLT assumption: we update the guess of g until the associated experience effects are the same for the two experience groups late in life presumed by the HLT assumption.

They obtain the estimates using an iterative procedure. First, start with a guess for the growth rate of a linear time trend. In practice, the guess is picked as the coefficient on the linear time trend term by regressing log wage on the set of dummies for experience groups and a linear time trend. Second, deflate the wage data using the guess of growth rate. Following Deaton's procedure laid out in Section D.1.2 but using log deflated wage as the dependent variable, we regress log deflated wage on a set of dummies for experience groups, cohort groups and d^*'s defined in the previous subsection D.1.2. Now check if the experience effects are sufficiently close between experience group 25 (or experience group 30) and experience group 35, according to a preset precision. If so, the convergence condition is satisfied. Otherwise, we update the guess for the growth rate by adding the growth rate of the currently estimated experience effects from experience group 25 (or experience group 30) to experience group 35 with a damping factor.

D.2 Skill-Biased Technological Change

Note that human capital price per efficiency unit under this production function is

$$p_s = \frac{\partial Y}{\partial H_s} = \left[(A_u H_u)^{\frac{\sigma-1}{\sigma}} + (A_s H_s)^{\frac{\sigma-1}{\sigma}} \right]^{\frac{\sigma}{\sigma-1}-1} (A_s H_s)^{\frac{\sigma-1}{\sigma}-1} A_s,$$

$$p_u = \frac{\partial Y}{\partial H_u} = \left[(A_u H_u)^{\frac{\sigma-1}{\sigma}} + (A_s H_s)^{\frac{\sigma-1}{\sigma}} \right]^{\frac{\sigma}{\sigma-1}-1} (A_u H_u)^{\frac{\sigma-1}{\sigma}-1} A_u,$$

with the time index t dropped for notational convenience but readers should bear in mind that all variables are allowed to change over time.

As a result, the relative human capital price estimated above reflects $\frac{p_s}{p_u} = \left(\frac{A_s}{A_u}\right)^{\frac{\sigma-1}{\sigma}} \left(\frac{H_s}{H_u}\right)^{\frac{\sigma-1}{\sigma}-1}$, or in log terms,

In the skill-biased technical change literature, (see excellent reviews by Acemoglu (2002) and Violante (2008)), it is often assumed that

$$Y(t) = \left[(B_s(t) L_s(t))^{\frac{\sigma-1}{\sigma}} + (B_u(t) L_u(t))^{\frac{\sigma-1}{\sigma}} \right]^{\frac{\sigma}{\sigma-1}}, \tag{6}$$

where L_s (L_u) is the labor supply of skilled (unskilled) workers, and the evolution in B_s/B_u is interpreted as the skill-biased technical change. Our formulation is consistent with it, and in fact generalizes it. Our formulation (5) distinguishes improvements in h_s/h_u with the technology that improves the productivity of skilled or unskilled human capital (i.e., A_s/A_u). These two forces together form the standard interpretation of skill-biased technical change B_s/B_u. To see this, rewrite the production function as

$$Y(t) = \left[(\underbrace{A_s(t) h_s(t)}_{B_s(t)} L_s(t))^{\frac{\sigma-1}{\sigma}} + (\underbrace{A_u(t) h_u(t)}_{B_u(t)} L_u(t))^{\frac{\sigma-1}{\sigma}} \right]^{\frac{\sigma}{\sigma-1}}.$$

In a competitive labor market, we have

$$w_s = \frac{\partial Y}{\partial L_s} = \left[(A_u h_u L_u)^{\frac{\sigma-1}{\sigma}} + (A_s h_s L_s)^{\frac{\sigma-1}{\sigma}} \right]^{\frac{\sigma-1}{\sigma}-1} (A_s h_s)^{\frac{\sigma-1}{\sigma}} L_s^{\frac{\sigma-1}{\sigma}-1},$$

$$w_u = \frac{\partial Y}{\partial L_u} = \left[(A_u h_u L_u)^{\frac{\sigma-1}{\sigma}} + (A_s h_s L_s)^{\frac{\sigma-1}{\sigma}} \right]^{\frac{\sigma-1}{\sigma}-1} (A_u h_u)^{\frac{\sigma-1}{\sigma}} L_u^{\frac{\sigma-1}{\sigma}-1}.$$

Therefore the college earnings premium can be written as $\frac{w_s}{w_u} = \left(\frac{A_s h_s}{A_u h_u}\right)^{\frac{\sigma-1}{\sigma}} \left(\frac{L_s}{L_u}\right)^{\frac{\sigma-1}{\sigma}-1}$, or in log terms,

$$\ln\left(\frac{w_s}{w_u}\right)=\frac{\sigma-1}{\sigma}\ln\left(\frac{A_s}{A_u}\right)+\frac{\sigma-1}{\sigma}\ln\left(\frac{h_s}{h_u}\right)-\frac{1}{\sigma}\ln\left(\frac{L_s}{L_u}\right).$$

Note that with the typical formulation (6) such as in Katz and Murphy (1992), we will have

$$\ln\left(\frac{w_s}{w_u}\right)=\frac{\sigma-1}{\sigma}\ln\left(\frac{B_s}{B_u}\right)-\frac{1}{\sigma}\ln\left(\frac{L_s}{L_u}\right).$$

In other words, B_s/B_u in their model is equivalent to $(A_s h_s)\,/\,(A_u h_u)$ in our formulation, and it is really a combination of the skill-biased technical change and the changes in relative human capital per worker between skilled and unskilled workers. Katz and Murphy (1992) estimate this equation, using 40 years of U.S. data and their benchmark estimate is $\sigma = 1.4$. In log changes, we have

$$\Delta\ln\left(\frac{w_s}{w_u}\right)=\frac{\sigma-1}{\sigma}\Delta\ln\left(\frac{A_s}{A_u}\right)+\frac{\sigma-1}{\sigma}\Delta\ln\left(\frac{h_s}{h_u}\right)-\frac{1}{\sigma}\Delta\ln\left(\frac{L_s}{L_u}\right).$$

As we can see, there are three factors that affect the college premium. An increase in relative labor supply L_s/L_u, holding everything else fixed, decreases relative wage. An increase in the relative human capital efficiency units h_s/h_u has two effects. First, it decreases the relative human capital prices p_s/p_u. Second, it increases skilled-labor's relative earnings capacity. The overall effect is positive if $\sigma > 1$ when the second effect dominates the first. The effect of the skill-biased technical changes (increasing A_s/A_u) on college premium depends on σ, too.