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

Hanming Fang† Xincheng Qiu‡

October 20, 2021

Preliminary

(*click here for latest version*)

Abstract

We document stark differences in the cross-sectional age-earnings profiles between the U.S. and China, the two largest economies in the world, during the past thirty years. We find that, first, the peak age in cross-sectional age-earnings profiles, which we refer to as the "golden age," stays almost constant at around 45-50 years old in the U.S., but decreases sharply from 55 to around 35 years old in China; second, the age-specific earnings grow drastically in China, but stay almost stagnant in the U.S.; and third, the cross-sectional and life-cycle age-earnings profiles look remarkably similar in the U.S., but differ substantially in China. We propose and empirically implement a unified decomposition framework to infer from the repeated cross-sectional earnings data the life-cycle human capital accumulation (the experience effect), the inter-cohort productivity growth (the cohort effect), and the human capital price changes over time (the time effect), under an identifying assumption that the growth of the experience effect stops at the end of the working career. The decomposition suggests that China has experienced a much larger inter-cohort productivity growth 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 also use the inferred components to revisit several important and classical applications in macroeconomics and labor economics, including the growth accounting and the estimation of the TFP growth, and the college wage premium and the skill-biased technical change.

Keywords: Age-Earnings Profiles, Human Capital, Life Cycle

JEL Codes: E24, E25, J24, J31, O47

---

*We would like to thank Hal Cole, Dirk Krueger, Ernest Liu, Iourii Manovskii, Nancy Qian, José-Víctor Ríos-Rull, Ke Shen, Michael Zheng Song, Feng Wang, Shang-Jin Wei, Yiqing Xu, Bernard Yeung, Xiaodong Zhu, and conference participants at the Chinese Economic Associatoin Annual Conference (2019), the NBER Chinese Economy Working Group (2020), the 8th Annual ABFER Conference (2021) and seminar participants at Fudan University, Peking University, ShanghaiTech University, USC Marshall, and University of Pennsylvania for helpful comments and suggestions. We are responsible for all remaining errors.

†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).[1] 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 level off, and eventually decline after reaching the peak earnings age. In this paper, we call the age group that achieves the highest average earnings in a cross-sectional age-earnings profile the "*golden age*." For instance, the "golden age" in the United States has stayed at around 50 years old, meaning that 50-year-old workers tend to have the highest average earnings among all age groups in a cross-sectional labor market data.

In this paper, we start with a systematic comparison of the age-earnings profiles between the U.S. and China, the two largest economies in the world. We document three striking differences between the two labor markets during the last thirty years:

- 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.
- Age-specific (real) earnings almost stagnate in the U.S. but grow drastically in China.
- The cross-sectional and life-cycle age-earnings profiles look remarkably similar in the U.S. but differ substantially in China.

We then 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 the inter-cohort productivity growth, or, the relative human capital level of a cohort of workers at the time when they enter the labor market; and 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. Lagakos et al. (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 growth in experience effect in a worker's late career, as implied by the standard human capital investment theory (Ben-Porath, 1967), which predicts no incentive to invest in human capital at the end of one's 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

[1] We do not attempt to review the large literature. See Heckman et al. (2006) for an excellent survey.

models with learning (Jovanovic, 1979).[2] This identification idea is exploited originally by Heckman et al. (1998) (hereafter, HLT), and more recently also by, Huggett et al. (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 the 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 older; when the inter-cohort productivity growth prevails, the "golden age" tends to be younger. It is instructive to think of two extreme cases. Suppose in one extreme that there is no inter-cohort productivity growth. Then the cross-sectional age-earnings profile simply reflects returns to experience and achieves its highest value at the oldest age. Suppose in the other extreme that there is no returns to experience. Then the cross-sectional age-earnings profile simply reflects differences in cohort-specific productivity and achieves it highest value at the youngest age, or the most recent cohort. We find that in China, the inter-cohort productivity growth was very rapid in the last thirty years, thus it wins the race against the experience effect. As a result, the golden age has experienced a gradual decline over the years. In contrast, in the U.S., the inter-cohort productivity growth is dimmed compared to a high return to experience, resulting in rather old golden ages.

Second, we find that the rental price to human capital (the time effect) increased much faster over the last thirty years in China than in 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.[3] Thus, the experience effect is the main driving force of both the cross-sectional and the life-cycle age-earnings profiles in the U.S. As a result, both cross-sectional and life-cycle profiles are close to the experience effects in such a stationary environment 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 drastically different from cross-sectional age-earnings profiles.

We also use our decomposition to revisit some important accounting exercises in macroeconomics and labor economics. First, by isolating the human capital prices changes obtained from the time effects, our decomposition delivers a notion of effective human capital quantities,

[2] 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.

[3] In fact, we find that the real rental price of human capital declined by about 1% per year, if anything, in the U.S. in the last thirty years.

which comprises both the experience and the cohort effects. The decomposition allows us to conduct a growth accounting exercise that properly accounts for the evolution of human capital. Adjusting for changes in the effective human capital, we obtain a series of estimated total factor productivity (TFP) growth that is lower than previous estimates.[4] Second, we also implement the same decomposition separately for college-educated and high-school-educated workers. We use the education-specific decompositions to obtain an estimated series for skill-biased technical change where relative human capital quantities between high-skilled and low-skilled workers are allowed to change over time.

The remainder of the paper is structured as follows. In Section 2, we describe the facts on age-earnings profiles in the U.S. and China; in Section 3, we present the theoretical framework and discuss issues on identification; in Section 4, we describe the main results from the decomposition and the applications; in Section 5, we apply the decomposition results to revisit several classical and important questions in macroeconomics and labor economics. Section 5.1 revisits the growth accounting exercise by adjusting for human capital changes based on our estimates; Section 5.2 extends the benchmark framework to present results by education groups and revisits skill-biased technical changes by accounting for potentially differential human capital changes of different skill groups; Section 5.3 simulates a counterfactual economy that starts to slow down after a fast-growing period. Finally, in Section 6, we conclude and discuss potential directions for future research.

2 Facts

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

We use the 1986-2012 waves of 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.[5]

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 impacts of extreme values, we drop outliers defined as earnings being in the top 2.5% and bottom 2.5% in each year. We normalize all earnings to the 2015 year using CPI. Individuals are weighted by the person-level ASEC weight. Figure 1a reveals that,

[4] This idea has also been exploited by Bowlus and Robinson (2012).

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

first, the "golden age" in the U.S. is relatively stable at around 50 years old during the past three decades; 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 between Figures 1a and 1b. First, Chinese workers have experienced a dramatic increase in real earnings in the past 30 years for all age groups. It is reflected in the large vertical upward 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. 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" in 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.[6]

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 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 shows 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 strong downward trend in the golden ages from 1986 to 2012, decreasing from more than 55 years old to around 35 years old.

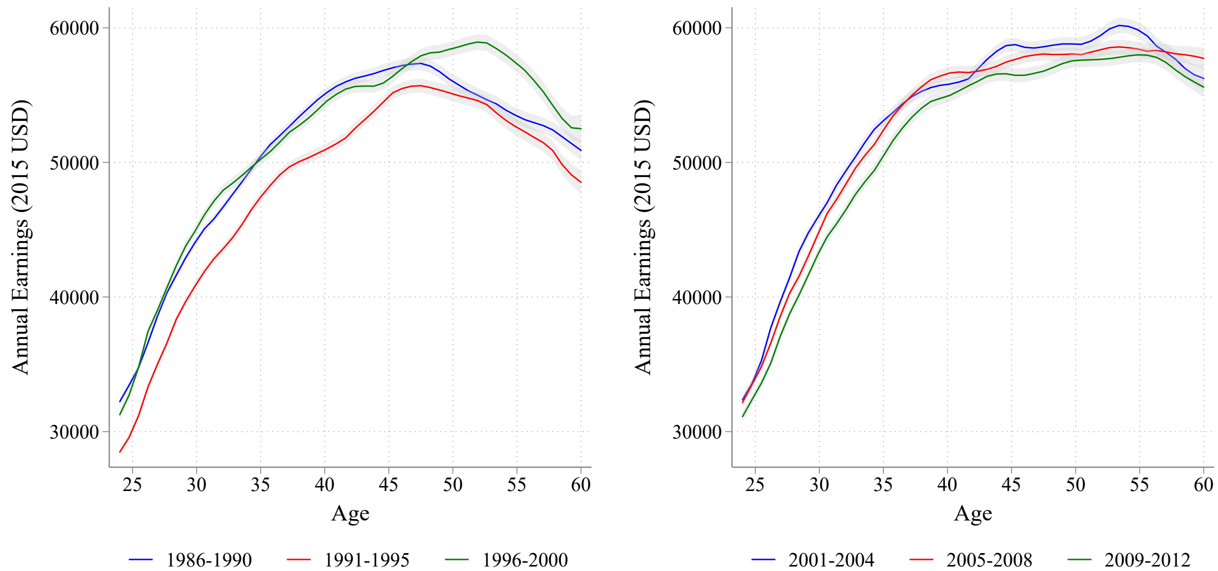
Robustness of the Emprical Facts. The above facts are robust to a series of alternative sample restrictions and estimation methods. First, as previously mentioned, by its design, UHS only covers urban households.[7] The stark difference in the evolution of cross-sectional age-earnings profiles between the U.S. and China, however, is not merely a result of the sample restriction on urban workers in UHS. In Figure A.1, we restrict our attention to CPS households that live in metropolitan areas, which is the closest geographic sample choice to urban households in UHS. There is virtually no difference in the shape of age-earning profiles, although the level of earnings

[6] Song and Yang (2010) notice a dramatic flattening of age-earnings profile during the fast-growing period of China. 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 income age.

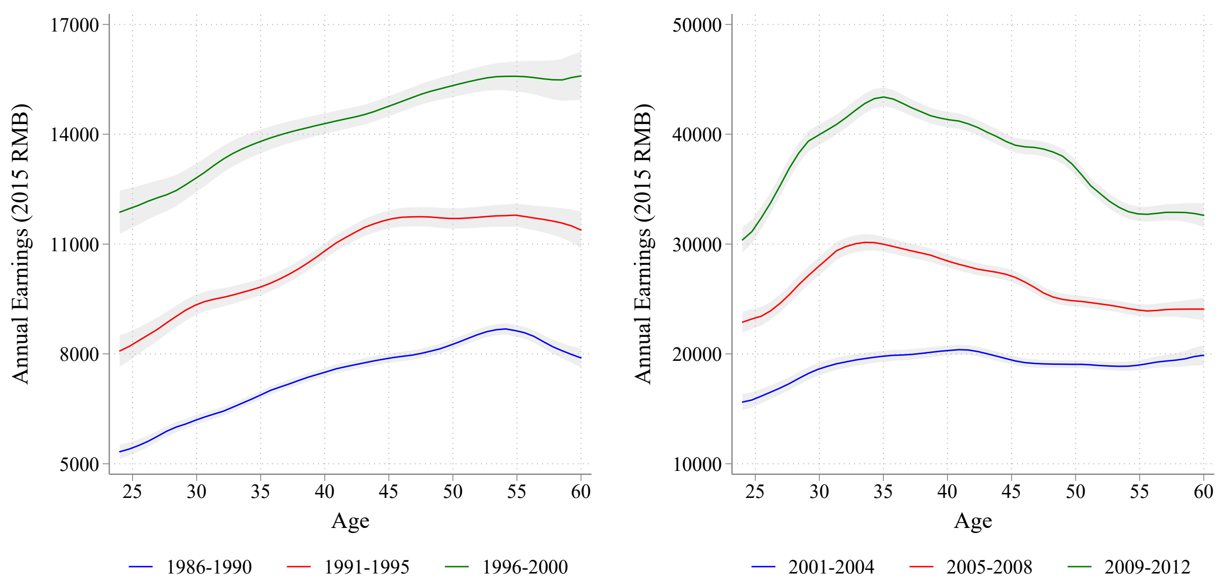
[7] Prior to 2002, UHS only covers households with local urban *hukou*. Although UHS started to include households without local urban *hukou* since 2002, the coverage is so low that non-local-*hukou* residents are under represented. See, for example, the discussion in Ge and Yang (2014).

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

(a) U.S.

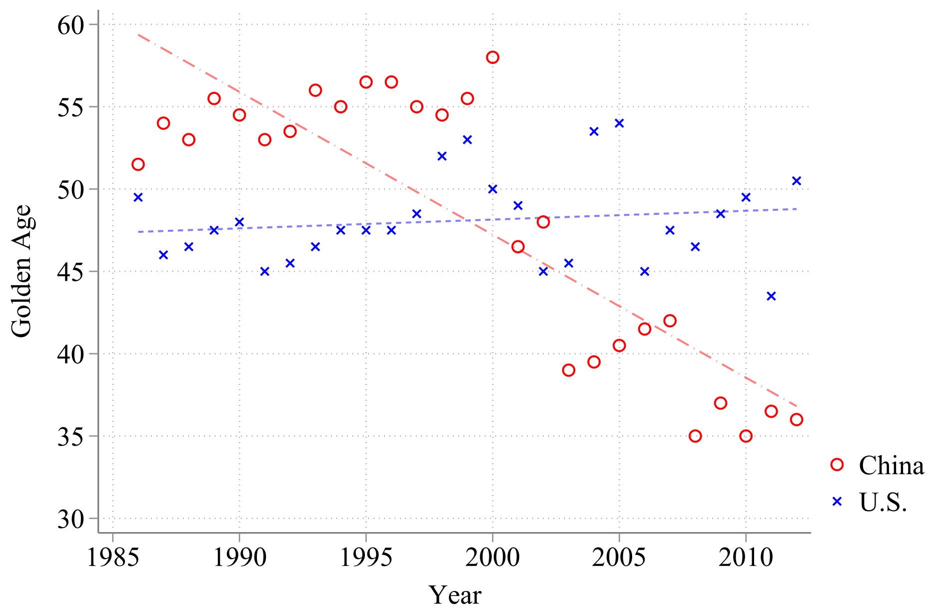


(b) China



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.

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



Notes: This figure shows the evolution of the cross-sectional "golden age" in the U.S. and China. The blue cross marker denotes the point estimate of the golden age in the U.S. and the red circle marker denotes the point estimate of the golden age in China. The blue short-dashed line and the red dash-dotted lines are the respective linear time trend in the evolution of the golden age in each country.

is on average higher for workers in metropolitan areas than those 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 investigate how the labor market evolves over time, it is crucial to provide a comprehensive set of evidence that spans a long period of time. So we choose not to drop any time periods in our main analysis. Instead, we verify 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 B.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 (Liaoning), East (Shanghai), South Central (Guangdong), Southwest (Sichuan), respectively. To mitigate the concern for representativeness, we replicate Figure 1b for a much larger set of 15 provinces covering all 6 regions in Figure A.2 in Appendix A.[8] The tradeoff is that, UHS micro data with the whole set of these 15 provinces is only available for the years from

[8]The 15 provinces 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.

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. During the early 2000s, the 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 address this question for a sub-period from 2002 to 2006 when UHS does collect information on "total number of hours worked last month." A typical month contains about $30/7 \approx 4.3$ weeks, so we use this number to convert the monthly measure of hours worked to a weekly measure in order to facilitate comparison with CPS.[9] Figure A.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 are 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 the U.S., and the bottom panel for China.

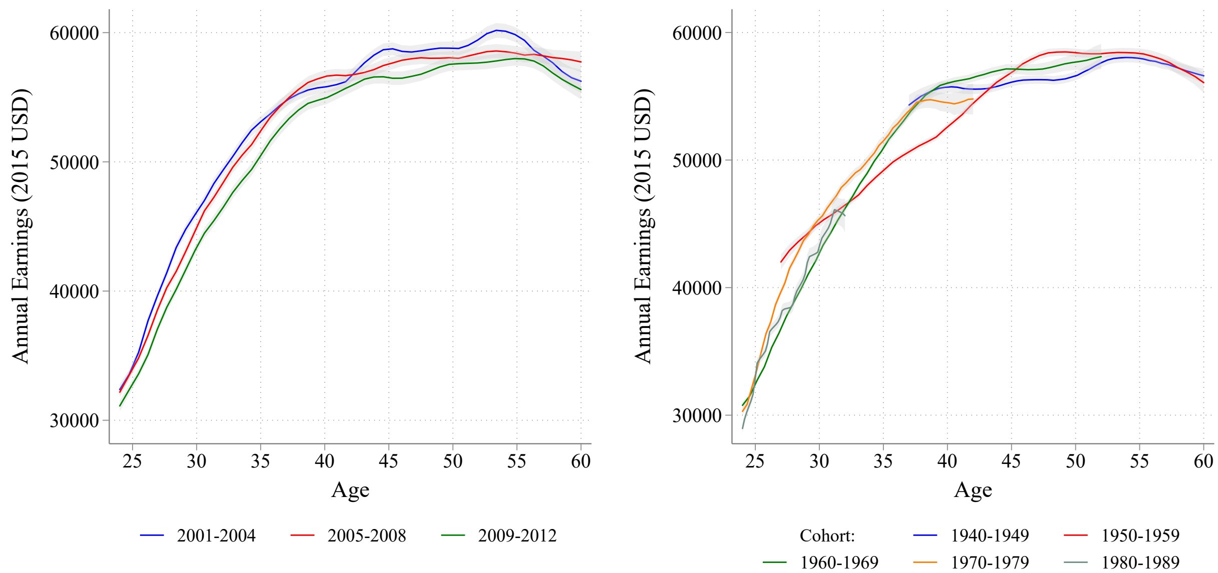
In the U.S. (Figure 3a), cohorts with year of birth expanding half a century share remarkably similar life-cycle earnings paths. Furthermore, life-cycle profiles on the right of Figure 3a resemble the cross-sectional profiles on the left (which is reproduced from the right panel of Figure 1a) closely, 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 30-year-old worker who wants to predict his (real) earnings 10 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 correct to interpret cross-sectional age-earnings profiles as life-cycle patterns, in practice they are close to each other for the U.S. case. In other words, stationarity is an reasonable assumption when studying the U.S. earnings profiles.

However, as shown in Figure 3b, the life-cycle patterns of different cohorts differ drastically for China. More recent cohorts enjoy both much higher earnings and steeper life-cycle earnings

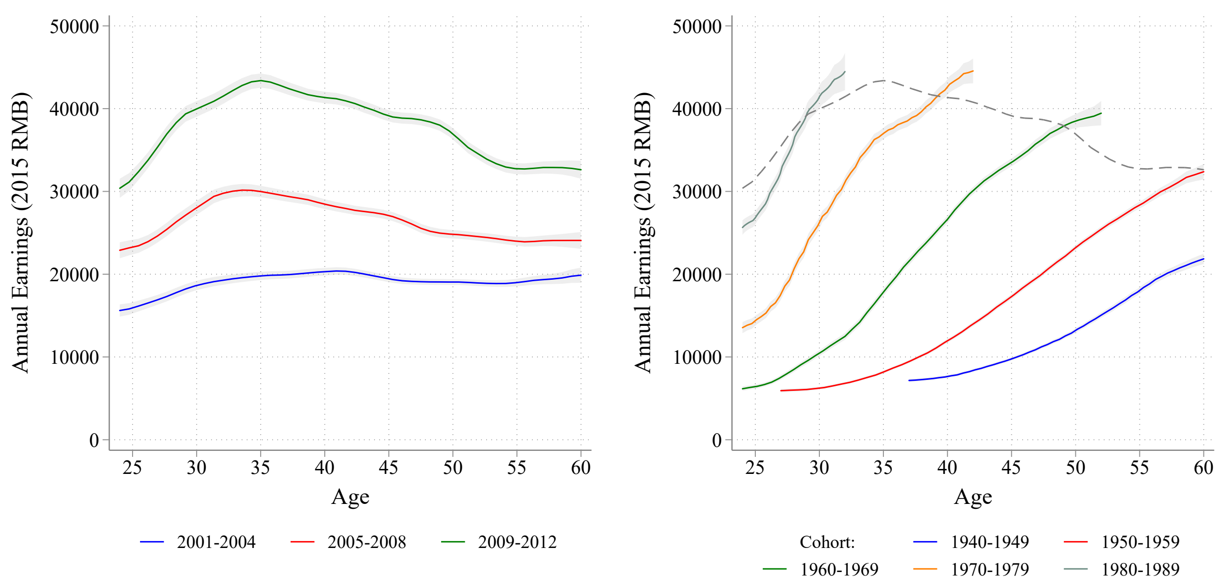
[9]The corresponding variable in CPS is "total number of hours usually worked per week over all jobs the year prior to the survey."

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.

growth. These life-cycle profiles also demonstrate no resemblance at all to the cross-sectional profiles, although they are actually linked to each other.[10] Note that both the left and the right panel are plotted using the same underlying data, but just with different ways to visualize the data. It is perhaps not surprising that in a fast-growing economy such as China, stationarity is not a valid approximation.

The next section provides a framework to organize the facts documented in this section.

3 A Decomposition Framework

3.1 Setup

We consider a competitive interpretation of of the observed wages such that the observed level of a worker's wage is the product of the *price* of human capital and the *quantity* of human capital this worker supplies. Denote $W_{i,t}$ as 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

$$W_{i,t} = P_t \cdot H_{i,t}. \tag{1}$$

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.[11] Taking logs on both sides of Eq. (1), we have:

$$w_{i,t} = p_t + h_{i,t}, \tag{2}$$

where for notational convenience, we use the lower case letters to represent the log values and the upper case letters to represent the 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,

$$h_{c,t} = \mathbb{E}_i \left[h_{i,t} | c\left(i\right) = c, t \right]$$

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

[10] Suppose we keep track of the same time period, say, 2010, across different life-cycle profiles. That is, we connect the point of age 30 in the life-cycle profile for cohort 1980, age 40 for cohort 1970, age 50 for cohort 1960 and so on, then we are able to reproduce the cross-sectional profile for 2010, as (conceptually) illustrated by the dashed gray line, which is reproduced from the 2009-2012 cross-sectional profile.

[11] This assumption rules out potential complementarity between different types of skills. We relax this assumption in Section 5.2.3.

(conditional on cohort c and time t). Therefore, we can rewrite equation (2) as

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

with $\mathbb{E}_i \left[\epsilon_{i,t} | c\left(i\right) = c, t\right] = 0$, for all c and t, where 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 $\{p_t + \lambda, h_{c,t} - \lambda\}$ for any constant λ ("normalization"), but also observationally equivalent to $\{p_t + \lambda_t, h_{c,t} - \lambda_t\}$ 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 components $h_{c,t} = s_c + r^c_{t-c}$, 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.[12] Using this notation, we can decompose (log) wages into time effects, cohort effects, and experience effects,

$$w_{i,t} = p_t + s_c + r^c_k + \epsilon_{i,t}, \tag{3}$$

with $\mathbb{E}_i \left[\epsilon_{i,t} | i \in c, t\right] = 0$, where (i) time effects p_t reflect the human capital prices, (ii) cohort effects s_c represent 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 common practice in the literature is to impose the returns to experience to be the same across cohorts, i.e., to restrict $r^c_k \equiv r_k, \forall c$, which gives rise to a variant of equation (3):

$$w_{i,t} = p_t + s_c + r_k + \epsilon_{i,t}. \tag{3'}$$

The main benefit of restricting $\left\{r^c_k\right\}$ 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-cycle in the data.[13] We follow this common practice mainly due to data limitations.

[12]We do not model the labor market entry decision and abstract from the difference between age and experience. In other words, workers of the same cohort are assumed to enter the labor market at the same age; thus, we use age and experience interchangably. In Section 5.2, 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 a robustness exercise in Table 1, we consider an alternative definition of experience as years since the first job, where workers of the same cohort are allowed to have different levels of experience even at the same age.

[13]It is worth noting that this restriction does not by itself solve the non-identification problem mentioned above. Even under this restriction, we still cannot isolate year effects, cohort effects, and experience effects without imposing

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).[14] The average (log) earnings of workers with experience k at time t is

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

where the expectation is taken over individuals i for given time t and experience k (and hence cohort is 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

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

where we move the subscripts to inside the brackets 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 the return to experience is a function of experience k.

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

$$\frac{\partial}{\partial k} w(k;t) = \dot{r}(k) - \dot{s}(t-k), \tag{4}$$

which is positive if $\dot{r}(k) > \dot{s}(t-k)$ and negative if $\dot{r}(k) < \dot{s}(t-k)$.[15] Note that both r and s are in logs, so the interpretations of $\dot{r}$ and $\dot{s}$ are 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:

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

Though straightforward, Proposition 1 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; hence the cross-sectional age-earnings profiles tend to be flat or even downward sloping. If life-cycle human capital growth dominates, then the older cohorts tend to have higher relative earnings, and then the cross-sectional age-earnings profiles tend to be steeply upward sloping. To further illustrate

additional assumptions, due to the perfect collinearity among year, cohort, and experience that $k = t - c$.

[14] We abstract from endogenous retirement decisions; hence the retirement age is set exogenously in this paper.

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

this proposition, it is instructive to consider two extreme cases. Consider a hypothetical economy where there is no inter-cohort human capital growth and each cohort is equally productive conditional on age. In this case, the oldest group will earn the highest wages as long as returns to experience are positive. Consider another hypothetical economy where there is no returns to experience but more recent cohorts are getting more productive. In this case, it is the youngest group that earns the most.

The cross-sectional "golden age" at time t can be defined as:

$$k^*(t) = \arg\max_{k\in[0,R]} w(k;t). \tag{5}$$

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, defined by (5), 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 simple framework 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 $\{\tilde{w}(k;c)\}_{k=0}^{R}$. The average (log) earnings of workers in cohort c with experience k is

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

where the expectation is taken over individuals i for given cohort c and experience k (and hence time is given by $t=c+k$). Due to the conditional mean zero property $\mathbb{E}_i\left[\epsilon_{i,t}|i\in c,t\right]=0, \forall c,t$, 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

$$\frac{\partial}{\partial k}\tilde{w}(k;c) = \dot{r}(k) + \dot{p}(c+k). \tag{6}$$

Comparing equation (4) with equation (6) 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 Eqs. (4) and (6) 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 the U.S. (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, such a tale serves as a promising description of what happened in the two labor markets in the past three decades.

This section has discussed conceptually how the returns to experience $\dot{r}$, inter-cohort human capital growth $\dot{s}$, and human capital price changes $\dot{p}$ affect the cross-sectional and life-cycle profiles, and provides a narrative of how they explain the striking differences we document for the U.S. and China. The next section will provide empirical estimates for these three elements. Before that, we discuss the identification of these three elements.

3.4 Identification

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

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

where i refers to an individual observation ,and t, 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 (3′) here:

$$w_{i,t} = p_t + s_c + r_k + \epsilon_{i,t}. \tag{3'}$$

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 $\mathbb{E}_i [\epsilon_{i,t} | i \in c, t] = 0, \forall c, t$.

Two issues are worth comments. 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 1935-39, year 1986, and experience $0-4$." The log earning of the baseline group is loaded on a constant term. 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.[16]

[16] 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 interrelated. We are facing an issue of near multicollinearity and

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 C.1, but summarize briefly the main message here. Deaton (1997) views time effects as capturing the cyclical fluctuations and impose an identifying assumption that time effects are orthogonal to a linear trend and sum up to zero. In this way, all growth is due to cohort effects.[17] 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 the theoretical justification for a natural identifying restriction such that *the growth of the experience effect is zero in the last few years of working life*, which is also the main identifying restriction in LMPQS.[18]

We now intuitively and constructively explain how this restriction facilitates identification of human capital prices (time effects), cohort-specific initial human capital (cohort effects), and human capital accumulation paths (experience effects). Suppose one assumes there is no human capital accumulation (say) from $R-1$ to R years old.[19] First, comparing the wages of $(R-1)$-year-old workers in year $t-1$ and R-year-old workers in t identifies the time effect from $t-1$ to t because (1) we are comparing the same cohort so by definition the cohort effect does not contribute to the wage differences; and (2) by the identification assumption there is no growth in the experience effect, thus the experience effect does not contribute to the wage differences either.

Second, comparing the wages of $(a-1)$-year-old workers in $t-1$ and a-year-old workers in t provides information for the experience effect from $a-1$ to a because (1) we are again comparing the same cohort so the cohort effect does not contribute to the wage differences, and (2) we have already backed out the year effect from $t-1$ to t from the previous step.

Third, further comparing the wages of $(a-1)$-year-old workers and a-year-old workers in t gives the cohort effect from cohort $c = t-a$ 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 $a-1$ to a.

In general, the HLT identification approach requires the researcher to pick her preferable values for a "flat region" of experience, for which there is no growth in the experience effect. It could also be extended to allow for a pre-specified human capital depreciation rate. We acknowledge that either input is somewhat arbitrary and cannot be inferred internally from data

the standard OLS estimator will generate imprecise estimates. Furthermore, the actual experience is an endogenous labor market outcome so that controlling for the actual experience may instead contaminate the estimates (see pp. 64-68 in Angrist and Pischke, 2009, for discussion).

[17] LMPQS also considers an opposite extreme that all growth is due to time, and a specification in between with cohort and time each contributes half.

[18] The same identification restriction was also adopted by Bowlus and Robinson (2012) and Huggett et al. (2011).

[19] The actual identifying assumption and algorithm is more sophisticated, but we provide the intuition in a nutshell here for transparency. See Appendix C.2 for a detailed explanation on the iterative procedure in implementation.

(which is rooted in the non-identification problem discussed previously). In particular, HLT assume a zero human capital depreciation rate, consistent with estimates reported by Browning et al. (1999). We follow this assumption, as many other papers studying life-cycle human capital accumulation also do (e.g., Kuruscu, 2006). The exact choice for the flat experience effect region is somewhat *ad hoc*, though the life-cycle models of human capital investment incentives provide guidance in this regard. We follow LMPQS by considering 40 years of experience and assuming there are no growth in the experience effects in the last ten years in the baseline specification. Bowlus and Robinson (2012) attempt to determine the flat experience effect age regions more carefully and prefer the flat age ranges to be 50-59 for college graduates and 46-55 for high school graduates. Our choice of the flat region largely overlaps with theirs. We have also investigated alternative specifications discussed below to rule out various concerns.

4 Decomposition

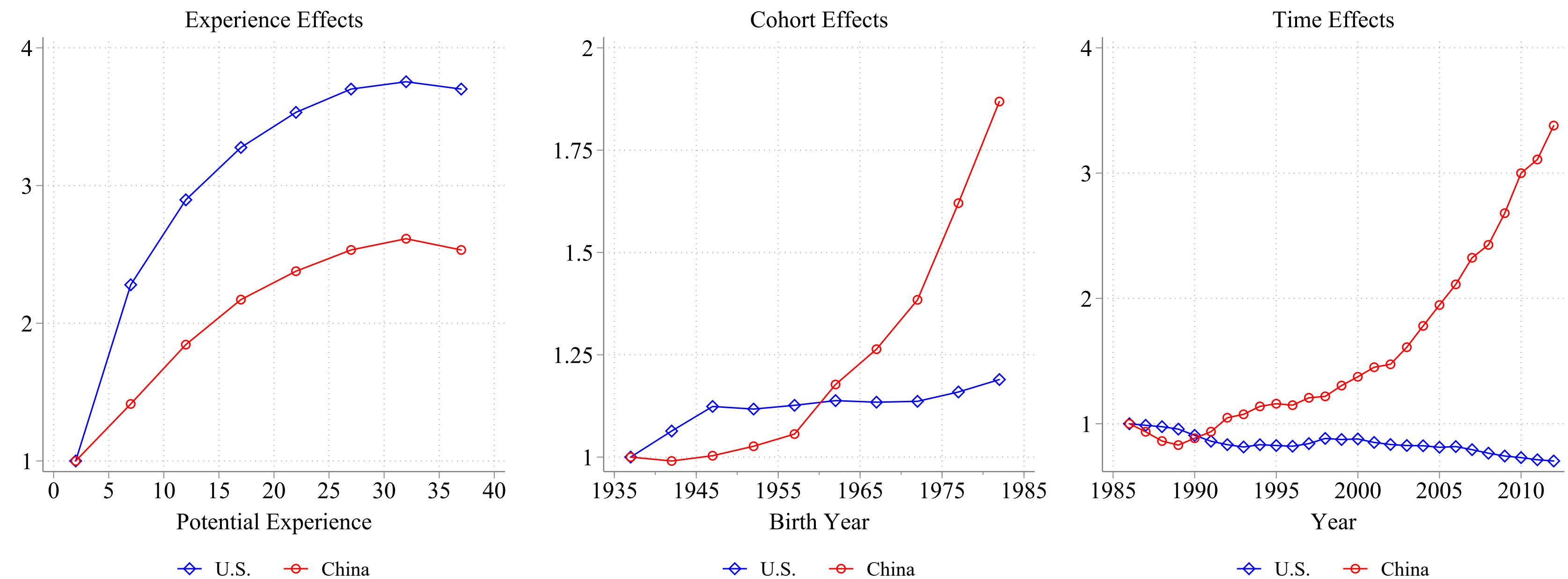
Figure 4 performs an HLT decomposition of earnings among experience, cohort, and time effects. Specifically, we estimate the experience effects (relative to labor market entry) in 5-year bins, cohort effects (relative to the 1935-1939 birth cohorts) in 5-year bins, and year effects (relative to 1986) year by year.

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 U.S. 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 the U.S., most of which happened from cohort 1935 to cohort 1950. In China, the inter-cohort productivity growth is almost 90%, most of which happened only since cohort 1960. (3) The time effect, i..e., the human capital price effect, grows by more than three folds in China from 1986 to 2012, while it is negligible in the U.S. (if anything, it declines at a rate of about 1% per year).

Robustness of the Decomposition Results. Before turning into the detailed discussion on the interpretations and the implications of the decomposition, we emphasize 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. In Row 3, we restrict attention to the only 4 provinces that are covered in the UHS sample throughout including 2010-2012.

Second, we consider alternative definitions for potential experience. In the baseline, potential experience is imputed as experience $:= \min\{\text{age} - \text{edu} - 6, \text{age} - 18\}$. 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

Figure 4: Decomposition



Notes: This figure shows the decomposition results of experience, cohort, and time effects in the U.S. (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 the U.S. and China. The first row is the baseline result as discussed in the main text. Row 2 controls for state fixed effect for the U.S. and provincial fixed effect for China, and Row 3 focuses on the 4 provinces 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 job, which is only available 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 the U.S.

definition for potential experience as experience := 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 with no growth in the 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 last five years within that range. Although the magnitude of experience effects varies somewhat across specifications as recognized by LMPQS, the general pattern we focus on is not affected by the choice of specification, especially in terms of the comparison between 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 and 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 the hours worked by the median men within each group are much more likely to be similar.[20] Hence in Row 9, we perform a quantile regression analysis to estimate the 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 5.2, 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 an increase in productivity for each education group (within-group effects). We will revisit the discussion of different education groups in Section 5.2.

Finally, since we do not have information on hours worked in UHS, we restrict attention mostly to earnings for the U.S. as well, for a 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 A.3 for direct evidence).

[20] Casanova (2013) documents a phenomenon of partial retirement, i.e., as workers age, an increasing fraction is transitioning from full-time into part-time work. Without proper information on hours worked or part-time employment status in the UHS, we cannot directly address the partial retirement issue. However, using median regression techniques minimizes the potential bias given that that a median worker is working full time.

That said, the estimated cohort effect and time effect are still in the ballpark consistent with other specifications.

4.1 Experience Effect: Life-Cycle Human Capital Accumulation

Consistent with the finding in Lagakos et al. (2018) and 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.[21]

In the classical life-cycle human capital accumulation literature, pioneered by Ben-Porath (1967) and 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.[22] 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 decomposition suggest that $P_t \equiv P, \forall t$ is not an innocuous assumption for the case of fast-growing economies like China, although it is a rather good approximation for the U.S.

Although we take a simple abstraction to model wages as being determined in a competitive labor market with perfect information, and hence interpret experience effects are life cycle human capital accumulation, it is worth pointing out that there are other models consistent with the estimated experience effects. For instance, one could introduce search frictions and allow for on-the-job search (e.g., Burdett and Mortensen, 1998). There the experience effects reflect

[21] 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 shown in Figure A.3, there is a large hours increase (or part-time to full-time transition) for very young workers in the U.S. We provide an additional decomposition using hourly wage in Figure A.4 and Row 11 of Table 1. The result is consistent with the experience effects reported by LMPQS, which reassures the validity of our decomposition.

[22] In Ben-Porath (1967)'s framework, time devoted to working and learning are distinct concepts in 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 et al. (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 such measurement challenge of time allocation between working and training.

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 about the match quality.

How can we explain the steeper returns to experience in the 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 with the moments reported in their paper. Yet another new, potential explanation for why the experience effect is higher in the U.S. than in China is that workers' skills are multidimensional, and the speed of accumulation may differ for different dimensions of skills (see Lise and Postel-Vinay, 2020, for example). If the more developed economies value cognitive skills more as supposed to manual skills than the less developed economies, and if cognitive skills have faster accumulation than manual skills, then the measured experience effect would be higher in developed economies. 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 capture the inter-cohort growth of human capital at any given age as well as the life-time human capital. The middle panel of Figure 4 shows that China has experienced rapid human capital growth among cohorts born after 1960. 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 several sources. For example, later cohorts receive more and/or higher-quality education, stay in better health conditions, or are equipped with more pertinent skills to perform the most 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, at a pace of around a 1% decline per year.

5 Applications

We now use our decomposition results to revisit several classical and important questions in macroeconomics and labor economics. First, we revisit the growth accounting exercise by adjusting for human capital changes based on our estimates. Second, we revisit skill-biased technical changes by accounting for potentially differential human capital changes of different skill groups. Lastly, we simulate a counterfactual economy that starts to slow down after a fast-growing period.

5.1 Growth Accounting and the Estimation of the TFP Growth

The first application of our decomposition is to fine tune a growth accounting. Consider a standard Cobb-Douglas aggregate production function

$$Y_t = A_t K_t^{\alpha_t} H_t^{1-\alpha_t}, \tag{7}$$

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 $X \in \{Y, K, H\}$ and L is the total number of workers. The output per worker can be expressed as $y_t = A_t k_t^{\alpha_t} h_t^{1-\alpha_t}$.

First, we could directly measure y_t, k_t, and α_t in the data — this is the standard part. Specifically, 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 s_t, all of which are from the Penn World Table 9.0 (Feenstra et al., 2015) provided by the Federal Reserve Bank of St. Louis website.[23,24] 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. Under the competitive framework, the labor share is equal to $1 - \alpha_t$, which we set to s_t.

Second, we construct human capital changes based on estimates from the decomposition in the Section 4 — this is the new part. Specifically, we construct the average human capital at time t (up to a normalization) as the weighted average of the human capital of each cohort group and experience group

$$h_t = \sum_c \sum_k \exp(s_c + r_k)\omega(c, k; t),$$

where $\omega(c, k; t)$ gives the employment share of workers of cohort c and experience k at time t,

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

[24] 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.

and estimates for s_c and r_k are obtained from our decomposition. We could therefore get an estimated series for changes in human capital per worker.[25]

Taking stock, TFP changes could be measured as a residual from

$$\mathrm{d}\ln y_t = \mathrm{d}\ln \tilde{A}_t + \alpha_t \,\mathrm{d}\ln k_t + (1-\alpha_t)\,\mathrm{d}\ln h_t, \tag{8}$$

where $\mathrm{d}\ln \tilde{A}_t := \mathrm{d}\ln A_t + (\ln k_t - \ln h_t)\,\mathrm{d}\alpha_t$. However, since our decomposition can only deliver changes but not levels, 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 $\mathrm{d}\ln A_t$.[26] In practice, as long as the annual labor share change $\mathrm{d}\alpha_t$ is small (it indeed is), this serves as a reasonable approximation to TFP changes.[27] Such approximation is commonly adopted in growth accounting (e.g., Fernald, 2014).

We present the contribution of each source — physical capital per worker, human capital per worker, and the residual — to the growth of GDP per worker in Figure 5. We find that all three sources contribute almost equally to the U.S. growth, with the contribution of human capital slightly dominating the other two sources. The picture is quite different in China. Although the absolute level of the growth in human capital is larger in China than in the U.S., the relative contribution of human capital turns out be the least important to China's growth. But this is merely a result of an even faster speed at which the physical capital and TFP grow in China. In fact, physical capital is responsible for almost 60% of the growth in GDP per worker, and TFP for almost another 30% in China.

Our findings should be viewed as a refinement of the existing growth accounting results in the literature by providing a more "under-the-hood" examination of what is often called the "black-box" TFP growth. Although one should not expect the levels of our TFP estimates to be identical to other estimates, because our growth accounting procedure incorporates inter-cohort human capital improvements and life-cycle human capital accumulation, and it is well-known that TFP is a model-based concept, it would give us additional reassurance in the accuracy of the method if it were able to track the broad movements over time in other prominent TFP estimates. Now we put the results to such a test. The growth accounting results in Figure 5 suggest little TFP growth in the U.S. since mid-2000s, which is consistent with the productivity slowdown during the same period according to estimates by Fernald (2015). For China, Figure 5 shows that TFP increases by close to 60% from 1986 to 2012, almost all of which are reaped since 2000. This is consistent with the estimates by Zhu (2012), who also find a much larger TFP growth occurred after the late 90s.[28] This is a period when many prominent economic reforms have happened,

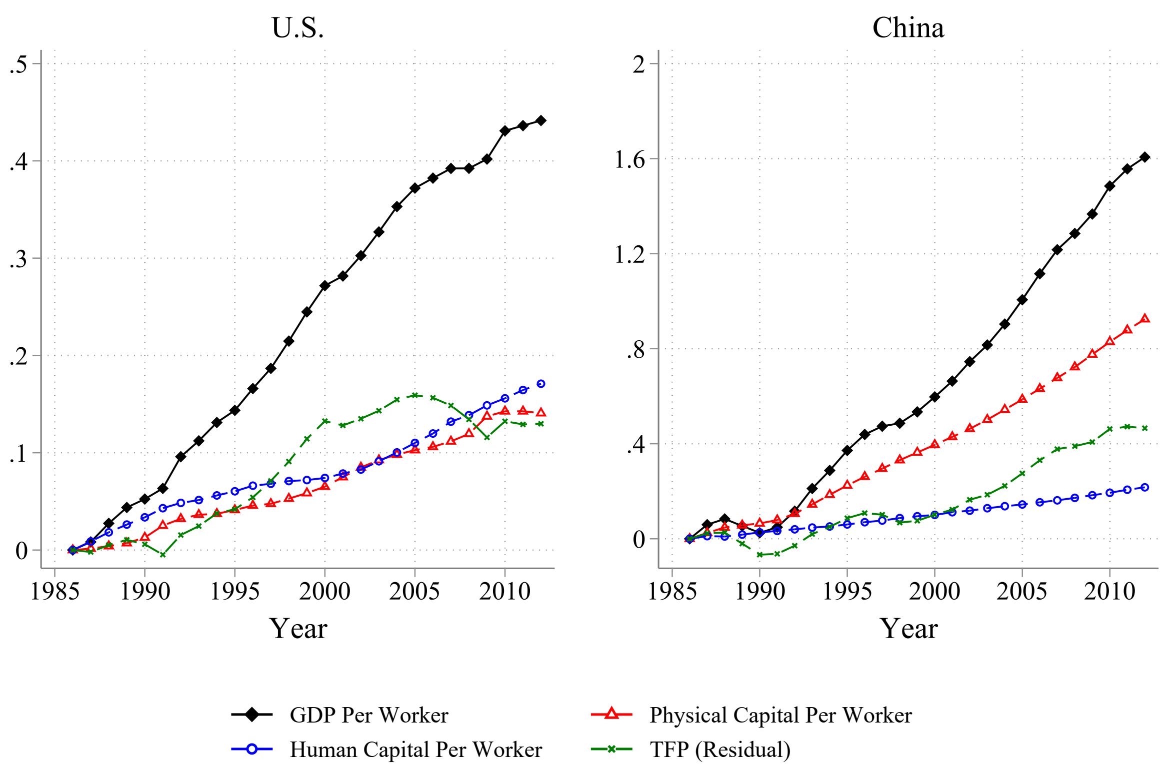
[25] For our estimated series from male earnings data to apply to the national growth accounting, one needs to assume that the human capital changes (not necessarily levels) are the same for males and females.

[26] Our method only delivers growth accounting and cannot perform levels accounting. This is because our decomposition delivers three series of changes relative to some base group, not the levels.

[27] Elsby et al. (2013) show that observed changes in the labor share barely affect the results of a growth accounting exercise.

[28] Zhu (2012) estimates the average annual total factor productivity growth in the nonagricultural sector to be 2.17% and 0.27% for the nonstate and state sectors during 1988-1998, but 3.67% and 5.50% for nonstate and state sectors

Figure 5: Growth Accounting



Notes: This figures decomposes the growth in GDP per worker (black diamond) into contributions of physical capital per worker (red triangle), human capital per worker (blue circle), and TFP residual (green cross). Note that the scales differ in the two figures.

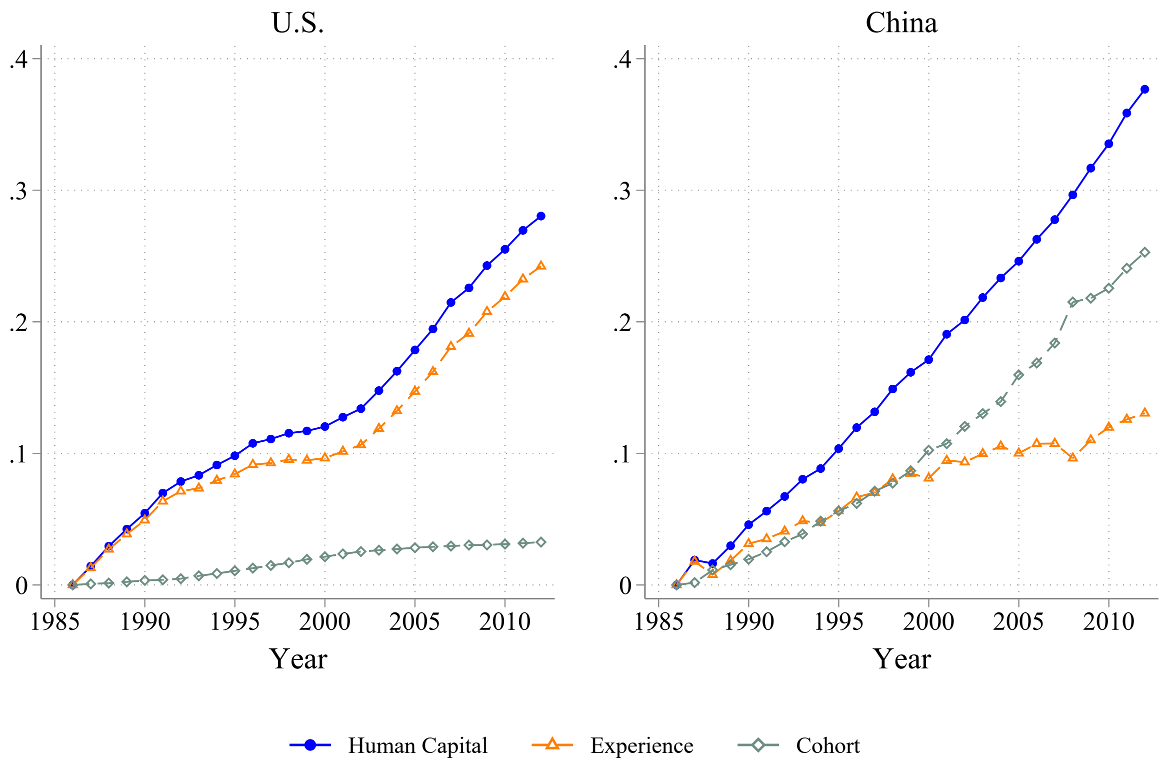
such as the privatization of the State-Owned Enterprises (SOE), the trade liberalization following the entry into the World Trade Organization (WTO), and the massive internal migration.[29]

There are several related attempts to adjust for human capital in development and growth accounting. Hall and Jones (1999) impute human capital as $h = \exp\{\phi(e)\}$, where e denotes years of schooling and $\phi'(e)$ in estimated as the return to schooling from a standard Mincerian log wage regression. Bils and Klenow (2000) extend the idea to include Mincerian return of experience as well. In addition, they introduce interdependence on the human capital of older cohorts to capture impacts from teachers. Such an approach based on Mincer regressions to measure human capital typically finds that cross-country differences in output per worker are largely driven by differences in TFP. A potential problem of those constructions based on educational attainment is that this approach implicitly assumes one additional year of schooling contains the

during 1998-2007.

[29]Chen et al. (2019) carefully addresses the selection issue in the privatization of SOEs and finds that privatization does lead to productivity gains. Brandt et al. (2017) provides evidence that trade liberalization — both input tariff cuts and output tariff cuts — raises firms' productivity. Tombe and Zhu (2019) quantify that the reduction in internal trade and migration costs accounts for 28% of China's growth.

Figure 6: Decomposition of Human Capital Growth into Experience and Cohort Effects



Notes: This figures decomposes the average human capital growth (blue circle) into contributions of the experience effect (orange triangle) and the cohort effect (gray diamond).

same quality of human capital across countries or over time. Manuelli and Seshadri (2014) take a different route. Instead of relying on the Mincer regression, they specify and calibrate a model of human capital acquisition with early childhood development, schooling, and on-the-job training, and then calculate human capital stocks from the calibrated model. They find a larger role for human capital and a smaller role for TFP in explaining the cross-country differences in output per worker. Our approach combines the merits of both approaches: it properly infers human capital while maintaining the simplicity and transparency of the procedure. The closest to our exercise is Bowlus and Robinson (2012), who use the HLT identifying restriction to tease out human capital price changes from human capital quantity changes. They are the first to apply the HLT in the context of growth accounting and find that adjusting the human capital input changes reduces the contribution of TFP to growth dramatically. Building on the insights of Bowlus and Robinson (2012), we take one step further by separating the role of experience accumulation and inter-cohort improvements in aggregate human capital growth, which we turn to now.

We calculate the role of experience to aggregate human capital by fixing the cohort effect at its base group level. Similarly, we calculate the role of cohort by fixing the experience effect at its base group level.[30] The left panel of Figure 6 shows that human capital per worker increases by almost 30% in the U.S. from 1986 to 2012, most of which is due to experience effects while little due to inter-cohort human capital improvements. This is perhaps not surprising given that the estimated cohort effect is small but the experience effect is very large in the U.S. Productivity gains from experience in an aging workforce would be large if the life-cycle human capital accumulation is fast. The right panel of Figure 6 shows that, in China, human capital per worker increases by almost 40% during the same period. Inter-cohort human capital improvement is a more important contributor. It accounts for two-thirds of the overall human capital growth while experience accumulation only accounts for the remaining one-third.

5.2 College Premium and Skill-Biased Technical Changes

5.2.1 Heterogeneous Human Capital by Education Groups

In the 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. It is straightforward to extend our framework to allow for different types of human capital. For example, college and high school graduates may possess different types of skills that are not perfect substitutes. To do so, we perform the decomposition as discussed in Section 4 separately for college workers and high school workers. 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 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 bins towards the end of working life. 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, effectively it is assumed that 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 overlapped with the "flat spot" proposed by Bowlus and Robinson (2012). After detailed investigation of the U.S. data, they conclude that a reasonable choice for the flat spot of the experience effect is around 50-59 for college graduates and 46-55 for high school graduates.

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 the previous literature that life-cycle wage growth tends to be faster for workers with more education (see Bagger et al., 2014, for example). The difference between the two education groups in their experience effect profiles, however, is much smaller compared to the difference in the

[30] The "experience" series in Figure 6 is calculated as $h_t^{\text{experience}} = \sum_c \sum_k \exp(r_k)\omega(c,k;t)$ and the "cohort" series as $h_t^{\text{cohort}} = \sum_c \sum_k \exp(s_c)\omega(c,k;t)$.

cohort effects that we are turning to.

Second, China and the 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 phenomenon in wage 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 the cohort-specific human capital that happened to 1980-1984 birth cohort college graduates. This is perhaps not surprising if one links to the institutional background this cohort experienced. The Chinese government expanded college enrollment at a large scale in 1999. In the following years, the expansion was unprecedentedly massive.[31] As a much large fraction of this cohort could enroll in college, thanks to the higher education expansion, the selectivity of college decreases significantly. Thus, the distribution of innate ability among this cohort of college students may shift downward compared to previous cohorts of college students. It is reassuring that our decomposition picks up this pattern.

Finally, the trends in time effects are broadly similar for both education groups. There appears to be a diverging trend of relative prices in the two countries. In China, the rental price to human capital increases rapidly for both education groups, and the rental price to college human capital increases even faster than that to high school human capital. In the U.S., the human capital price does not change much for either education group, but decreases slightly more for college workers.

5.2.2 Decomposing College Premium

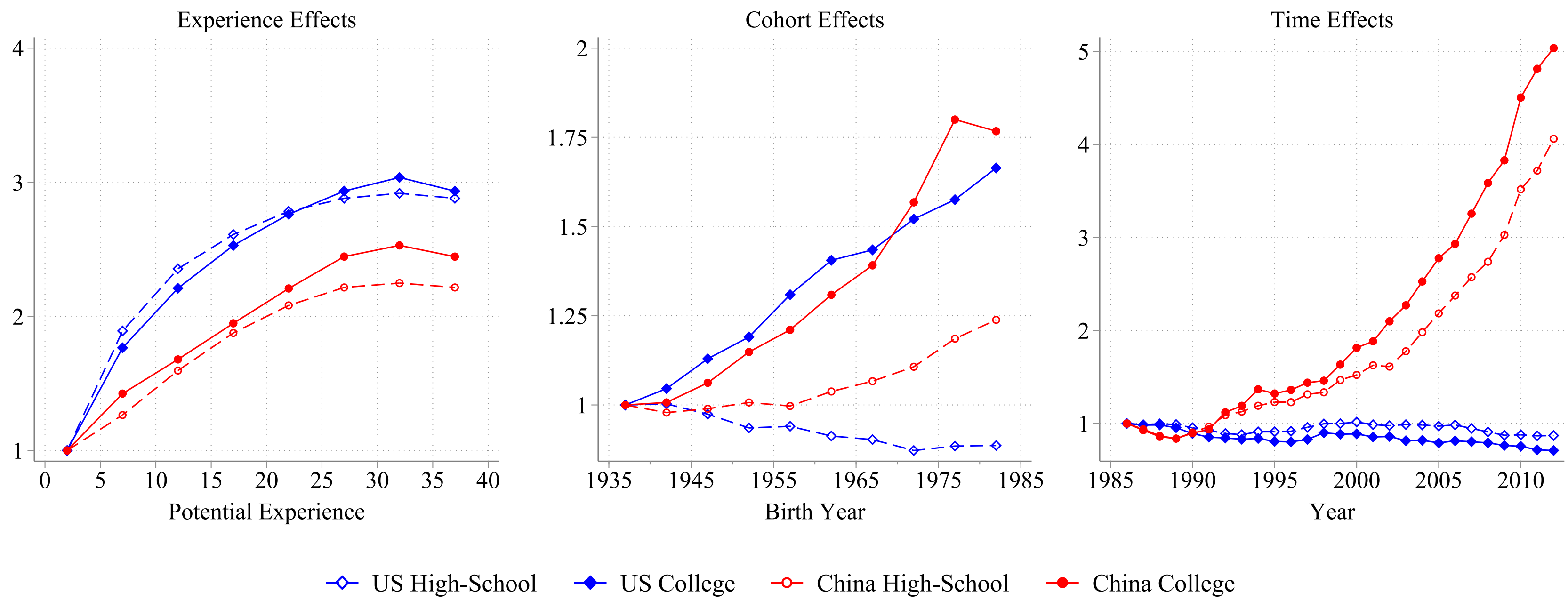
The wage ratio between college graduates and high school graduates is often interpreted as the relative price between college skills and high school skills. By this logic, evolution in the college wage premium is informative about changes in the relative skill prices. The implicit assumption is that the relative amount of human capital between education groups is unchanged. 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 the human capital of education group e at time t, and H_t^e is the average human capital for workers of education group e at time t. 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 $\xi_t = H_t^{\text{cl}} / H_t^{\text{hs}} \equiv \xi, \forall t$, can we interpret the changes in the college premium over time as reflecting entirely the changes in the relative price of college human capital and high school human capital. Under this

[31] 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 Workers



Notes: This figure shows the decomposition results of experience, cohort, and time effects in the U.S. (blue diamond) and China (red circle), separately for college workers (solid line and filled marker) and high school workers (dashed line and hollow marker).

implicit assumption, the seemingly puzzling 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. motivates the literature on the skill-biased technical changes (SBTC) (see Acemoglu and Autor, 2011; Violante, 2008, for excellent overviews). However, the so-called "skill-biased technical changes", which is essentially a residual object, may simply reflect changes in the relative human capital between education groups, and thus may not necessarily be related to technological changes. Our decomposition allows to estimate the changes in the relative human capital of college versus high school workers as well as the relative price of college versus high school skills. We construct relative human capital quantity series based on both experience and cohort effects, as we do in Section 5.1. We then decompose the evolution of average college premium into the relative changes in the price and quantity of human capital possessed by the two education groups.

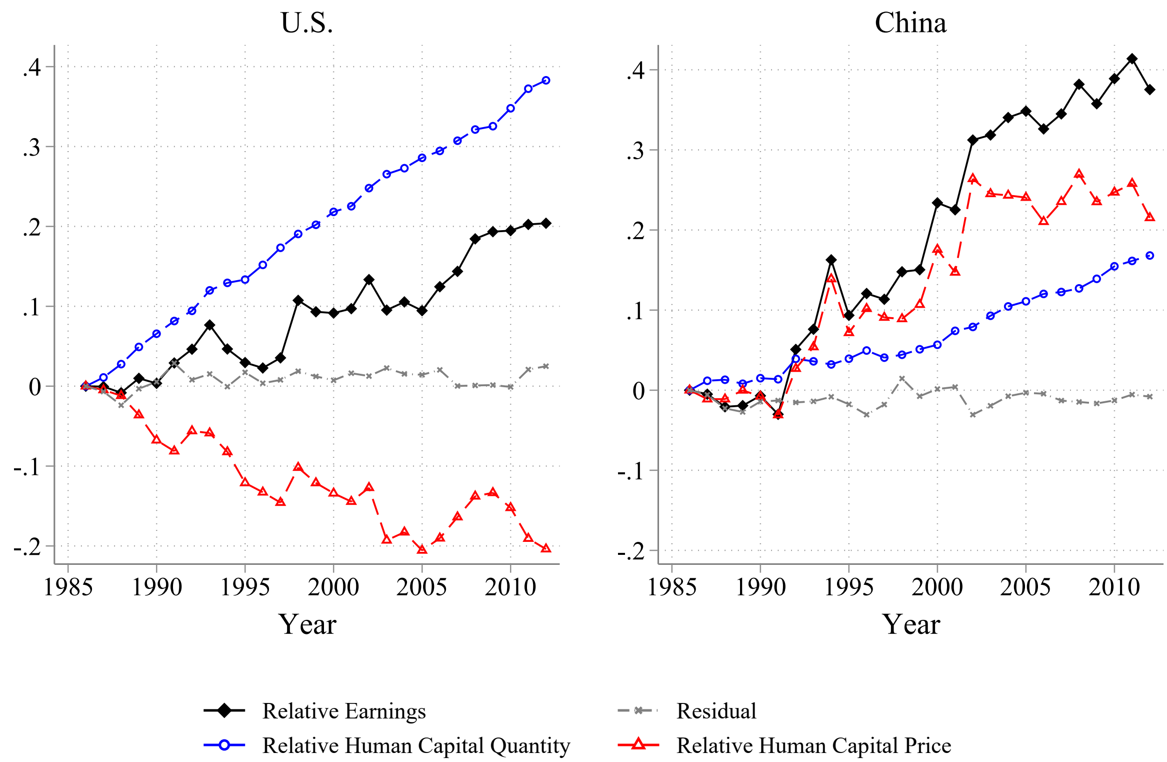
The results are plotted in Figure 8. The college premium is constructed based on the relative log earnings among prime age workers between 25 and 54 years old and we normalize the series to reflect changes relative to its 1986 level. As is shown in the left panel, in the U.S., the relative price between college human capital and high school human capital is actually declining. Rising college premium in the U.S. is mainly a result of the increase in the relative quantity of college human capital, i.e., an average college worker's human capital increases more than an average high school worker. In fact, the relative human capital quantity of an average college worker increases even more than offset the declining relative skill price so that the college premium still increases. The right panel of Figure 8 shows that in China, the increase in the college wage premium is driven by both the increase in the relative prices and the increase in relative quantity of college human capital over non-college human capital, but the relative price changes play a more important role. Note that in both figures, the residual term is tightly around zero, suggesting that the decomposition provides a good fit to the data.

5.2.3 Skill-Biased Technical Change

The finding in the previous subsection that most of the rise in the relative wage of college workers versus non-college workers in the U.S. is accounted for by the relative human capital quantity, rather than the relative skill price, is consistent with Bowlus and Robinson (2012). At first glance, this may seem to be a contradiction to skill-biased technical changes. Below, we take a step further to infer the extent of skill-biased technical changes in both countries. We find no contradiction between declining relative skill prices in the U.S. and the skill-biased technical changes. In fact, our decomposition results reveal large skill-biased technical changes in both countries, without which the relative price of college human capital would decline even more as the relative quantity of college human capital rose rapidly.

We revisit the magnitude of skill-biased technical change by 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

Figure 8: Decomposing Changes in College Premium



Notes: This figure decomposes changes in college premium (black diamond) into changes in relative human capital price (red triangle) and changes in relative human capital quantity (blue circle). The dashed gray line plots the residual of the decomposition.

high school human capital:

$$Y_t = \left[(A_t^s H_t^s)^{\frac{\sigma-1}{\sigma}} + (A_t^u H_t^u)^{\frac{\sigma-1}{\sigma}} \right]^{\frac{\sigma}{\sigma-1}}, \tag{9}$$

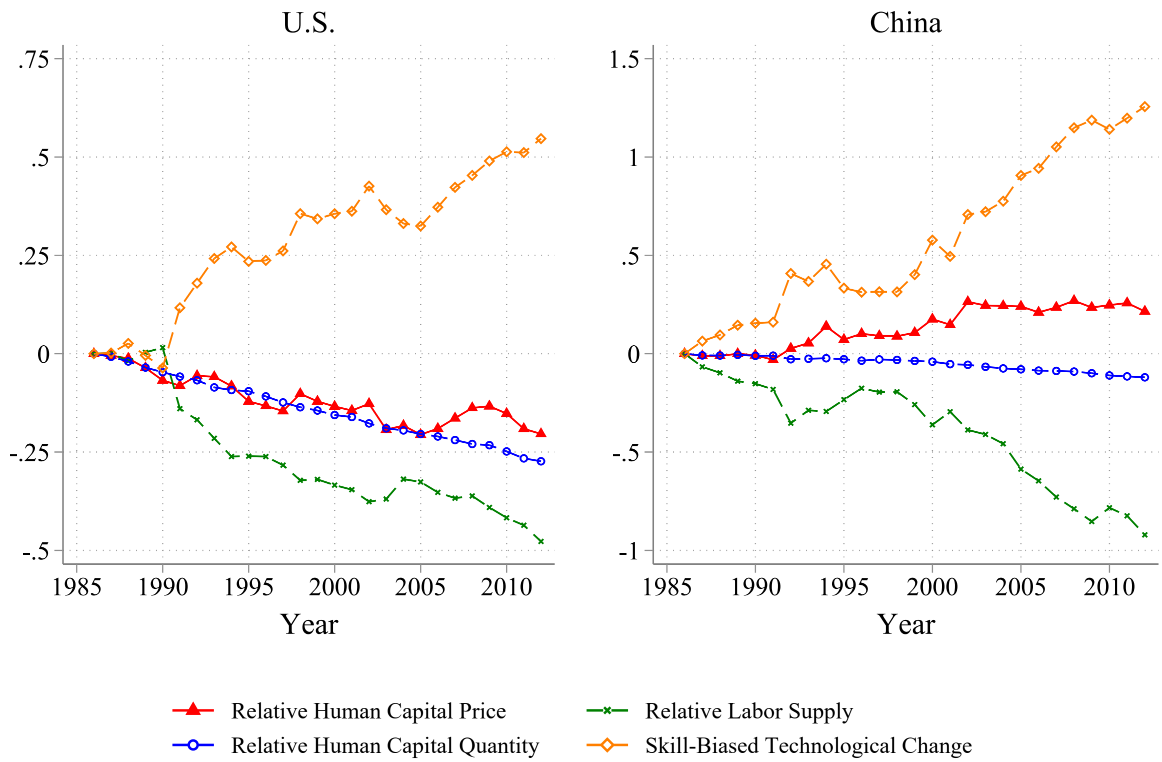
where H^s and H^u are the aggregate human capital quantity of the two types of skills, A^s and A^u are the respective skill augmenting technology, and $\sigma > 0$ is the elasticity of substitution between these two types of human capital.[32] Assume skills are paid by their marginal product. The relative price of the two types of skills is:

$$\ln\left(\frac{p^s}{p^u}\right) = \frac{\sigma-1}{\sigma}\ln\left(\frac{A^s}{A^u}\right) - \frac{1}{\sigma}\ln\left(\frac{h^s}{h^u}\right) - \frac{1}{\sigma}\ln\left(\frac{L^s}{L^u}\right), \tag{10}$$

where h^s and h^u are the efficiency units (human capital quantity) per worker of the two education groups, and L^s and L^u are the aggregate number of workers of the two education groups, such

[32]Since the focus here is on workers of different education, we abstract from capital in the production function. It is without loss of generality though, as the role of capital can be captured by A^s and A^u.

Figure 9: Decomposing Changes in Relative Human Capital Prices



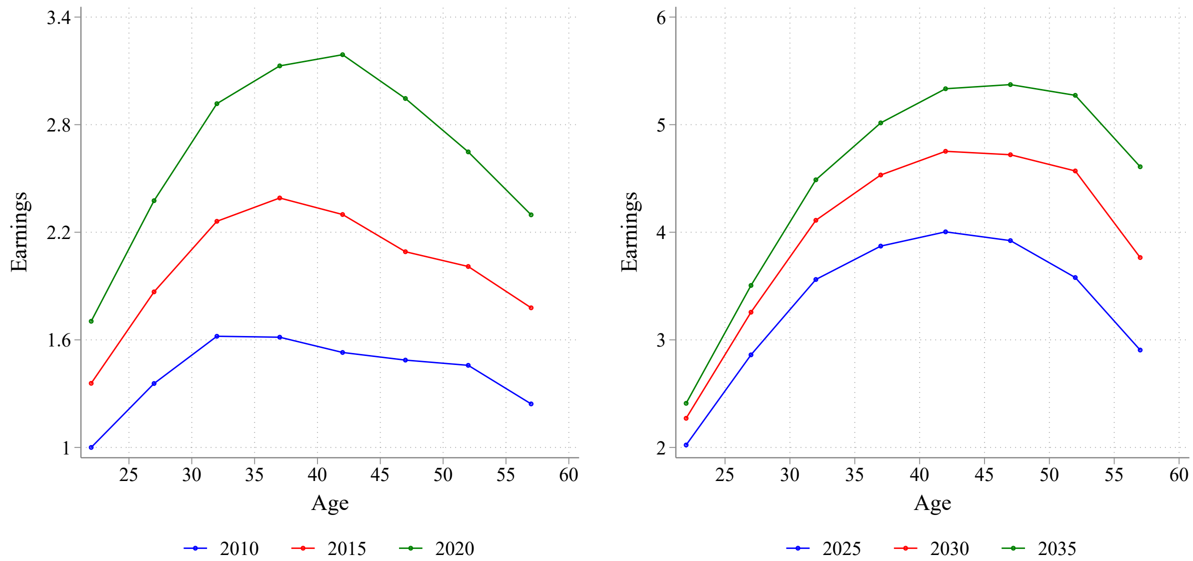
Notes: This figure decomposes changes in relative human capital prices into relative labor supply, relative human capital quantity per worker, and skill-biased technical change, under $\sigma = 1.4$ estimated by Katz and Murphy (1992).

that the aggregate supply is $H^s = h^s L^s$ and $H^u = h^u L^u$. The first term on the right-hand side captures the contribution of skill-biased technical changes to relative price changes. The second term reflects the role of 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) using 40 years of U.S. data, and obtain an estimated series for contributions of changes in A^s/A^u.[33] Since $\sigma > 1$, an increase in A^s/A^u (i.e., skill-biased technical 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 contributions of skill-biased technical changes can thus be obtained as a residual. We discuss the relation to previous estimates of skill-biased technical changes in Appendix C.3.

[33] Acemoglu and Autor (2011) conclude that most estimates in the literature agreed on a value of σ to be somewhere between 1.4 and 2. We report the decomposition results under $\sigma = 2$ in Figure A.5 in the Appendix. Although the exact numbers change a bit, the overall pattern is robust.

Figure 10: A Hypothetical Scenario for China's Earnings Profiles in 30 years



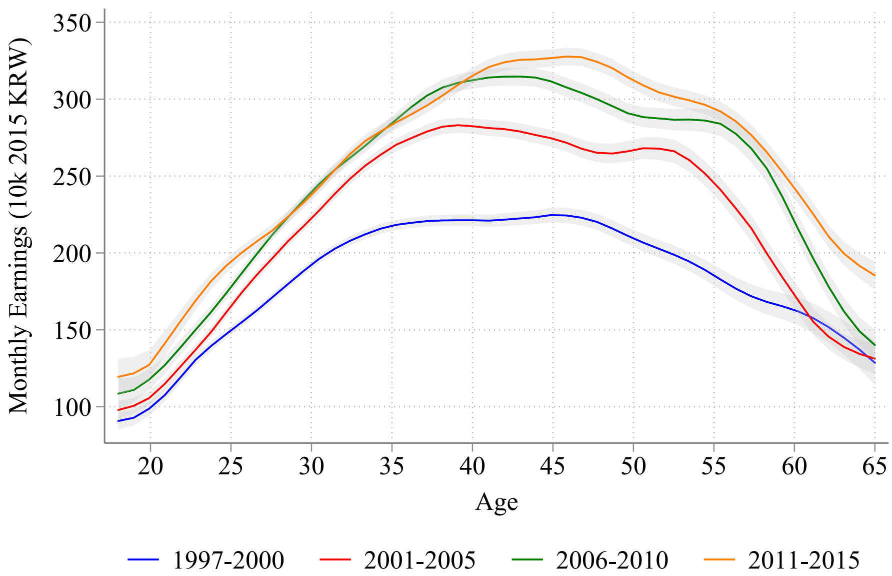
The contributions of relative labor supply, relative human capital per worker, and skill-biased technical change to the evolution of relative human capital prices is depicted in Figure 9. It clearly shows that in both U.S. and China, the relative quantity of college human capital grows rapidly, which would have led to sharp declines of the price of college human capital relative to non-college human capital. Due to skill-biased technical changes, the relative price of college human capital did not decline even more in the U.S. and actually increased in China in the last thirty years.

5.3 "New Normal" and the Golden Ages in China

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% ($= 1.87^{1/45} - 1$) 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 both close to 0 for US. However, the spectacular growth in China in the last forty years is not expected to last forever; in fact, since 2010, the growth rate in China has slowed down significantly, and many analysts expect the "new normal" growth rate in China to converging to rates similar to those in the U.S. (Barro, 2016). In this section, we perform a simple 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 Figure 10, we show that under this scenario, the vertical gaps between two consecutive

Figure 11: Cross-Sectional Age-Earnings Profiles of Korean Male Workers



cross-sectional age-earnings profiles will be shrinking, showing the slowdown in the time effects. Notably, the "golden age," which was around 30-35 in 2010, 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 is a result of the slowdown in the cohort effects (i.e., the inter-cohort human capital growth rate). If the Chinese economy indeed slows down and converges to the "new normal" growth rates similar to more mature developed economy such as the U.S. in the next thirty years, our simulation suggests that the cross-sectional age-earning profiles over time will exhibit older "golden ages", and reverse the pattern of ever-lowering "golden ages" in the the next thirty years.

Is this a realistic prediction? Only history will tell for sure, but interestingly, Figure 11 shows that such a pattern of increasing "golden ages" actually happened in Korea during the past ten years, using data from the Korean Labor and Income Panel Study (KLIPS). Korea experienced its fastest growth during the 1960s to 1990s. After that, it began to slowdown. Appendix Figure A.6 depicts the decomposition for Korea, together with the 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. 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 10.

6 Conclusion

In this paper, we first document stark differences in the cross-sectional age-earnings profiles between the U.S. and China, the two largest economies in the world, during the past thirty years. We find that, first, the peak age in cross-sectional age-earnings profiles, which we refer to as the "golden age," stays almost constant at around 45-50 years old in the U.S., but decreases sharply from 55 to around 35 years old in China; second, the age-specific earnings grow drastically in China, but stay almost stagnant in the U.S.; and third, the cross-sectional and life-cycle age-earnings profiles look remarkably similar in the U.S., but differ substantially in China.

To explain these striking differences, we propose and empirically implement a unified decomposition framework to the infer from the repeated cross-sectional earnings data the life-cycle human capital accumulation (the experience effect), the inter-cohort productivity growth (the cohort effect), and the human capital price changes over time (the time effect), under an identifying assumption that the growth of the experience effect stops at the end of the working career. The decomposition suggests that China has experienced a much larger inter-cohort productivity growth 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 also use the inferred components to revisit several important and classical applications in macroeconomics and labor economics, including the growth accounting and the estimation of the TFP growth, and the college wage premium and the skill-biased technical change. We find that once we adjust for the changes in the quantities of human capital, the estimated contribution of the TFP to GDP per capita growth is smaller than the previous estimates in the literature. We also find that the skill-biased technical change played an important role in the rising college premium to ensure that the relative price of college human capital does not drop as much as it would otherwise do when there is a large increase in the quantity of college human capital. A simple simulation exercise using our framework also suggests that, as the Chinese economy slows down to a "new normal" growth rate similar to that in the U.S., the golden ages of the cross-sectional age-earnings profile in China will start to increase to older ages, similar to what has happened in Korea in the last ten years.

The mostly descriptive findings in this paper suggest many potential directions for future research. First, in our analysis, we assume that workers are either perfect substitutes, or perfect substitutes within an education group. This rules out the possibility that there might be new vintages of physical capital that can only be combined with the human capital of newer cohorts; that is, the human capital of different cohorts are not substitutable. To distinguish technological changes that favor younger generations from the inter-cohort human capital growth would require better-suited data or richer model structures, but it is an exciting area for future research. Second, it is also important to link the decomposition results to specific institutions and reforms, and evaluate how much each policy contribute to the evolution of these components. For example, it is reasonable to hypothesize that the re-opening of the schools at the end of the Cultural Revolution and the 1999 college expansion in China are the most related to the inter-cohort productivity growth documented for China; 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. Third, throughout the paper we focused on urban males. A natural question is how structural change from agriculture to industry, increasing female labor participation, and internal migration (especially in China) are reflected in the decomposition. Fourth, understanding why the returns of experience is higher in developed economies than in less developed economies, which was documented in Lagakos et al. (2018) and confirmed in our study, is an interesting area for further explorations. Fifth, it is also important to examine the implications of the rapid inter-cohort productivity growth and human capital rental price in China on other programs such as the social security system.[34]

Finally, in this paper we focused on U.S. and China because they are the two largest economies in the world, and the labor market dynamics in these two countries are likely to play an out-sized influence on the global economy, but the decomposition framework in this paper can be fruitfully applied in other countries.

[34] For exampe, see Fang and Zhang (2021) for an exploratory study on the relationship between inter-cohort productivity growth and pension reform, particularly the delay of retirement age, in China.

References

Acemoglu, Daron and David Autor, "Skills, Tasks and Technologies: Implications for Employment and Earnings," in "Handbook of labor economics," Vol. 4, Elsevier, 2011, pp. 1043–1171.

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

Angrist, Joshua D and Jörn-Steffen Pischke, *Mostly Harmless Econometrics: An Empiricist's Companion*, Princeton University Press, 2009.

Bagger, Jesper, Fran Fontaine, Fabien Postel-Vinay, and Jean-Marc Robin, "Tenure, Experience, Human Capital, and Wages: A Tractable Equilibrium Search Model of Wage Dynamics," *American Economic Review*, 2014, *104* (6), 1551–96.

Barro, Robert J, "Economic Growth and Convergence, Applied Especially to China," *Working Paper 21872*, 2016.

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

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

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

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

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

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

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

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

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

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

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

Elsby, Michael WL, Bart Hobijn, and Ayşegül Şahin, "The Decline of the U.S. Labor Share," *Brookings Papers on Economic Activity*, 2013, *2013* (2), 1–63.

Fang, Hanming and Yi Zhang, ""Growing Pains" in the Chinese Social Security System," *Working Paper*, 2021.

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

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

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

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

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

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

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

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

Heckman, James J, Lance J Lochner, and Petra E Todd, "Earnings Functions, Rates of Return and Treatment Effects: The Mincer Equation and Beyond," *Handbook of the Economics of Education*, 2006, *1*, 307–458.

_ **, Lance Lochner, and Christopher Taber**, "Explaining Rising Wage Inequality: Explorations with a Dynamic General Equilibrium Model of Labor Earnings with Heterogeneous Agents," *Review of Economic Dynamics*, jan 1998, *1* (1), 1–58.

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

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

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

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

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

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

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

Lise, Jeremy and Fabien Postel-Vinay, "Multidimensional Skills, Sorting, and Human Capital Accumulation," *American Economic Review*, 2020, *110* (8), 2328–76.

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

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

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

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

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

Song, Zheng Michael and Dennis Tao Yang, "Life Cycle Earnings and Saving in a Fast-Growing Economy," *Unpublished paper, Chinese University of Hong Kong*, 2010, *3.*

Tombe, Trevor and Xiaodong Zhu, "Trade, Migration and Productivity: A Quantitative Analysis of China," *American Economic Review*, 2019, *109* (5), 1843–72.

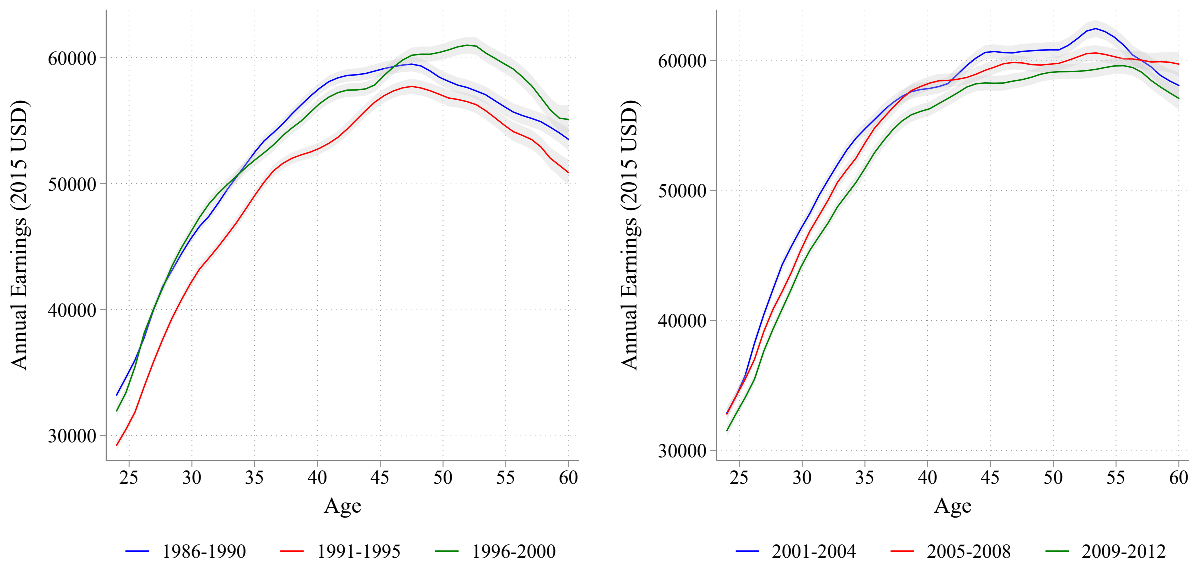
Violante, Giovanni L, "Skill-Biased Technical Change," *The New Palgrave Dictionary of Economics*, 2008, *2.*

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

Online Appendix

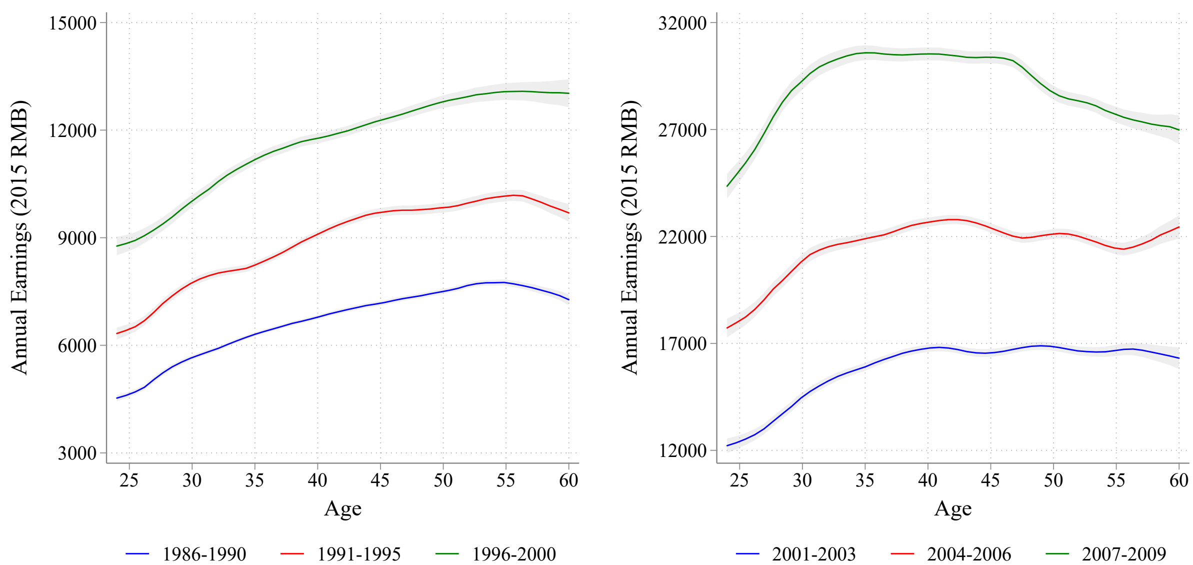
A Additional Figures

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



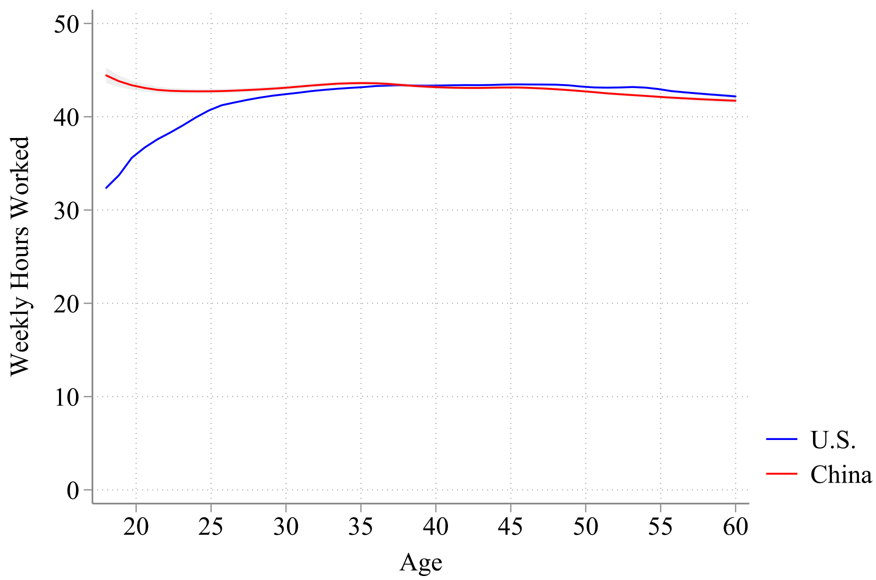
Notes: This figure plots the cross-sectional age-earnings profiles of U.S. male workers that live in metropolitan areas, using March CPS from 1986 to 2012. Each curve represents a cross section that pools adjacent years. The solid lines are kernel smoothed values and the gray shaded areas are the 95%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

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



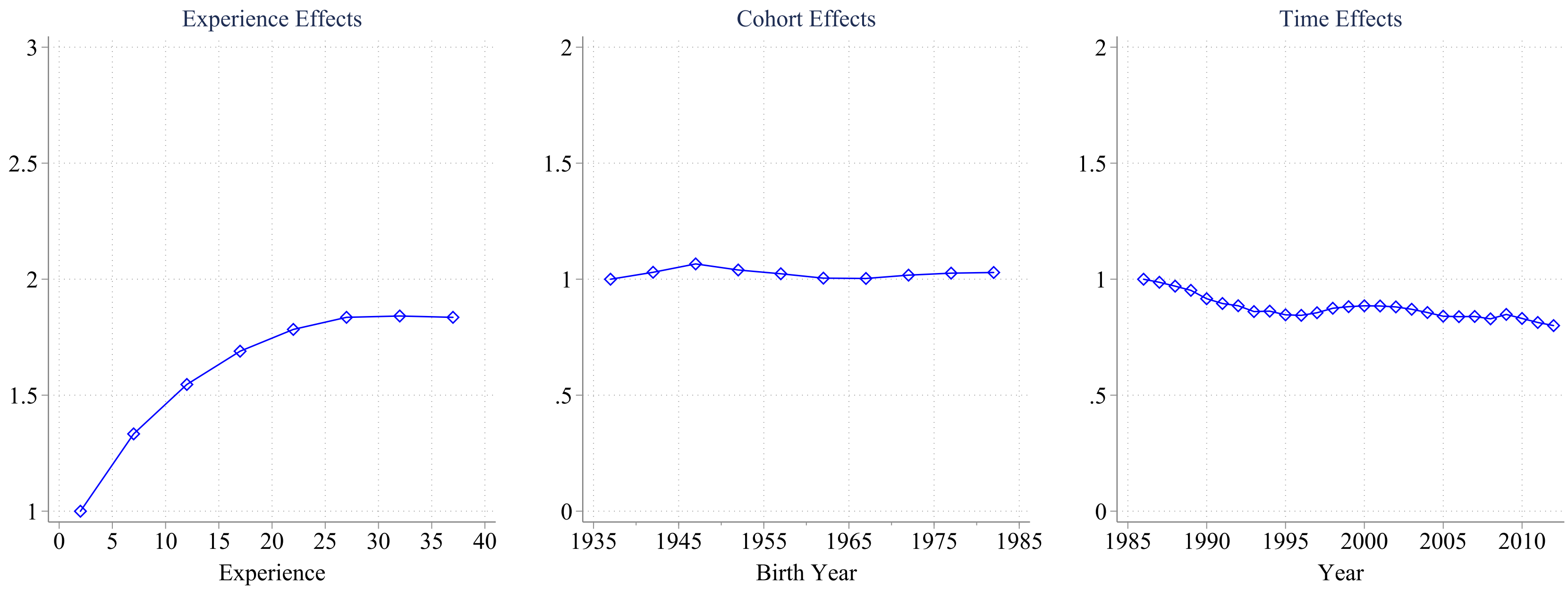
Notes: This figure plots the cross-sectional age-earnings profiles of Chinese Urban male workers in Beijing, Shanxi, Liaoning, Heilongjiang, Shanghai, Jiangsu, Anhui, Jiangxi, Shandong, Henan, Hubei, Guangdong, Sichuan, Yunnan, Gansu, covering 1986-1991 in the left panel and 2002-2009 in the right panel. 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. Also note that the time coverage is shorter than the baseline result—we only have data till 2009, instead of 2012, for these 15 provinces.

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



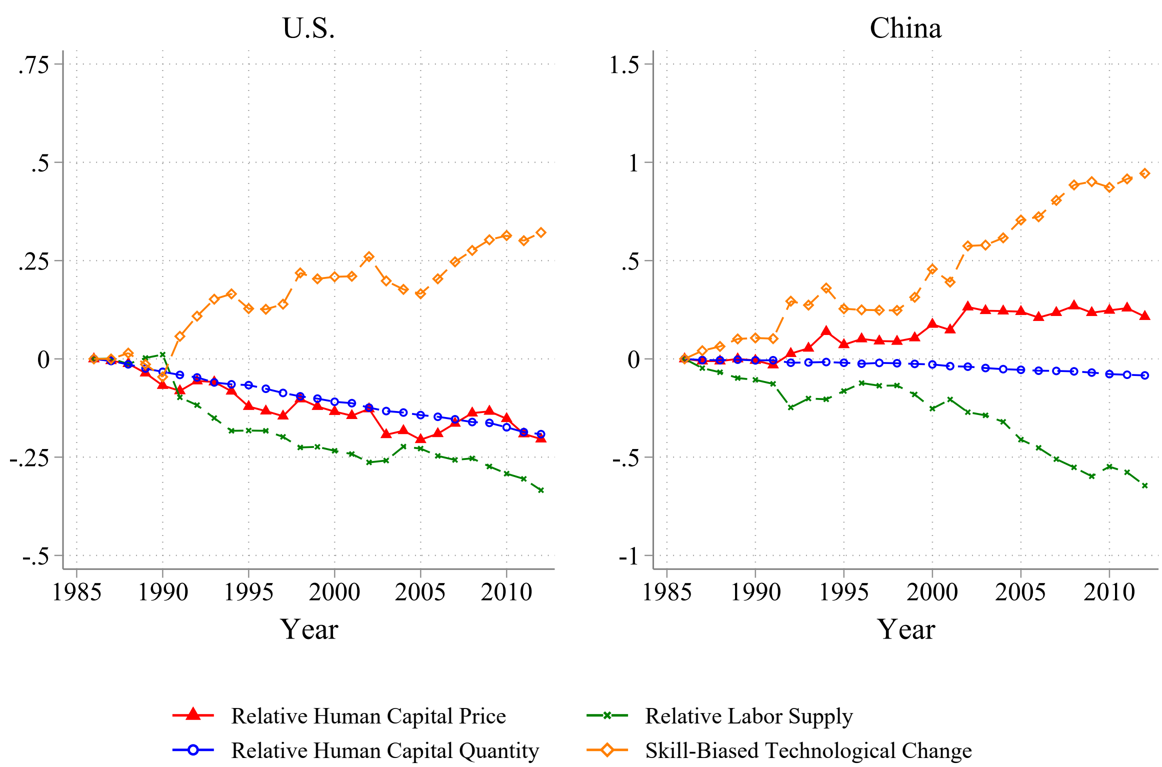
Notes: This figure plots the cross-sectional age-hours profiles of U.S. and Chinese male workers in 2002-2006. Hours worked per week is measured by the "total number of hours usually worked per week over all jobs the year prior to the survey" from CPS (for U.S.) and imputed as "total number of hours worked last month" divided by 4.3 from UHS (for China).

Figure A.4: Decomposition Using Hourly Wage for Full-Time Workers



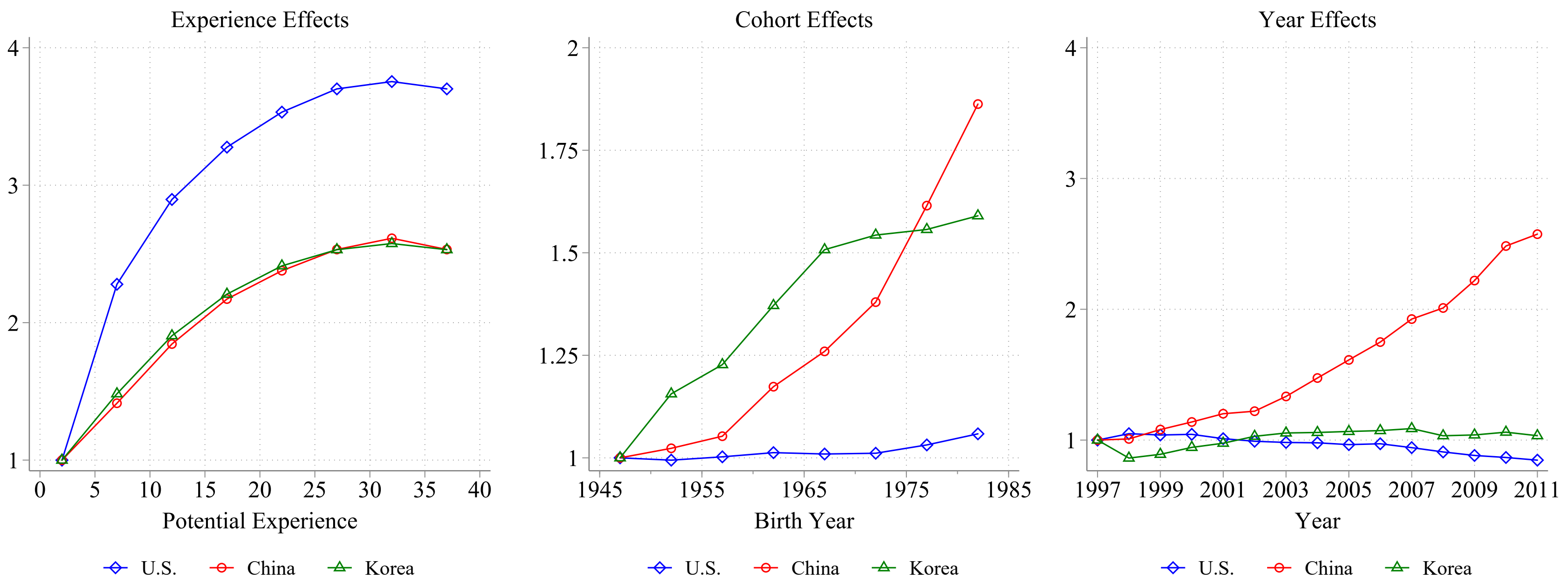
Notes: This figure shows the decomposition results of experience, cohort, and time effects in the U.S. based on hourly wage for full-time workers.

Figure A.5: 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 technical change, under $\sigma = 2$, the upper bound for σ estimated in the literature, according to Acemoglu and Autor (2011).

Figure A.6: 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.

B Additional Tables

Table B.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.

C Additional Discussion

C.1 Age-Cohort-Time Identification

C.1.1 McKenzie (2006)

Consider the following statistical model of age/experience, cohort, and time. Suppose the variable of interest is a linearly additive model of cohort (c), experience (k) and time (t) effects with

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

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

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

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:

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

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

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

Thus the change in the slope of the age-effect profile (i.e., a second difference) 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 this point, 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, by further normalizing one level each of two effects, one can recover all levels. But in this paper, what we care about are the slopes, i.e., 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).

C.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

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

where the level of the base group is load on to the constant term. 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 the collinearity across time, cohort, and age $t = c + k$:

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

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. Consider another set of parameter vectors defined by

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

which still satisfies the equation of interest. 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 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 $t = 3, .., T$,

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

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. LMPQS) investigate both, or a mixture of the two 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.

C.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

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

where $\beta\left(k, \theta\right)$ is derived from an economic model and θ is a vector of model fundamentals. To achieve identification, this approach requires the function $\beta\left(k, \theta\right)$ 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 or higher derivatives of the age effects. This approach ultimately facilitates identification of structural parameters associated with age effects without first identifying the age effects, by imposing parametric forms in the economic model.

C.1.4 Heckman et al. (1998)

Heckman et al. (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, in some sense, a perfect combination of the previous two approaches. On 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 from a structural model of human capital investment.

C.2 Algorithm

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). In the baseline specification, we impute potential experience as min ageedu6, age18 and consider a maximum of 40 years of experience. We group cohorts and experience into five-year bins. Assume there is no additional experience effect in the last two experience bins. The goal is to estimate

$$w_{i,t} = \text{constant} + s_c + r_k + p_t + \varepsilon_{i,t}$$

subject to the identifying restriction $r_{25\sim 29} = r_{35\sim 39}$.

Transform the above equation to

$$w_{i,t} = \text{constant} + s_c + r_k + gt + \tilde{p}_t + \varepsilon_{i,t},$$

where $\tilde{p}_t$ reflects fluctuations orthogonal to a linear trend such that $\sum_t \tilde{p}_t = 0$ and $\sum_t t\tilde{p}_t = 0$. That is, we rewrite an arbitrary time series p_t as a linear trend gt plus fluctuations $\tilde{p}_t$. The benefit of such algebraic manipulation is that once a value of g is obtained, we can run Deaton's procedure as explained in Section C.1.2 on the deflated wage $\tilde{w}_{i,t} := w_{i,t} - gt$ and get estimates for cohort, experience, time effects under this particular g. Then the problem boils down to pin down the value of g. The time trend is 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.

The algorithm can be summarized by an iterative procedure.

1. Start with a guess for the growth rate g^0 of the 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.

2. Suppose we are now at the m-th iteration. Deflate the wage data using the current guess of growth rate, g^m:
$$\tilde{w}_{i,t} := w_{i,t} - g^m t.$$

3. Rewrite the problem as a Deaton (1997) problem:
$$\tilde{w}_{i,t} = \text{ constant } + s_c + r_k + \tilde{p}_t + \varepsilon_{i,t}.$$

   Follow Deaton's procedure laid out in Section C.1.2 but use log deflated wage as the dependent variable. Regress log deflated wage on a set of dummies for experience groups, cohort groups and d^*'s as defined in the previous subsection C.1.2.

4. Check if the estimated experience effects are sufficiently close between experience group 25 (or experience group 30) and experience group 35, according to a preset precision.

5. If the convergence condition is satisfied, then we're done. Otherwise, we update the guess for the growth rate by annualized experience effect r^m_{end} in the specified flat region in the current iteration with a damping factor δ:

$$g^{m+1} = g^m + \delta r^m_{\text{end}},$$

and go back to step 2 with the updated guess g^{m+1}.

C.3 Skill-Biased Technical 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 all variables are allowed to change over time. Taking logs to the ratio of p^s and p^u gives Equation (10).

In the skill-biased technical change literature, it is often assumed that

$$Y_t = \left[(B^s_t L^s_t)^{\frac{\sigma-1}{\sigma}} + (B^u_t L^u_t)^{\frac{\sigma-1}{\sigma}} \right]^{\frac{\sigma}{\sigma-1}}, \tag{C.1}$$

where L^s (L^u) is the labor supply of college (high-school) 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, further decomposes it into two components. Our formulation (9) distinguishes relative improvements in human capital (i.e., h^s/h^u) from the technology that improves the productivity of the two types 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 (9) 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

$$w^s = \frac{\partial Y}{\partial L^s} = \left[(A^u h^u L^u)^{\frac{\sigma-1}{\sigma}} + (A^s h^s L^s)^{\frac{\sigma-1}{\sigma}} \right]^{\frac{\sigma-1}{\sigma}-1} (A^s h^s)^{\frac{\sigma-1}{\sigma}} (L^s)^{\frac{\sigma-1}{\sigma}-1},$$

$$w^u = \frac{\partial Y}{\partial L^u} = \left[\left(A^u h^u L^u\right)^{\frac{\sigma-1}{\sigma}} + \left(A^s h^s L^s\right)^{\frac{\sigma-1}{\sigma}} \right]^{\frac{\sigma-1}{\sigma}-1} \left(A^u h^u\right)^{\frac{\sigma-1}{\sigma}} \left(L^u\right)^{\frac{\sigma-1}{\sigma}-1}.$$

Therefore the college 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 changes,

$$\mathrm{d}\ln\left(\frac{w^s}{w^u}\right) = \frac{\sigma-1}{\sigma}\,\mathrm{d}\ln\left(\frac{A^s}{A^u}\right) + \frac{\sigma-1}{\sigma}\,\mathrm{d}\ln\left(\frac{h^s}{h^u}\right) - \frac{1}{\sigma}\,\mathrm{d}\ln\left(\frac{L^s}{L^u}\right).$$

Note that with the typical formulation (C.1) such as in Katz and Murphy (1992), we will have

$$\mathrm{d}\ln\left(\frac{w^s}{w^u}\right) = \frac{\sigma-1}{\sigma}\,\mathrm{d}\ln\left(\frac{B^s}{B^u}\right) - \frac{1}{\sigma}\,\mathrm{d}\ln\left(\frac{L^s}{L^u}\right).$$

In other words, B^s/B^u in the standard model is equivalent to $\left(A^s h^s\right) / \left(A^u h^u\right)$ in our formulation, and it is essentially a combination of the skill-biased technical change and the changes in relative human capital per worker between the two skill groups.

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.