Human Capital, Industrial Dynamics and Skill Premium*

Yong Wang† Xuemin Tang‡
Institute of New Structural Economics, Peking University

October 2021

Abstract

Motivated by several stylized facts about skill premium and industrial dynamics, we develop an endogenous growth model with infinite industries that are heterogeneous in both capital intensities and skill requirement. We analytically characterize the optimal investment in both physical capital and human capital, life-cycle dynamics of each of the infinite industries and the dynamics of skill premium. We highlight that (1) optimal human capital investment should be stage-dependent and synchronized with physical capital investment, and should match the skill demand from the underlying industries; (2) the aggregate skill premium and its dynamics are determined by the underlying industrial structures and their dynamics at the disaggregated level. Our model provides a structural microfoundation for the standard structureless macroeconomic frameworks for the analysis of skill premium and human capital investment in the pertinent macro literature.

Key Words: Human Capital; Industrial Dynamics, Skill Premium, Structural Change

*This paper has been presented at seminars or conferences at Renmin University, Peking University, International Conference on Trade, Financial Integration and Macroeconomic Dynamics & IEFS Japan 2019 Annual Meeting (Kobe University), Macroeconomics Workshop: 2018 UIBE, Shanghai Macro Workshop (SUFE 2019), CCER Summer Institute 2018, Shanghai Jiaotong University, Xi'an Jiaotong University, Tianjin University, Wuhan University, International Conference on Income Distribution and Economic Development (SWUFE, 2018), International Conference on Macroeconomic Development (Dalian, 2019), and China and the World Economy Annual Conference 2019. We are particularly grateful to Joseph Kaboski, Justin Yifu Lin, Ernest Liu, Erzo Luttmer, Rachel Ngai, Gene Grossman, José-Víctor Ríos-Rull, Ping Wang, Zhu Wang, Wen Yao, Daniel Xu, Akos Valentiyi and Xiaodong Zhu for their valuable suggestions and comments. All potential deficiencies are our own. The financial support from the China National Social Science Foundation (grant no. 20AJL017) is gratefully acknowledged.

†Contact author. Institute of New Structural Economics, Peking University. Email Address: yongwang@nsd.pku.edu.cn

‡Institute of New Structural Economics, Peking University. Email address: xuemintang@nsd.pku.edu.cn

1. Introduction

Industries are heterogeneous in their demand for human capital and physical capital. Underinvestment in human capital would result in low labor productivity and slow economic growth, whereas overinvestment in human capital might lead to brain drain or even social instability as no local jobs are available for "over-educated" people. Mismatch between skill supply and skill demand caused by inappropriate human capital investment would be further complicated by physical capital investment, not only because the two kinds of investment compete for limited resources but also because workers of different skill levels exhibit different substitutability with phyical capital. In addition, investment in both tangible and intangible capital feeds back on the evolution of the composition of underlying industries, measured by value added shares or employment shares of different industries in an economy. Therefore, it is important to understand how investment in human capital and physical capital should be balanced in the context of industrial dynamics as an economy undergoes structural change. Here, industrial dynamics refers to the time variation in output and employment of a given industry: when and how an industry enters a market, develops and perhaps exits the market eventually (see Klepper, 1996, 1997; Hopenhayn 1991; Restuccia and Rogerson, 2008). Structural change refers to the change in compositions of industries, especially at the disaggregated level such as the 6-digit NAICS level within manufacturing. Surprisingly, however, the answer to this question still remains unclear despite the enormous progress in the pertinent growth literature, because the existing literature on the interaction between human capital and physical capital mostly adopts a one-sector structureless framework (see, for example, Lucas, 1988, Mankiw, Romer and Weil, 1992, Barro and Sala-i-Martin, 1996, *etc.*). A case in point is the macroeconomic study on skill premium. There are two main approaches in the literature. One is to emphasize complementarity between physical capital and skilled labor and substitutability between physical capital and unskilled labor at the aggregate level: physical capital accumulation raises the marginal productivity of skilled labor but reduces the demand for unskilled labor, so skill premium widens with capital accumulation (see, for example, Stokey, 1996; Krusell *et al.*, 2000). The other approach is to highlight the role of skill-biased technological progress at the aggregate level: marginal productivity of skilled labor increases faster than unskilled labor because the rate of technological progress is higher in productions utilizing skilled labor, so skill premium keeps widening with the biased technical change (see, for example, Acemoglu, 2003).

Whereas both approaches provide valuable insights and useful quantitative frameworks, there are at least two important limitations. First, both approaches assume that the mathematical form of the aggregate production function is exogenous and time-invariant, but the choice of functional forms may affect quantitative results significantly. Moreover, the functional form of the aggregate production function could be endogenously changing as the composition of underlying industries may change over the course of economic development (Ju, Lin and Wang, 2015). Second, both approaches attribute changes in skill premium entirely to the *quantitative* changes in aggregate variables, and hence remain agnostic about whether and how skill premium (and its dynamics) at the aggregate level may be affected by potential *structural* changes. Recall that in growth models with endogenously different compositions of goods or industries (see Stokey, 1989; Romer, 1990), human capital intensities are typically assumed identical across goods/industries, which is counterfactual.[1] Such models are neither

[1]Notable exceptions include Buera, Kaboski, Rogerson and Vizcaino (2021), where skill intensities are

designed to nor able to distinguish, quantitatively or qualitatively, separate roles of a wide array of industry-specific shocks or frictions in factor relocation across industries, because all these micro-level structural differences are lumped together and treated as quantitative differences in the aggregate skill-biased technology or some other aggregate variable.

In light of these dissatisfying features, we propose an alternative growth framework which allows us to jointly explore skill premium, industrial dynamics, and optimal investment on both human capital and physical capital simultaneously. Our framework deviates from the standard assumptions that the aggregate production function has an exogenous and time-invariant functional form and that the economy consists of only one sector or mutiple symmetric sectors. Instead, industries are heterogeneous in physical capital to skilled labor ratios in our model, and the aggregate production function is endogenously derived from the industrial structures, which in turn are endogenously determined by the endowment structure, namely, the proportions of different production factors. We focus our theoretical analysis mainly on two tasks. One is to explicitly characterize how optimal human capital investment and physical capital investment should be made when industrial structures may change endogenously over time. The other task is to illustrate how the level and dynamics of skill premium at the aggregate level is actually determined by the industrial structures and industrial dynamics at the disaggregated level.

Our paper is modelwise closest to Ju, Lin and Wang (2015) (or JLW model thereafter), but there are three key differences. (1) There are two production factors in the JLW model: homogeneous labor and physical capital, whereas our model has three production factors: skilled labor, unskilled labor, and physical capital. (2) There is no human capital investment in the JLW model, whereas human capital investment is explicitly introduced into our model, which enables us to examine skill premium and human capital investment in the context of industrial dynamics. In fact, the JLW model is a special case of the current model when all labor is skilled labor. (3) The JLW model fails to explain the empirical pattern of shakeout, namely, the expansion period takes longer than the decline period in the life cycle of an industry, which is documented in details and explored in the literature (see, for example, Jovanovic and MacDald, 1994, Bertomeu, 2009, Klepper, 1996, 1997), but the current model is able to explain this fact because the marginal product of skilled labor increases faster than that of unskilled labor due to the capital-skill complementarity. Like the JLW model, the driving force of structural change in this paper is changes in factor endowment, which is different from the other mechanisms highlighted in the structural change literature such as non-homothetic preferences (see Kongasmut, Rebelo and Xie, 2001; Boppart; 2014; Comin, Lashkari, and Mestieri. 2021), unbalanced productivity growth (Ngai and Pissarides, 2007; Acemoglu and Guerrieri, 2008; Buera, Kaboski, Rogerson and Vizcaino, 2021) or international trade (Matsuyama, 2009; Uy, Yi and Zhang, 2013). Moreover, the driving force of life-cycle industry dynamics in our model is physical capital and human capital investment rather than technological innovation, so the mechanism is also different from the standard literature on industry dynamics (see Jovanovic and MacDald, 1994, Bertomeu, 2009, Klepper, 1996, 1997, Wang, 2008). The new mechanism proposed in this paper is more appealing to

---

asymmetric across sectors. However, they do not investigate determinants or dynamics of skill premium, nor do they study endogenous human capital investment decisions or industrial dynamics. Mulligan and Sala-i-Martin (1993) explore transitional dynamics in a two-sector endogenous growth model, but the two sectors refer to physical capital sector and human capital sector instead of different industries that producing different consumption goods. They do not investigate skill premium or life-cycle industry dynamics.

developing countries, where the relevant choice is how to adopt the appropriate technologies from an exisitng menu rather than invent new technologies (Basu and Weil, 2003; Lin and Wang, 2019).

The rest of the paper is structured as follows. In section 2, we document six stylized facts about skill premium, factor endowment, industrial heterogeneity, and industrial dynamics by using both the US and cross-country disaggregate industry level data. Motivated by these facts, we propose a multi-factor and multi-sector endogenous growth model to explain these facts simultaneously. In our model, industries are asymmetric in capital intensity and skill requirement. The composition of industries are endogenously determined by the factor endowment structure, namely the composition of physical capital, skilled labor and unskilled labor. We first show in a static model how a given structure of factor endowment determines the optimal industrial structure, the aggregate production function, and all the factor prices including skill premium (Section 3). Then we further develop a dynamic model, in which both physical capital and human capital investments are endogenously determined. The physical capital sector features investment-specific technological progress in an AK fashion, which yields sustainable growth. Human capital investment transforms unskilled labor into skilled labor, so the factor endowment structures also evolve endogenously. We analytically characterize how changes in endowment structures drive the dynamic changes in industrial compositions, life-cycle dynamic path of each underlying industry and the evolution of the skill premium along the aggregate growth path (Section 4). In particular, we show how the dynamic model could simultaneously explain all the stylized facts documented in Section 2. To better understand the underlying mechanisms, we first develop a simple model which allows for endogenous investment in physical capital but there is no human capital investment, so the composition of skilled labor and unskilled labor remains time-invariant (Section 4.1). Then we introduce endogenous decisions on human capital investment to the above setting, which allows us to explore how to balance these two different types of investment (Section 4.2). We show that the optimal human capital investment should be congruent with the demand for skills from the underlying industries, the composition of which varies endogenously in response to the changes in factor endowment structures. Skill premium is shown to change at the same frequency as the industrial compositions. After that, we further allow for the existence of positive externality in human capital investment, as is often examined in the literature, and explores its implications (Section 4.3). Section 5 concludes.

2. Stylized Facts

We document the following stylized facts about skill premium, factor endowments and industrial dynamics, which will be discussed in details in this section.

- **Fact 0 (positive correlation):** There exists a positive correlation between skill premium and physical capital to skilled-labor ratio (or capital-skill ratio, hereafter) for the aggregate manufacturing sector.
- **Fact 1 (cross-industry heterogeneity):** There exists tremendous cross-industry heterogeneity in capital-skill ratios.

- **Fact 2 (hump-shaped dynamics):** An industry typically exhibits a hump-shaped life cycle: its value-added share (or employment share) first increases, reaches the peak, and then declines.
- **Fact 3 (timing fact):** An industry with higher capital-skill ratio reaches its peak later.
- **Fact 4 (congruence fact):** The further away an industry's capital-skill ratio deviates from the economy's endowment structure (measured by the total capital-skill ratio), the smaller value added share (and employment share) is the industry.
- **Fact 5 (Asymmetric Duration fact):** The rising period of an industry (in terms of employment and/or value added) is longer than the declining period.

2.1 Evidence from US data

We use the NBER-CES Manufacturing Industry Data for the US. This data set adopts the 6-digit NAICS codes and covers 473 sub-industries within the manufacturing sector from 1958 to 2016. At this disaggregated level, the rank in terms of capital-skill ratio between two industries is frequently reversed over time, which creates challenges when empirically testing predictions from models that assume time-invariant (rank of) factor intensities for industries. To address this data issue that has been a long-time headache for empirical tests of the Hechscher-Ohin trade model, we follow Schott (2002) by redefining industries according to their capital to skilled-labor ratios. We first rank all the 27,751 observations consisting of 473 industries for 59 years (156 observations are dropped due to missing values in employment) by the capital–skill ratios in an increasing order, and then equally divide all these observations into 99 bins (newly-defined industries). Within each newly-defined industry, there are 280 observations. By construction, the capital to skill ratio is lowest in the first bin, called "industry 1" and highest in the last bin, or "industry 99". Moreover, the ranking of capital to skill ratio across the newly defined industries is time invariant.

Positive correlation between skill premium and capital skill ratio

Figure 1 plots skill premium against capital to skilled-labor ratio for the whole manufacturing sector from 1958 to 2016. Different dots represent observations in different years. Here due to lack of better measures of skilled labor and unskilled labor, we follow the literature by taking production workers as unskilled labor and non-production workers as skilled labor [[add relevant literature that also does this for the US manufacturing data]]. Skill premium is measured by the wage ratio between skilled and unskilled workers. A positive correlation between skill premium and the ratio of capital to skilled labor is discernible.[2]

[INSERT Figure 1: Positive correlation between skill premium and logarithm of capital to skilled labor ratio]

[2] A more commonly used measure of skilled and unskilled labor is eduaction level of workers, which is not available for the US manufacturing sector for the whole period of 1958-2016. However, this information is available for the US manufacturing sector from 1995 to 2009 in the WOID manfuacturign database. The positive correlation between skill premium and capital skilled labor ratio is even stronger. Refer to Figure 1A in the appendix.

Figure 2 below is a scatter plot showing that the positive correlation between skill premium and capital to skilled-labor ratio is robust using the WIOD Socio-Ecoomic Accounts (SEA) manufacturing database for the US between 1995 and 2009.

[INSERT FIGURE 2 HERE]

Figure 3 shows that the logarithm of physical capital to skilled-labor ratio increases over the period from 1958 to 2016.

[INSERT FIGURE 3 HERE]

Cross-industry Heterogeneity

There exists tremendous cross-industry heterogeneity in the capital to skilled-labor ratio among the 99 newly-defined industries. Table 1 shows that, among all industries in 1958, the highest capital–skilled labor ratio is 1638 US dollars per worker, which is 67 times larger than the lowest one in the same year. In 2016, the highest capital to skilled-labor ratio is still about 69 times higher than the lowest one. The standard deviation across industries is about 285 in 1958, and monotonically increases in each decade, reaching 602 in 2016.[3]

[INSERT TABLE 1]

Hump-shaped Life Cycle Pattern and Timing Fact

Figure 4 plots the time series of the HP-filtered employment shares of three newly-defined industries in the total manufacturing sector from 1958 to 2016.

[INSERT FIGURE 4 HERE]

It shows that the employment share decreases over time in the industry with the lowest capital to skilled-labor ratio (industry 1), exhibits a hump shape in the industry with a "middle" level of capital to skilled-labor ratio (industry 43), and increases over time in the industry with the highest capital to skilled-labor ratio (industry 99). Similar patterns are also observed when the employment share is replaced by the value-added share. It suggests that each industry exhibits a hump-shaped life cycle pattern (Fact 2) and that more capital-skill intensive industries reach their peaks later (Fact 3).

To establish Fact 2 and Fact 3 more rigorously, we run the following regression:

$$Y_{it} = \beta_0 + \beta_1 t + \beta_2 t^2 + \beta_3 k_i \cdot t + \beta_4 k_i + \beta_5 T_{it} + \beta_6 D_i + \beta_7 GDPGR_t + \varepsilon_{it}, \quad (1)$$

where Y_{it} is the employment or value-added share of industry i in the total manufacturing sector at year t; k_i is the year-average of the capital to skilled-labor ratios of industry i, T_{it} is the labor productivity of industry i at year t, respectively; D_i is the industry dummy; $GDPGR_t$ is the GDP growth rate; and ε_{it} is the error term. If the hump-shaped dynamic pattern is statistically valid, one should expect the coefficient for the quadratic term, β_2, to be negative and significant. In addition, after controlling for the labor productivity and

[3] If we do the same exercises based on the 473 originally defined industries, the cross-industry heterogeneity is even larger. See Table A1 in the Appendix.

the industry fixed effect, we know from (1) that industry i reaches its peak at $t_i^{\max} \equiv -\frac{\beta_1+\beta_3 k_i}{2\beta_2}$. That is, $\frac{\partial Y_{it}}{\partial t} > 0$ if and only if $t < t_i^{\max}$. If the timing fact is statistically valid, we should expect $-\frac{\beta_2}{\beta_3}$ to be positive, or equivalently, β_3 should be positive when β_2 is negative. Moreover, the peak time $t_i^{\max}$ must be positive when β_1 is positive. Table 2 reports all the (GLS and OLS) regression results, which all confirm the hump-shaped pattern and the timing fact.

[INSERT TABLE 2]

Another way to establish the timing fact is to directly regress the peak time of a newly-defined industry's share (either employment share or value-added share) on its capital to skilled-labor ratio. The results are reported in Table 3. Column (1) and column (3) show that the peak time of an industry is positively correlated with its capital to skilled-labor ratio. For comparison purpose, column (2) and column (4) show that more capital intensive industries reach their peaks later, where capital intensity is measured by the capital expenditure share (measured by one minus labor income share). It confirms the finding in Ju, Lin and Wang (2015). It suggests that capital skill ratio and capital intensity are two alternative good predictors of an industry's peak time.

[INSERT Table 3: Industries with higher capital to skill ratio reach peaks later.]

Congruence fact

To further understand what determines the industrial structures and their dynamics, we run the following regression:

$$Y_{it} = \beta_0 + \beta_1 \left| \frac{K_{it}/L_{sit} - K_t/L_{st}}{K_t/L_{st}} \right| + \beta_2 \cdot productivity_{it} + \beta_3 D_i + \varepsilon_{it}, \tag{2}$$

where Y_{it} is the employment (or value added) share of newly defined industry i in the whole manufacturing sector at year t. $\left| \frac{K_{it}/L_{sit} - K_t/L_{st}}{K_t/L_{st}} \right|$ is the absolute value of a normalized difference between industry i's capital to skilled-labor ratio and the aggregate capital to skilled-labor ratio for the manufacturing sector at year t, $productivity_{it}$ is the labor productivity of industry i at year t, D_i is dummy for industry i. The results are reported in Table 4.

[INSERT TABLE 4]

Column (1) (or Column (2)) shows that β_1 is negative and significant, indicating that the employment (or value added) share of an industry is smaller if the capital to skilled-labor ratio of the industry is less congruent with the endowment structure. We refer to this finding as the congruence fact. Column (3) and Column (4) report the results when we use the original NAICS industry classification. It shows that the results are robust. (what we need more is the positive correlation between labor productivity and capital skill ratio across industries, which is used to support an assumption in our model).

Asymmetric Duration Fact

Now we document a new fact about industry life cycle, that is, the rising period of an industry (in terms of employment and/or value added) is generally longer than the declining period. We call it asymmetric duration fact.

To measure the duration of the rising and declining period of industry n, we first plot the HP-filered time series of the industry's employment, then we find the time when the employment reaches the global maximum (denoted by $t_n^{\max}$). Next, we find the lowest possible horizontal line of employment that has two intersections with the employment curve. Denote the leftmost and rightmost intersection time by t_n^b and t_n^e, respectively. The duration of the rising period is $t_n^{\max} - t_n^b$ and the duration of the decling period is $t_n^e - t_n^{\max}$. Ideally, when the time series is long enough, t_n^b (t_n^e) would be the time point when industry n just enters (exits) the market. However, data limitation makes it impossible to observe the whole life span of every industry. Typically, the period from 1958 to 2016 only covers an incomplete part of the life cycle of an industry. That is, $t_n^e - t_n^b$ is typically shorter than the whole life span of industry n. Define $\mathcal{N}(\theta) \equiv \{n \,|t_n^e - t_n^b \geq \theta\}$, where θ is a positive constant. So $\mathcal{N}(\theta)$ is the set of industries whose rising period plus declining period observed in the data sample is at least θ. Obviously, the set $\mathcal{N}(\theta)$ is smaller when the cutoff value θ increases. We calculate the percentage of industries whose rising period is longer than the declining period among all the industries in set $\mathcal{N}(\theta)$, that is,

$$\Pr(t_n^{\max} - t_n^b > t_n^e - t_n^{\max} \,|n \in \mathcal{N}(\theta)\,) = \frac{\#\{n \in \mathcal{N}(\theta) \,|t_n^{\max} - t_n^b > t_n^e - t_n^{\max}\}}{\#\mathcal{N}(\theta)}.$$

Table 5 reports the percentage values for both the newly defined 99 industries and the original 6-digit NAICS 373 industries with different cutoff value θ.

[INSERT TABLE 5]

It shows that the rising period is systematicaly longer than the declining period in most industries, which we call asymmetric duration fact.

2.2 Evidence from cross-country data

We now turn to the evidence from the cross-country data. The WIOD Socio-Economic Accounts (SEA) manufacturing data set is at the two-digit level and consists of 14 sectors from 1995 to 2009. In WIOD (SEA) skill type is defined on the basis of the level of educational attainment. More specifically, the data set uses the 1997 International Standard Classification of Education (ISCED) classification to define low, medium and high skilled labor. Here we added up the medium skilled labor to high skilled labor, so the definition of two skill types is given in Table 6.

[INSERT TABLE 6: definitions of skills]

We replicate the same exercise to check all the five facts for other countries and we find that all of them are still valid. To save space, we delegate all the details into the appendix.

This concludes the empirical part of the paper. Motivated by these stylized facts, we now develop a theoretical model, which takes Fact 1 as exogenously given and endogenously generates Facts 0, 2, 3, 4 and 5 simultaneously. We start with the static model, in which all production factors are exogenously given.

3. Static Model

The model setting extends the model in Ju, Lin and Wang (2015) (JLW model thereafter) by differentiating two different types of labor for the purpose of studying skill premium and human capital investment. More concretely, the economy is inhabited by a continuum of identical households with measure equal to one. Each household is endowed with capital E, skilled labor L_s and unskilled labor L_u. The total labor $(L_s + L_u)$ can be equivalently interpreted as the total family size of each household with each family member endowed with one unit of labor. The JLW model is a special case of this model when all workers are skilled $(L_u = 0)$. The production function of the final commodity is

$$X = \sum_{n=0}^{\infty} \lambda^n x_n, \tag{3}$$

where x_n denotes intermediate good produced by industry n, λ^n is the marginal productivity of intermediate good x_n in the final good production. We require $\lambda > 1$ and $x_n \geq 0$ for any $n = 0, 1, 2, 3, ...$ Only the final commodity X can be used for consumption. The utility function is $CRRA$:

$$U(C) = \frac{C^{1-\sigma} - 1}{1-\sigma}, \text{ where } \sigma \in (0, 1], \tag{4}$$

where C denotes consumption per household.

All technologies exhibit constant returns to scale. Let $F_n(E, l_s, l_u)$ denote the production function for industry $n \geq 0$, where E , l_s and l_u denote physical capital, skilled labor and unskilled labor, respectively. Good 0 is produced with labor only, and we normalize the units such that one unit of labor produces one unit of good 0. Moreover, skilled labor and unskilled labor are perfectly substitutable with equal labor productivity when producing good 0. Thus $F_0(E, l_s, l_u) = l_u + l_s$. For each good $n \geq 1$, there are two alternative ways of production depending on whether physical capital is used: If capital is not used in production, one unit of labor, skilled or unskilled, produces $\frac{1}{b^n}$ units of good n, where $b > 1$. If capital is used in production, skilled labor is required because only skilled labor can operate the "machine", in which case it requires one unit of skilled labor and a^n units of physical capital to produce one unit of good n. In other words, capital and skilled labor are complementary. These two alternative ways can be used simultaneously, so for $n \geq 1$, $F_n(E, l_s, l_u)$ is equal to the following value:

$$F_n(E, l_s, l_u) = \max_{l_{s1}, l_{s2}} \{ \frac{l_u + l_{s1}}{b^n} + \min\{ \frac{E}{a^n}, l_{s2} \} \}$$

subject to

$$l_{s1} + l_{s2} \leq l_s, l_{s1} \geq 0, l_{s2} \geq 0.$$

It implies that

$$F_n(E, l_s, l_u) = \begin{cases} \frac{l_u}{b^n} + \min\{\frac{E}{a^n}, l_s\} = \frac{l_u}{b^n} + l_s, & \text{if } E \geq a^n l_s \\ \frac{l_u + l_s + \frac{(b^n - 1)}{a^n} E}{b^n} & \text{if } E < a^n l_s \end{cases}.$$

It shows that the marginal product of skilled labor, when equipped with enough capital, becomes strictly higher than that of unskilled labor or that of skilled labor without capital. So capital not only substitutes unskilled labor but also substitutes "unequipped" skilled labor.
To make the analysis non-trivial, we assume[4]

$$\min\{a - 1, b\} > \lambda > 1. \tag{5}$$

It implies that, without loss of generality, the industry production functions can be rewritten as

$$x_n = F_n(E, l_s, l_u) = \begin{cases} l_u + l_s, & \text{if } n = 0 \\ \frac{l_u}{b^n} + \min\{\frac{E}{a^n}, l_s\}, & \text{if } n \geq 1 \end{cases}. \tag{6}$$

All technologies are freely available.

Let the final commodity X be the numeraire. Let r , w_s and w_u denote the rental price of capital and wage rates for skilled and unskilled labor, respectively. All the markets are perfectly competitive. All firms maximize their profits by taking all prices as given, and each household maximizes her utility function (4) subject to the following budget constraint

$$C \leq w_u L_u + w_s L_s + rE.$$

By resorting to the Second Welfare Theorem, we obtain the following proposition.
Proposition 1. *Given factor endowment* E, L_u and L_s, *there exists a unique perfectly competitive market equilibrium, in which industrial output* $\{x_n\}_{n=0}^{\infty}$ *and the final output* X *are characterized in the following table:*

[INSERT TABLE 7: Quantities in Static Equilibrium]

Table 7: Quantities in Static Equilibrium

$0 \leq E < aL_s$	$a^n L_s \leq E < a^{n+1} L_s$ for $n \geq 1$
$x_0 = L_u + L_s - \frac{E}{a}$	$x_0 = L_u$
$x_1 = \frac{E}{a}$	$x_n = \frac{L_s a^{n+1} - E}{a^{n+1} - a^n}$
	$x_{n+1} = \frac{E - a^n L_s}{a^{n+1} - a^n}$
$x_j = 0$,for $\forall j \neq 0, 1$	$x_j = 0$,for $\forall j \neq 0, n, n+1$
$X = L_u + L_s + \frac{\lambda - 1}{a} E$	$X = L_u + \frac{\lambda^n (a - \lambda)}{a - 1} L_s + \frac{\lambda^{n+1} - \lambda^n}{a^{n+1} - a^n} E$

[4]Observe that if $\lambda \leq 1$, (3) , (5) and (6) imply that no equilibrium exists because a higher-indexed intermediate good (larger n) is always strictly more desirable to produce than any lower-indexed intermediate good as the former is more costly to produce but less productive in the final good production. $a > \lambda$ is imposed for the same reason. Similarly, no equilibrium exists if $b < \lambda$, because otherwise all unskilled labor will be allocated to the highest-indexed intermediate good, which does not exist. When $b = \lambda$, indeterminacy arises as unskilled can be allocated to produce any intermediate good in equilibrium and any industry could exist.

Proof. Refer to the appendix. Q.E.D

Observe that the active underlying industries are different when the capital to skilled-labor ratios (endowment structures) are different. Industry 0 is always active and all unskilled labor is employed in industry 0. The output of other active industries depends on $\frac{E}{L_s}$. More precisely, when $\frac{E}{L_s} \in (a^n, a^{n+1})$, for $n \geq 1$, the only two other active industries are industry n and industry $n+1$, whose capital to skilled labor ratios are closest to the endowment structure. When $\frac{E}{L_s} \in (0, a)$, only industry 0 and industry 1 coexist, and capital to skilled labor ratios of these two sectors are also closest to the endowment structure. These equilibrium results are consistent with Fact 4 (the congruence fact) documented in Section 2. The results in Table 6 are graphically illustrated in the following diagram.

[INSERT FIGURE 5: Quantities in Static Equilibrium]

Observe that the output of industry 0 first decreases with $\frac{E}{L_s}$ and then remains constant when all skilled labor is "absorbed away" from industry 0. For any $n \geq 1$, output of industry n increases with $\frac{E}{L_s}$ when $\frac{E}{L_s} \in (a^{n-1}, a^n)$ and then decreases with $\frac{E}{L_s}$ when $\frac{E}{L_s} \in (a^n, a^{n+1})$, which is consistent with Fact 2 (hump-shaped pattern) documented in Section 2. For each $n \geq 1$, industry n reaches its peak of output when $\frac{E}{L_s} = a^n$, which means that higher-indexed industries reach their peaks of output at higher levels of $\frac{E}{L_s}$, consistent with Fact 3 (the timing fact).

Table 6 shows that the aggregate production function $X(E, L_u, L_s)$ has endogenously different functional forms, depending on the endowment structure $\frac{E}{L_s}$. It is because endowment structures determine the composition of underlying industries. This feature is different from standard macro models, where the functional form of the aggregate production function is exogenously given and also assumed to be time invariant.

Observe that r, w_s and w_u are equal to the marginal products of capital, skilled labor and unskilled labor, respectively, so they can be directly derived from the endogenous aggregate production function shown in Table 6. For instance, when $a^n L_s \leq E < a^{n+1} L_s$ for any $n \geq 1$, the factor prices are, respectively, given by

$$r = \frac{\partial X}{\partial E} = \frac{\lambda^{n+1} - \lambda^n}{a^{n+1} - a^n}; w_s = \frac{\partial X}{\partial L_s} = \frac{\lambda^n (a - \lambda)}{a - 1}; w_u = \frac{\partial X}{\partial L_u} = 1.$$

Collorary Let p_n denote the market price for good n, and θ denote the labor income share in total GDP. All prices and factor shares are summarized in Table 8:

Table 8: Prices and Factor Income Shares in Static Equilibrium

$0 \leq E < aL_s$	$a^n L_s \leq E < a^{n+1} L_s$ for $n \geq 1$
$\frac{w_s}{w_u} = 1$	$\frac{w_s}{w_u} = \frac{\lambda^n(a-\lambda)}{a-1}$
$\frac{r}{w_s} = \frac{\lambda-1}{a}$	$\frac{r}{w_s} = \frac{\lambda-1}{a^n(a-\lambda)}$
$p_0 = w_u = 1$	$p_0 = w_u = 1$
$p_1 = \lambda$	$p_n = \lambda^n$
	$p_{n+1} = \lambda^{n+1}$
$\theta = \frac{L_u + L_s}{L_u + L_s + \frac{\lambda-1}{a} E}$	$\theta = \frac{L_u + \frac{\lambda^n(a-\lambda)}{a-1} L_s}{L_u + \frac{\lambda^n(a-\lambda)}{a-1} L_s + \frac{\lambda^{n+1}-\lambda^n}{a^{n+1}-a^n} E}$

It shows that skill premium $\frac{w_s}{w_u}$ is positively correlated with $\frac{E}{L_s}$ because both are weakly increasing in n. It is qualitatively consistent with Fact 0 documented earlier. More precisely, skill premium remains constant when the set of active industries remains unchanged and it exhibits a discountinous upward jump when a new (with a higher capital-skill ratio) industry first appears. After that, $\frac{w_s}{w_u}$remains constant again as $\frac{E}{L_s}$ increases, until an even newer industry first appears, at which point $\frac{w_s}{w_u}$ exhibits a positive jump. Although the factor income share is not the focus of our paper, but our model implies that labor income share in GDP fluctuates endogenously as capital increases. More specifically, as E increases, the labor income share strictly declines when the active industries do not change, and then the labor income share discontinuously jumps up when a new industry first becomes active, and then the labor income share strictly decreases until a new industry first appears, at which point the labor income share jumps up again, so on and so forth. The decreasing part is due to the fact that all factor prices remain constant whereas GDP increases with E. The jump is due to the fact that skilled wage w_s has a discontinuous upward jump when a new industry first appears. As $n \to \infty$, θ converges to the interval $(\frac{a-\lambda}{(a-1)\lambda}, \frac{a-\lambda}{a-1})$. Similarly, we can conduct comparative static analysis with respect to L_s instead of E. Obviously, θ increases with L_s when the active industries do not change, and then it jumps down discontinuously when a new industry first appears.

Next, we extend the static model into a dynamic one by allowing production factors to change endogenously over time.

4. Dynamic Model

The dynamic model consists of three parts. In Part A, we study the case in which physical capital changes endogenously but there is no human capital investment, in other words, the amount of skilled labor and unskilled labor is exogenous and time-invariant. In Part B, we further allow for endogenous human capital investment, which transforms unskilled labor into skilled labor, and we explore the interaction between endogenous physical capital investment and human capital investment in determining industrial dynamics and skill premium. In Part C, we further generalize the model in Part B by allowing for the existence of positive externalities in human capital investment.

Part A: No Human Capital Investment

In this part, we let capital stock K grow endogenously but keep L_s and L_u fixed over time. Time is continuous and households are infinitely lived. Like in the JLW model, there are two sectors. One sector produces capital goods, which cannot be consumed directly. More specifically, one unit of capital inherited from the past produces ξ units of new working capital, where parameter ξ captures the investment-specific technological progress. The other sector produces the final commodity and all the intermediate goods. The final commodity is for consumption and is not storable. This sector is characterized in the static model in Section 3. To produce consumption good $C(t)$ at time point t, it requires $E(C(t))$ physical capital, where function $E(\cdot)$ is obtained by solving the inverse function from the bottom row in Table 7. More specifically,

$$E(C)=\left\{\begin{array}{cc} 0, & C=L_u+L_s \\ E_{0,1}(C), & \text{if } L_u+L_s<C<\lambda L_s+L_u \\ E_{0;n}(C), & \text{if } C=\lambda^n L_s+L_u, \text{for } n\geq 1 \\ E_{0,n,n+1}(C), & \text{if } \lambda^n L_s+L_u<C<\lambda^{n+1}L_s+L_u \text{ for } n\geq 1 \end{array}\right., \tag{7}$$

where

$$\begin{aligned} E_{0,1}(C) &\equiv \frac{a}{\lambda-1}\left(C-L_u-L_s\right), \\ E_{0;n}(C) &\equiv a^n L_s, \text{for } n\geq 1, \\ E_{0,n,n+1}(C) &\equiv \frac{a^{n+1}-a^n}{\lambda^{n+1}-\lambda^n}\left[C-L_u-\frac{\lambda^n(a-\lambda)}{a-1}L_s\right], \text{for } n\geq 1. \end{aligned}$$

By the Second Welfare Theorem, we can characterize the decentralized market equilibrium by solving the following artificial social planner problem:

$$\max_{C(t)}\int_0^\infty \frac{C(t)^{1-\sigma}-1}{1-\sigma}e^{-\rho t}dt, \text{ where } \sigma\in(0,1),$$

subject to

$$\dot{K}=\xi K-\delta\cdot E(C(t)), \tag{8}$$

where ρ is the time discount rate, δ is the depreciation rate, $E(C(t))$ is given by (7), and $K(0)=K_0$ is given. (8) states that the newly produced working capital net of the depreciated capital in the production of the consumption good is used for capital accumulation. Let t_n^b and t_n^e denote, respectively, the first and the last time points when output of good n reaches the highest level for $n\geq 1$. Let t_0^e denote the final time point when only good 0 is produced. To ensure positive growth and exclude explosive growth, we assume $0<\xi-\rho<\sigma\xi$. The optimization problem can be rewritten as

$$\max_{C(t)}\int_0^{t_0^e}\frac{C(t)^{1-\sigma}-1}{1-\sigma}e^{-\rho t}dt+\sum_{n=0}^{\infty}\int_{t_n^e}^{t_{n+1}^b}\frac{C(t)^{1-\sigma}-1}{1-\sigma}e^{-\rho t}dt+\sum_{n=0}^{\infty}\int_{t_{n+1}^b}^{t_{n+1}^e}\frac{C(t)^{1-\sigma}-1}{1-\sigma}e^{-\rho t}dt$$

subject to

$$\dot{K}=\left\{\begin{array}{ccc} \xi K & \text{when} & 0\leq t\leq t_0^e \\ \xi K-\delta E_{0,1}(C), & \text{when} & t_0^e\leq t\leq t_1^b \\ \xi K-\delta E_{0;1}(C), & \text{when} & t_1^b\leq t\leq t_1^e \\ \xi K-\delta E_{0,n,n+1}(C), & \text{when} & t_n^e\leq t\leq t_{n+1}^b, \text{ for } n\geq 1 \\ \xi K-\delta E_{0,n+1}(C) & \text{when} & t_{n+1}^b\leq t\leq t_{n+1}^e, \text{ for } n\geq 1 \end{array}\right.,$$

K_0 is given.

Following the Kamien and Schwartz (1991), we could solve this Hamiltonian dynamic system with endogenously switching state equations. The reason why the state equations are switching is that the underlying industrial structures are different for different output level of consumption goods.

Proposition 2. *Suppose K_0 is sufficiently small. In equilibrium consumption changes by alternating between a stagnation period and a growth period. More pricisely, we have:*

$$C(t) = \begin{cases} L_u + L_s, & \text{when} \quad 0 \le t \le t_0^e \\ (L_u + L_s)e^{g_C(t-t_0^e)}, & \text{when} \quad t_0^e \le t \le t_1^b \\ L_u + \lambda L_s, & \text{when} \quad t_1^b \le t \le t_1^e \\ (L_u + \lambda^n L_s)e^{g_C(t-t_n^e)}, & \text{when} \quad t_n^e \le t \le t_{n+1}^b, \text{ for } n \ge 1 \\ L_u + \lambda^{n+1} L_s & \text{when} \quad t_{n+1}^b \le t \le t_{n+1}^e, \text{ for } n \ge 1 \end{cases},$$

where $g_C = \frac{\xi-\rho}{\sigma}$.
Proof. See the appendix.
The following graph plots the equilibrium time path of consumption.

[insert new figure 6]

As we can see, initially no capital is used for producing consumption goods because capital is too scarce and expensive, so all capital is used for cumulation in the capital goods sector till time t_0^e. After that, good 0 and good 1 are simultaneously produced and some capital is used for producing consumption goods because capital becomes cheap enough to be used for consumption goods, aggregate consumption grows at a constant rate $g_C = \frac{\xi-\rho}{\sigma}$, which is the standard Euler equation, till time t_1^b . During this growth period $(t_0^e, t_1^b]$, the rental price of capital relative to consumption good is equal to $\frac{\lambda-1}{a}(= \frac{\partial X}{\partial E}$, where $X = L_u + L_s + \frac{\lambda-1}{a}E)$ according to Table 7, and relative factor price is $\frac{r}{w_s} = \frac{\lambda-1}{a}$ (from Table 8) remains unchanged over time. Starting from t_1^b, only good 1 is produced and consumption remains constant at level $L_u + \lambda L_s$ till point t_1^e, and capital used for consumption goods remains constant because capital is too scarce and hence too expensive to support the production of good 2, which is more capital intensive. Notice that at time t_1^b, the rental price of capital relative to consumption good is equal to $\frac{\lambda-1}{a}$but at time t_1^e the rental price of capital relative to consumption good is equal to $\frac{\lambda^2-\lambda}{a^2-a}(= \frac{\partial X}{\partial E}$,where $X = L_u + \frac{\lambda(a-\lambda)}{a-1}L_s + \frac{\lambda^2-\lambda}{a^2-a}E)$. (5) implies that $\frac{\lambda^2-\lambda}{a^2-a} < \frac{\lambda-1}{a}$, so capital has to be further accumulated in the capital goods sector during this period until it is abundant enough to engage the more capital-intensive technology (good 2) in consumption production. Observe that relative factor price $\frac{r}{w_s} = \frac{\lambda-1}{a}$at t_1^b and $\frac{r}{w_s} = \frac{\lambda-1}{a(a-\lambda)}$ at t_1^e based on Table 8. The discontinuity in the relative factor price at consumption level $L_u + \lambda L_s$ in the static model, or equivalently, the non-differentiability of function $E(C)$ at $C = L_u + \lambda L_s$, translates into a period with constant consumption in the dynamic model. During this period, the relative factor price changes continuously over time. Consumption grows at rate $\frac{\xi-\rho}{\sigma}$again during $[t_1^e, t_2^b]$, followed by a constant consumption period, ad infinitum.
The following proposition characterizes the life cycle dynamics of each industry. Define $m_n \equiv t_{n+1}^b - t_n^e$, $\forall n \ge 0$.
Proposition 3. *Suppose K_0 is sufficiently small. The economy starts with only industry 0 producing till time t_0^e. For each industry $n \ge 1$, its output exhibits a hump-shaped life-cycle dynamic pattern: it first enters the market at time t_{n-1}^e, its output rises for a period of m_{n-1}, reaches the peak at time t_n^b, and then the output remains constant till time t_n^e,* after

which the output *declines for a period of* m_n*, industry* n *disappears after* t^e_{n+1}*, so its whole life span is* $m_{n-1}+t^e_n-t^b_n+m_n$ *, where*

$$m_n = \frac{\sigma \log \frac{L_u+\lambda^{n+1}L_s}{L_u+\lambda^n L_s}}{\xi-\rho}, \forall n \geq 0. \tag{9}$$

and

$$t^e_n - t^b_n = \frac{\ln \frac{a}{\lambda}}{\xi-\rho}, \forall n \geq 1.$$

Proof. Refer to the Appendix. **Q.E.D.**

The industrial dynamics characterized in the above theorem can be more intuitively illustrated in the following diagram.

[INSERT FIGURE 7]

It shows that the model predictions are qualitatively consistent with the stylized facts documented in Section 2. More concretely, each industry $n \geq 1$ exhibits a hump-shaped life cycle pattern (Fact 2), industries with higher capital-skill ratios reach their peak values later (Fact 3), active industries are those whose capital-skill ratios are closest to the ratio of total physical capital and skilled labor in the whole consumption good sector (Fact 4). Moreover, observe that $\frac{\partial m_n}{\partial n} < 0$, that is, the booming period is longer than the decline period ($m_{n-1} > m_n$), consistent with the asymmetric duration fact (Fact 5). As $t \to \infty$, $m_n \to \frac{\log \lambda}{g_C}$. By comparison, the booming and declining periods of industry n are equal ($m_{n-1} = m_n = \frac{\log \lambda}{g_C}$) in the JLW model (obtained by substituting $L_u = 0$ into (9)), so the life span of every industry is equal to $\frac{(2\sigma-1)\log\lambda+\ln a}{\xi-\rho}$, which is longer than the life span of the counterpart in the current model.
Combining the static results in Table 6 and the dynamic results summarized in Proposition 2 and Proposition 3, we obtain the following proposition about how skill premium interacts with industrial dynamics over time.
Proposition 4. *The skill premium* $\frac{w_s}{w_u}$ *is equal to one before time* t^e_0, and *it is equal to* $\frac{\lambda^n(a-\lambda)}{a-1}$ *when* $t \in [t^e_n, t^e_{n+1})$, for all $n \geq 1$. *Alternatively speaking, during the life span of industry* $n \geq 1$*, skill premium remains constant at* $\frac{\lambda^n(a-\lambda)}{a-1}$ *from* t^e_{n-1}*, the time point when industry* n *first enters the market, till time* t^e_n*, at which point the skill premium jumps up discontinuously to* $\frac{\lambda^{n+1}(a-\lambda)}{a-1}$ *and industry* n *starts to decline. The skill premium remains at level* $\frac{\lambda^{n+1}(a-\lambda)}{a-1}$ *during the whole declining period of industry* n *till it exits the market at* t^b_{n+1}*..*
The following figure plots the dynamics of skill premium.

[*INSERT FIGURE* 8]

We can see that the skill premium remains constant over time and jumps up discontinuously only when a new (more capital-intensive) industry just enters the market. Moreover, the skill premium changes more and more frequently over time, eventually converging to a constant frequency.

Part B: Human Capital Investment

Now we introduce human capital investment into the model. Unskilled labor can be transformed into skilled labor via human capital investment. Both physical capital investment and human capital investment are endogenously decided by private agents. By the Second Welfare Theorem, we characterize the decentralized market equilibrium by solving the following artificial benevolent social planner problem:

$$\max_{C(t),G(t)} \int_0^\infty \frac{C(t)^{1-\sigma}-1}{1-\sigma} e^{-\rho t} dt, \text{ where } \sigma \in (0,1),$$

subject to

$$\begin{aligned} \dot{K} &= \xi K - \delta \cdot E(C(t), L_s, L_u) - G(t), && (10) \\ \dot{L}_s &= \phi L_u^{1-\kappa} \cdot G(t)^\kappa, && (11) \\ \dot{L}_u &= g_b \cdot (L_u + L_s) - \dot{L}_s, && (12) \end{aligned}$$

and that $K(0), L_u(0), L_s(0)$ are all given, where $E(C(t), L_s, L_u)$ is still given by (7), but L_s and L_u are now explicitly listed as endogenous state variables, $G(t)$ is human capital investment (per household) at time t and g_b is the exogenous birth rate. We assume

$$g_b < \xi < \frac{\xi - \rho}{\sigma}, \tag{13}$$

which will be explained soon.

(10) states that human capital investment is costly and paid in terms of physical capital. (11) says that how many new skilled labor is "produced" shall depend positively on the total size of the current pool of unskilled labor $L_u(t)$ and the human capital investment $G(t)$ per household. The parameter $\kappa \in (0,1)$ captures the relative importance of human capital investment in "producing" skilled labor. A larger κ means that the process of transforming unskilled labor into skilled labor is more intensive in human capital investment. The strictly positive parameter ϕ measures the general efficiency of the skill transformation process, capturing all relevant factors such as the quality of the training program. (12) states that the increase in the number of unskilled labor is equal to the newly born labor $g_b \cdot (L_u + L_s)$ (who are assumed to be unskilled labor automatically) net of those who are just transformed into skilled labor $\dot{L}_s$. Let $\widehat{t}_n^b$ and $\widehat{t}_n^e$ denote the time point when output of good n first and last reaches the highest level, respecitvely, for $n \geq 1$. Let $\widehat{t}_0^e$ denote the last time point when only good 0 is produced.

Proposition 5. *When both physical capital investment and human capital investment are endogenous, the consumption alternates between a high growth period with constant growth rate given by*[5]

$$g_C = \frac{\xi - \rho}{\sigma}, \forall t \in [\widehat{t}_n^e, \widehat{t}_{n+1}^b], \text{ for all } n \geq 0$$

[5] Note that the total measure of household remains constant equal to one but the population increases over time. Now $C(t)$ is consumption per household and per capita consumption is $\frac{C(t)}{L_u(t)+L_s(t)}$.

and *a low growth period with growth rate* g_b *for* $\forall t \in [\widehat{t}_n^b, \widehat{t}_n^e]$ for all $n \geq 1$. *there exists a unique but different temporary Balanced Growth Path (BGP) for each different stage of development (i.e., different industrial structure), on which the following is true*

$$\frac{\dot{L}_s}{L_s} = \frac{\dot{L}_u}{L_u} = \frac{\dot{G}}{G} = g_b \tag{14}$$

for $n \geq 1$. *Moreover, on the temporary BGP when industries* n and $n+1$ coexist ($t \in (\widehat{t}_n^e, \widehat{t}_{n+1}^b)$), *the following is true*

$$\frac{L_s}{L_u} = \frac{\phi \chi_n^\kappa}{g_b}, \tag{15}$$

$$\frac{G(t)}{L_u(t)} = \chi_n, \tag{16}$$

where χ_n *is uniquely determined by*

$$(1-\kappa)\,\phi\chi_n^\kappa = \frac{a^n}{\lambda-1}\left\{(a-\lambda) - \frac{a-1}{\lambda^n}\right\}\phi\kappa\delta{\chi_n}^{\kappa-1} - \xi, \tag{17}$$

for $n \geq 1$. *In the very long run,* $\frac{L_s}{L_u+L_s} \to 1$*, and the economy converges to one with only skilled labor and physical capital, which is characterized in Ju, Lin and Wang (2015).*
Proof. See the appendix. **Q.E.D.**

The proposition shows that the household consumption growth rate after industrialization is still constant and equal to $g_C = \frac{\xi-\rho}{\sigma}$, so the growth rate of per capita consumption becomes $g_C - g_b$, which is positive due to (13). It implies that the capital devoted to the production of consumption goods increases at a rate higher than g_C, so that capital per skilled labor increases sufficiently fast to support the consumption growth rate g_C. (14) shows that human capital investment, skilled labor and unskilled labor all grow at the same constant rate as the birth rate g_b on the *temporary* BGP when industry n and industry $n+1$ coexist. Moreover, (15) and (16) state that the skill structure of the labor pool $\frac{L_s}{L_u}$ and human capital investment per unskilled labor $\frac{G}{L_u}$ both remain constant on this temporary BGP.
The transitional dynamics between two neighboring temporary BGPs is the following: The temporary BGP with the coexistence of industries $0, n$ and $n+1$ is asymptotically reached when t gets sufficiently close to $\widehat{t}_{n+1}^e$, and then at time point $\widehat{t}_{n+1}^e$, industry n disappears and industry $n+2$ is about to enter, the skill premium discontinuously jumps up, and optimal human capital investment per unit of unskilled labor $\frac{G}{L_u}$discontinuously jumps (up) from χ_n to χ_{n+1}, so L_s grows faster than g_b for a while until a new temporary BGP (with industries $0, n+1$ and $n+2$ coexisting) is asymptotically reached as t gets sufficiently close to $\widehat{t}_{n+2}$, so on and so forth.
Corollary 1. *With endogenous human capital investment, the peak value of output of each industry* $n \geq 1$ *is strictly larger than the counterpart without human capital investment so long as the population growth rate is non-negative. That is,* $\widehat{x}_n(\widehat{t}_n^b) > x_n(t_n^b)$ *for any* $n \geq 1$*, if* $g_b \geq 0$.
Proof. Observe that $\widehat{x}_n(\widehat{t}_n^b) = L_u(\widehat{t}_n^b) + \lambda^n L_s(\widehat{t}_n^b)$ and $x_n(t_n^b) = L_u + \lambda^n L_s$. When $g_b \geq 0$, we have $L_u(\widehat{t}_n^b) + L_s(\widehat{t}_n^b) \geq L_u + L_s$ (with $=$ holding if and only if $g_b = 0$). Moreover, $L_s(\widehat{t}_n^b) > L_s$

because of positive human capital investment as implied in Proposition 5, so $\widehat{x}_n(\widehat{t_n^b}) > x_n(t_n^b)$ as $\lambda > 1$. **Q.E.D**

Corollary 2. *The following is true on the temporary BGP when industries* $0, n$ *and* $n+1$ *coexist for any* $n \geq 1$*:*

$$\frac{\partial \chi_n}{\partial n} > 0; \frac{\partial \chi_n}{\partial \delta} > 0; \frac{\partial \chi_n}{\partial \phi} > 0; \frac{\partial \chi_n}{\partial a} > 0; \frac{\partial \chi_n}{\partial \xi} < 0. \tag{18}$$

Proof. Immediately implied by (17). **Q.E.D**

(18) implies that $\frac{L_s}{L_u}$ and $\frac{G}{L_u}$ both become strictly higher when the underlying supporting industries have higher capital-skill ratios (n). As a result, $\frac{L_s}{L_u}$ also increases when industrial upgrading occurs. Moreover, (15) and (17) jointly imply

$$\lim_{n\to\infty} \frac{L_s}{L_u + L_s} = 1,$$

which means that all labor will be skilled labor in the very long run, the scenario as characterized in Ju, Lin and Wang (2015).

(18) also shows that $\frac{L_s}{L_u}$ and $\frac{G}{L_u}$ both increase when capital depreciation rate δ becomes larger, or when the efficiency of skill transformation ϕ becomes higher, or when the capital-skill ratio ladder of two neighboring industries a becomes larger. The reason is that a higher δ weakens the desirability of using physical capital to produce consumption goods, so skilled labor and hence human capital investment will be more favored. As a result, $\frac{G}{L_u}$ increases and $\frac{L_s}{L_u}$ increases. Similarly, an increase in the efficiency of skill transformation ϕ enhances the marginal return of human capital investment, so it induces a higher G and higher L_s. A higher a means that industrial upgrading becomes permanently more costly, therefore, to ensure a constant positive consumption growth, it would be better to enhance human capital investment to produce more skilled labor, leading to higher $\frac{L_s}{L_u}$ and $\frac{G}{L_u}$.

Both $\frac{L_s}{L_u}$ and $\frac{G}{L_u}$ decrease with ξ because a larger ξ implies that the physical capital production is more efficient, so physical capital becomes more valuable relative to human capital, and therefore, the incentive to invest human capital is weakened. As a result, both $\frac{G(t)}{L_u(t)}$ and $\frac{L_s}{L_u}$ decrease with ξ.

Furthermore, $\frac{L_s}{L_u}$ decreases with g_b because faster growth of the unskilled labor pool means less human capital investment for each unskilled labor for any given amount of total human capital investment, hence a smaller fraction of skilled labor in the steady state. However, $\frac{G}{L_u}$ is independent of g_b as G can discontinuously jump up whenever new industries emerge, but L_s has to change continuously.

Observe that $\frac{\chi_{n+1}}{\chi_n} > a, \forall n \geq 0$ because (17) implies that

$$\frac{\chi_{n+1}^{\kappa}}{\chi_n^{\kappa}} = \frac{\frac{a^{n+1}}{\lambda-1}\left\{(a-\lambda) - \frac{a-1}{\lambda^{n+1}}\right\}\phi\kappa\delta\chi_{n+1}^{\kappa-1} - \xi}{\frac{a^{n}}{\lambda-1}\left\{(a-\lambda) - \frac{a-1}{\lambda^{n}}\right\}\phi\kappa\delta\chi_{n}^{\kappa-1} - \xi} > \frac{\frac{a^{n+1}}{\lambda-1}\left\{(a-\lambda) - \frac{a-1}{\lambda^{n+1}}\right\}\phi\kappa\delta\chi_{n+1}^{\kappa-1}}{\frac{a^{n}}{\lambda-1}\left\{(a-\lambda) - \frac{a-1}{\lambda^{n}}\right\}\phi\kappa\delta\chi_{n}^{\kappa-1}} > \frac{a\chi_{n+1}^{\kappa-1}}{\chi_n^{\kappa-1}}.$$

Thus, we conclude from (16) that the human capital investment per unskilled labor on the new temporary BGP becomes more than a times larger than that on the previous temporary BGP. Similarly, (15) implies that skilled to unskilled labor ratio becomes a^{κ} times higher on the new temporary BGP than the previous temporary BGP.

Define $\widehat{m}_n \equiv \widehat{t}^b_{n+1} - \widehat{t}^e_n, \forall n \geq 0$.

Proposition 6. *For each industry $n \geq 1$, its output exhibits a hump-shaped life-cycle dynamic pattern: it enters the market at time $\widehat{t}^e_{n-1}$, its output rises for a period of $\widehat{m}_{n-1}$, and reaches the peak at time $\widehat{t}^b_n$, after which its output declines for a period of $\widehat{m}_n$ and the industry exit the market after $\widehat{t}^e_{n+1}$. During industry n's whole life span, skill premium is $\frac{\lambda^{n-1}(a-\lambda)}{a-1}$ when industry n is booming (i.e.,* when *$t \in [\widehat{t}^e_{n-1}, \widehat{t}^b_n)$) and the skill premium jumps to $\frac{\lambda^n(a-\lambda)}{a-1}$ when industry n* declines *(i.e., when $t \in [\widehat{t}^e_n, \widehat{t}^b_{n+1})$), where*

$$\widehat{m}_n \approx \frac{\log \frac{g_b+\phi\lambda^{n+1}\chi_n^\kappa}{g_b+\phi\lambda^n\chi_{n-1}^\kappa}}{\frac{\xi-\rho}{\sigma} - g_b}, \forall \text{ sufficiently large } n > 1, \tag{19}$$

where χ_n is determined by (17). Moreover, $\widehat{m}_n > m_n$ for all $n > 1$, where m_n is given by (9).

Proof. Since C grows at a constant rate $g_C = \frac{\xi-\rho}{\sigma}$ on the BGP, and

$$C(\widehat{t}_n) = L_u(\widehat{t}_n) + \lambda^n L_s(\widehat{t}_n), \text{for } n \geq 1,$$

thus we have

$$\begin{aligned}
\widehat{m}_n &= \frac{\log \frac{L_u(\widehat{t}_{n+1})+\lambda^{n+1}L_s(\widehat{t}_{n+1})}{L_u(\widehat{t}_n)+\lambda^n L_s(\widehat{t}_n)}}{g_C} = \frac{\log \frac{L_u(\widehat{t}_{n+1})}{L_u(\widehat{t}_n)} \frac{1+\lambda^{n+1}\frac{L_s(\widehat{t}_{n+1})}{L_u(\widehat{t}_{n+1})}}{1+\lambda^n \frac{L_s(\widehat{t}_n)}{L_u(\widehat{t}_n)}}}{g_C} \\
&\approx \frac{\log e^{g_b(\widehat{t}_{n+1}-\widehat{t}_n)} \frac{1+\lambda^{n+1}\frac{\phi\chi_n^\kappa}{g_b}}{1+\lambda^n\frac{\phi\chi_{n-1}^\kappa}{g_b}}}{g_C} = \frac{g_b\widehat{m}_n + \log \frac{1+\lambda^{n+1}\frac{\phi\chi_n^\kappa}{g_b}}{1+\lambda^n\frac{\phi\chi_{n-1}^\kappa}{g_b}}}{g_C},
\end{aligned}$$

where the last equation uses the definition of $\widehat{m}_n$. Solving out $\widehat{m}_n$ yields (19). The third semi-equation comes from (14) and (15). It is not exactly equal for two reasons. First, the growth rate of L_u is constant at rate g_b only on the BGP, but not always so off the BGP. For example, during the transition period between two different temporary steady state levels of $\frac{G}{L_u}$, say χ_n and χ_{n+1}, the growth rate of L_u is lower than g_b because human capital investment G and L_s both grow at a rate higher than g_b, so $\frac{L_u(\widehat{t}_{n+1})}{L_u(\widehat{t}_n)}$ is actually smaller than $e^{g_b(\widehat{t}_{n+1}-\widehat{t}_n)}$. Second, $\frac{L_s(\widehat{t}_{n+1})}{L_u(\widehat{t}_{n+1})}$ cannot jump discontinuously at time point $\widehat{t}_{n+1}$ although $G(t)$ jumps up at point $\widehat{t}_{n+1}$, so the ratio $\frac{L_s(\widehat{t}_{n+1})}{L_u(\widehat{t}_{n+1})}$ is the same as the one on the BGP when industry n and industry $n+1$ coexist.

To show $\widehat{m}_n > m_n$, notice that, first, the aggregate consumption growth rate g_c is always the same. Second, both L_u and L_s remain constant in Part A but grow at a positive rate (equal to g_b on the BGP) in Part B, so the capital devoted to consumption production (E) increases more slowly in Part B than in Part A. As a result, $\frac{E}{L_s}$ grows more slowly in Part B than in Part A, which means it takes a longer time for $\frac{E}{L_s}$ to increase from a^n to a^{n+1}, so $\widehat{m}_n > m_n$. **Q.E.D**

This proposition largely resembles Proposition 2 and Proposition 3 in that all predictions are still consistent with Facts 0-4. Moreover, same as in Part A, the "shakeout" pattern of industrial dynamics is also preserved, as implied by (19). The major difference is that, now with endogenous human capital investment, industry life spans are longer than before, so the skill premium changes less frequently than in Part A. This would be still true even when $g_b = 0$ because L_s keeps increasing in Part B, so $\frac{E(C)}{L_s}$ grows more slowly than in Part A because C still grows at the speed $g_C = \frac{\xi-\rho}{\sigma}$. Observe that $\underset{n\to\infty}{lim}\, m_n = \frac{\log\lambda}{g_C - g_b}$, which means that life spans of industries, hence the frequencies of skill premium adjustment, eventually will be identical.

Part C. Human Capital Externality

Now we introduce externality in human capital investment. Everything is identical to Part B except that (11) is replaced by the following:

$$\dot{L}_s = \phi \left[L_u^{1-\kappa} \cdot G(t)^{\kappa}\right]^{1-\zeta} \overline{G}^{\zeta}(t), \tag{20}$$

where $\overline{G}(t)$ is the average household spending on human capital investment at time t and the parameter $\zeta \geq 0$. Part B is a special case when $\zeta = 0$. When $\zeta > 0$, it captures the positive externality in human capital investment. For instance, when your neighbors make private investment on human capital for their family members such as participating in training programs or going to professional schools, it helps create market demand for those services, so both the supply and quality of such services will improve due to market competition. Moreover, watching and interacting with your neighbours who purchase those services may also help increase the skills of your family members. You benefit from your neighbors' private invesment in human capital without compensating them.

Proposition 6. *In the laissez affaire decentralizaed market equilibrium with positive externality in human capital investment ($\zeta > 0$), there exists a unique but different temporary Balanced Growth Path (BGP) for each different stage of development. More specifically, denote human capital investment per unskilled labor by $\widetilde{\chi}_n(t) \equiv \frac{G(t)}{L_u(t)}$ when industry n and industry $n+1$ coexist with industry 0 (that is, $t \in [\widetilde{t}_n, \widetilde{t}_{n+1})$). The following is true on the temporary BGP for the corresponding stage of development:*

$$g_C = \frac{\xi - \rho}{\sigma}$$

$$\frac{\dot{L}_s}{L_s} = \frac{\dot{L}_u}{L_u} = \frac{\dot{G}(t)}{G(t)} = g_b \tag{21}$$

$$\frac{L_s}{L_u} = \frac{\phi}{g_b}\widetilde{\chi}_n^{\kappa(1-\zeta)+\zeta}, \tag{22}$$

where $\widetilde{\chi}_n$ is uniquely determined by

$$\frac{\xi}{\phi\kappa(1-\zeta)}\widetilde{\chi}_n^{(1-\kappa)(1-\zeta)} = \delta\left[\frac{a^n(a-\lambda)}{\lambda-1} - \frac{a^{n+1}-a^n}{\lambda^{n+1}-\lambda^n}\right] - \frac{(1-\kappa)}{\kappa}\widetilde{\chi}_n, \tag{23}$$

which implies the following properties:

$$\frac{\partial \widetilde{\chi}_n}{\partial \xi} < 0; \frac{\partial \widetilde{\chi}_n}{\partial \delta} > 0; \frac{\partial \widetilde{\chi}_n}{\partial n} > 0; \frac{\partial \widetilde{\chi}_n}{\partial \phi} > 0; \frac{\partial \widetilde{\chi}_n}{\partial \zeta} < 0; \frac{\partial \widetilde{\chi}_n}{\partial a} > 0 \tag{24}$$

for $\forall n \geq 1$*. In the very long run,* $\frac{L_s}{L_u+L_s} \to 1$*, and the economy converges to one with only skilled labor and physical capital, which is characterized in Ju, Lin and Wang (2015).*

Proof: See details in the appendix. **Q.E.D**

Comparing (24) with (18) in Part B, we see that human capital investment per unskilled labor $\widetilde{\chi}_n$ decreases with the degree of externality ζ, as it encourages free riding and dampens private incentive to invest in human capital. It implies that $\widetilde{\chi}_n$ is strictly lower than χ_n as characterized in Part B. All other comparative static properties of $\widetilde{\chi}_n$ are same as those of χ_n. Comparing (15) and (22) shows that, when $\widetilde{\chi}_n$ increases by one per cent, $\frac{L_s}{L_u}$ increases by $\kappa(1-\zeta)+\zeta$ per cent, which is more elastic than the case in part B. It is because the former benefits from externalities, so the gain is larger for the same amount of private human capital invesment. Moreover, following the same method as before, we can show that $\frac{\widetilde{\chi}_{n+1}}{\widetilde{\chi}_n} > a$, *for* $\forall n \geq 1$.

The counterpart to Proposition 6 is the following:

Define $\widetilde{m}_n \equiv \widetilde{t}_{n+1} - \widetilde{t}_n$, $\forall n \geq 0$.

Proposition 7. *In the laissez affaire decentralizaed market equilibrium with positive externality in human capital investment* ($\zeta > 0$)*, for each industry* $n \geq 1$*, its value added exhibits a hump-shaped life-cycle dynamic pattern: it appears at time* $\widetilde{t}_{n-1}$*, its output rises for a period of* $\widetilde{m}_{n-1}$*, and reaches the peak at time* $\widetilde{t}_n$*, after which its output declines for a period of* $\widetilde{m}_n$ *and disappears after* $\widetilde{t}_{n+1}$*. During industry* n*'s whole life span* $\widetilde{m}_{n-1} + \widetilde{m}_n$ *, skill premium is* $\frac{\lambda^{n-1}(a-\lambda)}{a-1}$ *when industry* n *is booming (i.e.,* when $t \in [\widetilde{t}_{n-1}, \widetilde{t}_n)$*) and the skill premium jumps to* $\frac{\lambda^n(a-\lambda)}{a-1}$ *when industry* n declines *(i.e., when* $t \in [\widetilde{t}_n, \widetilde{t}_{n+1})$*), where*

$$\widetilde{m}_n \approx \frac{\log \frac{g_b + \phi\lambda^{n+1}\widetilde{\chi}_n^{\kappa(1-\zeta)+\zeta}}{g_b+\phi\lambda^n \widetilde{\chi}_{n-1}^{\kappa(1-\zeta)+\zeta}}}{\frac{\xi-\rho}{\sigma} - g_b}, \forall \text{ sufficiently large } n > 1, \tag{25}$$

where $\widetilde{\chi}_n$ *is determined by (23). Moreover,* $\widetilde{m}_n > \widehat{m}_n$ *for all* $n > 1$*, where* $\widehat{m}_n$ *is given by (19).*

The proposition states that the life span of an industry is longer than that of the corresponding industry without human capital externality ($\widetilde{m}_n > \widehat{m}_n$ *for all* $n > 1$). Why? Because

$$\widetilde{m}_n = \frac{\log \frac{L_u(\widetilde{t}_{n+1})+\lambda^{n+1}L_s(\widetilde{t}_{n+1})}{L_u(\widetilde{t}_n)+\lambda^n L_s(\widetilde{t}_n)}}{g_C} = \frac{\log \left[1 + \frac{L_u(\widetilde{t}_{n+1})-L_u(\widetilde{t}_n)+\lambda^{n+1}L_s(\widetilde{t}_{n+1})-\lambda^n L_s(\widetilde{t}_n)}{L_u(\widetilde{t}_n)+\lambda^n L_s(\widetilde{t}_n)}\right]}{g_C}$$

the consumption is that Observe that when $\zeta \to 1$, no household has incentives to make any private human capital investment because all households want to free ride on other households' investment, leading to zero human capital investment, which is precisely the case explored in Part A (when $g_b = 0$).

The existence of externality in human capital investment implies that the market equilibrium is not Pareto efficient.

Proposition 8. *With positive externality in human capital investment ($\zeta > 0$), the first best human capital investment per unskilled labor $\frac{G(t)}{L_u(t)}$ is higher than that in the Laissez affaire decentralized market equilibrium on the temporary Balanced Growth Path (BGP) for each corresponding stage of development. More specifically, denote human capital investment per unskilled labor by $\widetilde{\chi}_n^*(t) \equiv \frac{G(t)}{L_u(t)}$ when industry n and industry $n+1$ coexist with industry 0 (that is, $t \in [\widetilde{t}_n^*, \widetilde{t}_{n+1}^*)$). The following is true on the temporary BGP for the corresponding stage of development:*

$$g_C = \frac{\xi - \rho}{\sigma}$$

$$\frac{\dot{L}_s}{L_s} = \frac{\dot{L}_u}{L_u} = \frac{\dot{G}(t)}{G(t)} = g_b \tag{26}$$

$$\frac{L_s}{L_u} = \frac{\phi}{g_b}\widetilde{\chi}_n^{*\kappa(1-\zeta)+\zeta}, \tag{27}$$

where $\widetilde{\chi}_n^*$ is uniquely determined by

$$\frac{\xi}{\phi\left[\kappa\left(1-\zeta\right)+\zeta\right]}\widetilde{\chi}_n^{*(1-\kappa)(1-\zeta)} = \delta\left[\frac{a^n(a-\lambda)}{\lambda-1} - \frac{a^{n+1}-a^n}{\lambda^{n+1}-\lambda^n}\right] - \frac{(1-\kappa)}{\left[\kappa + \frac{\zeta}{(1-\zeta)}\right]}\widetilde{\chi}_n^*, \tag{28}$$

which implies the following properties:

$$\frac{\partial\widetilde{\chi}_n^*}{\partial\xi} < 0; \frac{\partial\widetilde{\chi}_n^*}{\partial\delta} > 0; \frac{\partial\widetilde{\chi}_n^*}{\partial n} > 0; \frac{\partial\widetilde{\chi}_n^*}{\partial\phi} > 0; \frac{\partial\widetilde{\chi}_n^*}{\partial\zeta} > 0; \frac{\partial\widetilde{\chi}_n^*}{\partial a} > 0 \tag{29}$$

for $\forall n \geq 1$. In the very long run, $\frac{L_s}{L_u+L_s} \to 1$, and the economy converges to one with only skilled labor and physical capital, which is characterized in Ju, Lin and Wang (2015).

Proof: See details in the appendix. **Q.E.D**

Observe that $\frac{\partial\widetilde{\chi}_n^*}{\partial\zeta} > 0$, opposite to the case in the laissesz affair decentralized market equilibrium$\frac{\partial\widetilde{\chi}_n}{\partial\zeta} < 0$. This is because the social return to human capital investment increases with the degree of positive externaly (captured by ζ), so the first best amount of investment increases with ζ. However, in the laissez affair decentralized market equilibrium, larger externality (higher ζ) implies a stronger private incentive to free ride on other people's human capital investment, leading to a lower private investment and hence underinvestment in human capital in equilibirum.

Proposition 9. *With positive externality in human capital investment ($\zeta > 0$), the first best industry dynamics is as following: For each industry $n \geq 1$, its value added exhibits a hump-shaped life-cycle dynamic pattern: it appears at time $\widetilde{t^*}_{n-1}$, its output rises for a period of $\widetilde{m}_{n-1}^*$, and reaches the peak at time $\widetilde{t^*}_n$, after which its output declines for a period of $\widetilde{m}_n^*$ and disappears after $\widetilde{t^*}_{n+1}$. During industry n's whole life span $\widetilde{m}_{n-1}^* + \widetilde{m}_n^*$, skill premium is*

$\frac{\lambda^{n-1}(a-\lambda)}{a-1}$ *when industry* n *is booming (i.e.,* when $t \in [\widetilde{t}^*_{n-1}, \widetilde{t}^*_n)$) *and the skill premium jumps to* $\frac{\lambda^n(a-\lambda)}{a-1}$ *when industry* n declines (*i.e., when* $t \in [\widetilde{t}^*_n, \widetilde{t}^*_{n+1})$), *where*

$$\widetilde{m}^*_n \approx \frac{\log \frac{g_b+\phi\lambda^{n+1}\widetilde{\chi}_n^{*\kappa(1-\zeta)+\zeta}}{g_b+\phi\lambda^n\widetilde{\chi}_{n-1}^{*\kappa(1-\zeta)+\zeta}}}{\frac{\xi-\rho}{\sigma}-g_b}, \forall \text{ sufficiently large } n>1, \tag{30}$$

where $\widetilde{\chi}_n$ *is determined by (??). Moreover,* $\widetilde{m}^*_n > \widetilde{m}_n > \widehat{m}_n$ *for all* $n > 1$*, where* $\widehat{m}_n$ *and* $\widetilde{m}_n$ are *given by (19) and (25), respectively.*

The industry life span is even longer than the laissez faire counterpart, mainly because the amount of skilled labor increases faster whereas physical capital investment becomes smaller. All the equilibrium patterns are still consistent with all the stylized facts established earlier.

5. Conclusion

In this paper, we document several stylized facts about skill premium, endowment structures (capital, skilled labor and unskilled labor), and industrial dynamics at disaggregated levels using the US and cross-country manufacturing data. Motivated by these stylized facts, we build a tractable endogenous growth model with infinite industries, which are heterogeneous in capital-skill ratios. The model predictions are qualitatively consistent with all the stylized facts. Our model shows explicitly how skill premium dynamics at the aggregate level is logically connected to the life-cycle dynamics of underlying industries at the disaggregated level, so it provides a structural micro-foundation for the aggregate models with exogenous skill-biased technological progress, which are currently the dominant framework for the analysis of skill premium in the macro literature. Our model implies that the optimal human capital investment is stage-dependent and varies with the underlying industrial structures. Moreover, human capital investment should be synchronyzed with physical capital investment in a way that best accommodates the changing demand for skill and capital from the underlying industries, the structure of which also changes endogenously over time. We highlight that the driving force for structural change at the disaggregated industry level in our model is the endogenous change in factor endowment structures (capital-skill ratios), different from the standard mechanisms in the literature of structural transformation.
For most part of the model, there is no role of government as the first welfare theorem applies. However, if we deviate from the first-best environment by introducing externalities (as shown in Section 4.3), or other relevant frictions such as financial frictions and labor market frictions, the endogenous skill premium dynamics will be presumably different and there would be scopes for discussing welfare-enhancing roles of government. For all these promising directions for future research, a prerequisite is a good understanding of the first-best benchmark model developed in this paper. Other interesting avenues for future research include introducing international trade (see Parro, 2013, Burstein and Vogel, 2017), non-competitive market structures (Klepper and Graddy,1990; Bertomeu, 2009), and/or embedding heterogeneous firms to study firm dynamics together with industry dynamics (Dinlersoz and MacDonald, 2009).

References

[1] Acemoglu, Daron. 2003. "Patterns of Skill Premia", *Review of Economic Studies* 70 (2): 199–230

[2] Autor, David, H., Lawrence F. Katz, and Melissa S. Kearney. 2006. "The Polarization of the U.S. Labor Market." *American Economic Review* 96 (2): 189-194.

[3] Bertomeu, Jeremy. 2009. "Endogenous Shakeouts", *International Journal of Industrial Organization*, Volume 27 (3):435-440

[4] Boppart, Timo. 2014. "Structural Change and the Kaldor Facts in a Growth Model with Relative Price Effects and Non-Gorman Preferences", Econometrica 47: 1231-1246.

[5] Buera, Francesco, Joseph Kaboski, Richard Rogerson and Juan Vizcaino. 2021. "Skill-biased Structural Change", *Review of Economics Studies*. forthcoming

[6] Burstein, Ariel and Jonathan Vogel, 2017. "International Trade, Technology, and the Skill Premium," *Journal of Political Economy* 125 (5): 1356-1412.

[7] Comin, Diego, Danial Lashkari, and Marti Mestieri. 2021 "Structural Change with Long-run Income and Price Effects." Econometrica 89(1): 311-374

[8] Dinlersoz, EM and G MacDonald. 2009. "The industry life-cycle of the size distribution of firms", *Review of Economic Dynamics* 12: 648-667

[9] Eicher, Theo S., and Cecilia Garcıa-Penalosa. 2001. "Inequality and growth: the dual role of human capital in development." *Journal of Development Economics* 66 (1): 173-197.

[10] He, Hui. 2012. "What drives the skill premium: Technological change or demographic variation?" *European Economic Review* 56 (8): 1546-1572.

[11] Jovanovic, Boyan and Glenn M. MacDonald, 1994. "The Life Cycle of a Competitive Industry," *Journal of Political Economy* 102(2): 322-347.

[12] Ju, Jiandong, Justin Yifu Lin, and Yong Wang. 2015. "Endowment Structures, Industrial Dynamics, and Economic Growth", *Journal of Monetary Economics*: 244-263

[13] Klepper, Steve. 1996. "Entry, exit, growth, and innovation over the product life cycle", *American Economic Review* 86(3): 562-583.

[14] ___, 1997. "Industry Life Cycles", *Industrial and Corporate Change*, Volume 6 (1): 145–182

[15] Klepper, Steven and Elizabeth Graddy. 1990. "The Evolution of New Industries and the Determinants of Market Structure", *The RAND Journal of Economics* 21(1):27-44

[16] Mastuyama, Kiminori. 2009. "Structural Change in an Interdependent World: A Global View of Manufacturing Decline," *Journal of the European Economic Association*, 7 (April-May): 478-486

[17] Kongsamut, Piyabha, Rebelo, Sergio and Danyang Xie. 2001. "Beyond Balanced Growth." *Review of Economic Studies* 68: 869-882;

[18] Krusell, Per Lee, E. Ohanian, José-Víctor Ríos-Rull, Giovanni L. Violante, 2000. "Capital-skill Complementarity and Inequality: A Macroeconomic Analysis", *Econometrica* 68 (5): 1029-1053

[19] Lin, Justin Yifu and Yong Wang. 2019. "Remodeling Structural Change", in *Oxford Handbook of Structural Transformation*, Oxford University Press.

[20] Lucas, Robert E. Jr., 1988. "On the Mechanics of Economic Development". *Journal of Monetary Economics* 22: 3-42.

[21] Mankiw, N. Gregory, David Romer and David N. Weil. 1992. "A Contribution to the Empirics of Economic Growth", *Quarterly Journal of Economics* 107(2): 407-437

[22] Mulligan, Casey B. and Xavier Sala-i-Martin. 1993. "Transitional Dynamics in Two-Sector Models of Endogenous Growth". *Quarterly Journal of Economics* 108(3): 739-773.

[23] Ngai, L. Rachel, and Christopher A. Passarides. 2007. "Structural Change in a Multi-Sector Model of Growth." *American Economic Review* 97 (January): 429-443

[24] Parro, Fernando. 2013. "Capital-Skill Complementarity and the Skill Premium in a Quantitative Model of Trade." *American Economic Journal: Macroeconomics* 5 (2): 72-117.

[25] Pavcnik, Nina. 2003. "What explains skill upgrading in less developed countries?" *Journal of Development Economics* 71(2): 311-328.

[26] Stokey, Nancy. 1996. "Free Trade, Factor Returns, and Factor Accumulation," *Journal of Economic Growth*, 1 (1996): 421-447

[27] Uy, Timothy, Kei-Mu Yi and Jing Zhang. 2013. "Structural Transformation in an Open Economy." *Journal of Monetary Economics* 60(6): 667-682

[28] Wang, Zhu. 2008. "Income distribution, market size and the evolution of industry", *Review of Economic Dynamics*, 11(3): 542-565

Mathematical Appendix

Proof for Proposition 1.

Proof.

$$\max_{\{k_n\},\{l_{un}\},\{l_{sn}\}} \sum_{n=0}^{\infty} \lambda^n F_n(k_n, l_{sn}, l_{un}),$$

subject to

$$\begin{aligned} \sum_{n=0}^{\infty} k_n &\leq K; \sum_{n=0}^{\infty} l_{un} \leq L_u; \sum_{n=0}^{\infty} l_{sn} \leq L_s; \\ k_n &\geq 0; l_{un} \geq 0; l_{sn} \geq 0 \end{aligned}$$

;where

$$x_n = F_n(k_n, l_{sn}, l_{un}) = \begin{cases} l_{un} + l_{sn}, & \text{if} \quad n = 0 \\ \frac{l_{un}}{b^n} + \min\{\frac{k_n}{a^n}, l_{sn}\}, & \text{if} \quad n \geq 1 \end{cases}.$$

First of all, we show that L_u must be fully used and only used to produce good 0. Observe that L_u must be fully used up because it can be at least used to produce x_0. Now suppose some unskilled labor is used to produce some good $n \geq 1$, that is, $l_{un} > 0$ for some $n \geq 1$, then the marginal product of unskilled labor in the production of the final good is $\frac{\lambda^n}{b^n}$, which is strictly smaller than one because $\lambda < b$, so it is a contradiction. As a result, $l_{un} = 0$ for all $n \geq 1$ in equilibrium, the production functions of industries can be rewritten as

$$x_n = F_n(k_n, l_{sn}, l_{un}) = \begin{cases} l_{un} + l_{sn}, & \text{if} \quad n = 0 \\ \min\{\frac{k_n}{a^n}, l_{sn}\}, & \text{if} \quad n \geq 1 \end{cases},$$

When $0 \leq K < aL_s$,

Suppose $x_j > 0$ and $x_{j'} > 0$. WLOG, assume $1 \leq j < j'$. We show by contradiction that it must be true that $j' = j + 1$. Suppose $j' \geq j + 2$. Let w_s denote marginal product of skilled labor, w_u denote marginal product of unskilled labor, and r denote marginal product of physical capital. Let p_n denote price of good n. Then from the previous analysis, we know that

$$x_n = F_n(k_n, l_{sn}, 0) = \min\{\frac{k_n}{a^n}, l_{sn}\}, \quad \text{if} \quad n \geq 1 ,$$

which implies

$$p_n = a^n r + w_s.$$

So when $x_j > 0$ and $x_{j'} > 0$, we must have

$$\frac{p_{j'}}{p_j} = \frac{\lambda^{j'}}{\lambda^j}$$

that is,

$$\begin{aligned} \frac{a^{j'}r + w_s}{a^j r + w_s} &= \frac{\lambda^{j'}}{\lambda^j} = \lambda^{j'-j} \\ a^{j'}r + w_s &= \left(a^j r + w_s\right)\lambda^{j'-j} \\ w_s &= \frac{a^j r(a^{j'-j} - \lambda^{j'-j})}{(\lambda^{j'-j} - 1)} \end{aligned}$$

Suppose $j' - j \geq 2$, then

$$\begin{aligned} \frac{p_{j+1}}{p_j} &= \frac{a^{j+1}r + w_s}{a^j r + w_s} = \frac{a^{j+1}r + \frac{a^j r(a^{j'-j} - \lambda^{j'-j})}{(\lambda^{j'-j}-1)}}{a^j r + \frac{a^j r(a^{j'-j} - \lambda^{j'-j})}{(\lambda^{j'-j}-1)}} \\ &= \frac{a(\lambda^{j'-j} - 1) + (a^{j'-j} - \lambda^{j'-j})}{(\lambda^{j'-j} - 1) + (a^{j'-j} - \lambda^{j'-j})} \\ &= \lambda^{j'-j}\frac{a-1}{a^{j'-j} - 1} \end{aligned}$$

$$\begin{aligned} \lambda^{j'-j}\frac{a-1}{a^{j'-j} - 1} &> \lambda \\ \lambda^{j'-j-1} &> \frac{a^{j'-j} - 1}{a-1} \end{aligned}$$

Proof for Proposition 2.

When $t \in [t_n^e, t_{n+1}^b]$, industry n declines over time and industry $n+1$ rises over time, we know that during this period $C(t) \in [\lambda^n L_s + L_u, \lambda^{n+1} L_s + L_u]$ for any $n \geq 1$, and we have the following discounted-value Hamiltonian:

$$H_{0,n,n+1} = \frac{C(t)^{1-\sigma} - 1}{1-\sigma} e^{-\rho t} + \widehat{\lambda}_{n,n+1}\left[\xi K - \delta \cdot E_{0,n,n+1}(C(t))\right] \tag{31}$$

First order conditions

$$\frac{\partial H_{0,n,n+1}}{\partial C} = 0 \Rightarrow C(t)^{-\sigma} e^{-\rho t} = \widehat{\lambda}_{n,n+1}\delta\frac{a^{n+1} - a^n}{\lambda^{n+1} - \lambda^n} \tag{32}$$

$$\Rightarrow -\sigma\frac{\dot{C}}{C} - \rho = \frac{\dot{\widehat{\lambda}}_{n,n+1}}{\widehat{\lambda}_{n,n+1}} \tag{33}$$

$$\dot{\widehat{\lambda}}_{n,n+1} = -\frac{\partial H_{0,n,n+1}}{\partial K} \Rightarrow \frac{\dot{\widehat{\lambda}}_{n,n+1}}{\widehat{\lambda}_{n,n+1}} = -\xi \tag{34}$$

which jointly imply

$$\frac{\dot{C}}{C} = \frac{\xi - \rho}{\sigma}.$$

In particular, when $t = t^b_{n+1}$, we should have

$$C(t^b_{n+1})^{-\sigma} e^{-\rho t^b_{n+1}} = \widehat{\lambda}_{n,n+1}(t^b_{n+1}) \delta \frac{a^{n+1} - a^n}{\lambda^{n+1} - \lambda^n}. \tag{35}$$

When $t = t^e_{n+1}$, we should have

$$C(t^e_{n+1})^{-\sigma} e^{-\rho t^e_{n+1}} = \widehat{\lambda}_{n+1,n+2}(t^e_{n+1}) \delta \frac{a^{n+2} - a^{n+1}}{\lambda^{n+2} - \lambda^{n+1}}. \tag{36}$$

Now we show by contradiction that $t^b_{n+1} < t^e_{n+1}$. Suppose $t^b_{n+1} = t^e_{n+1} = t_{n+1}$, the optimality conditions require

$$H_{0,n,n+1}(t_{n+1}) = H_{0,n+1,n+2}(t_{n+1}),$$

and

$$\widehat{\lambda}_{n,n+1}(t_{n+1}) = \widehat{\lambda}_{n+1,n+2}(t_{n+1}).$$

Substituting the second equation into (35) and (36), and revoking the fact that

$$C(t^b_{n+1}) = C(t^e_{n+1}) = \lambda^{n+1} L_s + L_u, \tag{37}$$

we obtain

$$\frac{a^{n+1} - a^n}{\lambda^{n+1} - \lambda^n} = \frac{a^{n+2} - a^{n+1}}{\lambda^{n+2} - \lambda^{n+1}},$$

a contradiction. Therefore, we prove that $t^b_{n+1} < t^e_{n+1}$. We know that $C(t) = \lambda^{n+1} L_s + L_u$ for any $t \in [t^b_{n+1}, t^e_{n+1}]$. Now we pin down this duration $t^e_{n+1} - t^b_{n+1}$. When $t \in [t^b_{n+1}, t^e_{n+1}]$, $E_{0,n+1}(C(t)) = \lambda^{n+1} L_s$, and we have the following discounted-value Hamiltonian:

$$H_{0,n+1} = \frac{C(t)^{1-\sigma} - 1}{1 - \sigma} e^{-\rho t} + \widehat{\lambda}_{n+1} \left[\xi K - \delta \cdot a^{n+1} L_s\right] + \eta_{n+1}(\lambda^{n+1} L_s + L_u - C) \tag{38}$$

$$\begin{aligned} \frac{\partial H_{0,n+1}}{\partial C} &= 0 \Rightarrow C(t)^{-\sigma} e^{-\rho t} = \eta_{n+1} \\ -\dot{\widehat{\lambda}}_{n+1} &= \frac{\partial H_{0,n+1}}{\partial K} = \xi \widehat{\lambda}_{n+1} \end{aligned}$$

thus we have

$$\widehat{\lambda}_{n+1}(t) = \widehat{\lambda}_{n+1}(t^b_{n+1}) e^{-\xi(t - t^b_{n+1})}$$

for any $t \in [t^b_{n+1}, t^e_{n+1}]$. In particular,

$$\widehat{\lambda}_{n+1}(t^e_{n+1}) = \widehat{\lambda}_{n+1}(t^b_{n+1}) e^{-\xi(t^e_{n+1} - t^b_{n+1})},$$

Note that the optimality condition requires

$$\begin{aligned} \widehat{\lambda}_{n,n+1}(t^b_{n+1}) &= \widehat{\lambda}_{n+1}(t^b_{n+1}); \\ \widehat{\lambda}_{n+1}(t^e_{n+1}) &= \widehat{\lambda}_{n+1,n+2}(t^e_{n+1}), \end{aligned}$$

therefore, the previous three equations jointly imply that

$$\widehat{\lambda}_{n+1,n+2}(t^e_{n+1}) = \widehat{\lambda}_{n,n+1}(t^b_{n+1})e^{-\xi(t^e_{n+1}-t^b_{n+1})}$$

which, together with (35), (36) and (37), yields

$$t^e_{n+1} - t^b_{n+1} = \frac{\ln \frac{a}{\lambda}}{\xi - \rho}.$$

Thus the whole life cycle of industry $n+1$ consists of a rising period (t^e_n, t^b_{n+1}), a peak period $[t^b_{n+1}, t^e_{n+1}]$, and a decline period $(t^e_{n+1}, t^b_{n+2}]$. Its life span is equal to $t^b_{n+2} - t^e_n$. It is easy to see that

$$t^b_{n+1} - t^e_n = \frac{\ln \frac{\lambda^{n+1}L_s+L_u}{\lambda^n L_s+L_u}}{\frac{\xi-\rho}{\sigma}}; \forall n \geq 1.$$

Proof for Proposition 5.

Proof. Substituting (11) into (12) yields

$$\dot{L}_u = g_b \cdot (L_u + L_s) - \phi L_u^{1-\kappa} \cdot G(t)^{\kappa}.$$

Suppose $C(t) \in (\lambda^n L_s + L_u, \lambda^{n+1}L_s + L_u)$ for any $n \geq 1$. Establish the discounted-value Hamiltonian as follows:

$$\begin{aligned} H &= \frac{C(t)^{1-\sigma}-1}{1-\sigma}e^{-\rho t} + \widehat{\lambda}_{n,n+1}\left[\xi K - \delta \cdot E(C(t)) - G(t)\right] + \eta\left[g_b \cdot (L_u + L_s) - \phi L_u^{1-\kappa} \cdot G(t)^{\kappa}\right] + \psi\phi L_u^{1-\kappa} \cdot \\ &= \frac{C(t)^{1-\sigma}-1}{1-\sigma}e^{-\rho t} + \widehat{\lambda}_{n,n+1}\left\{\xi K - \delta \cdot \left[\frac{a^{n+1}-a^n}{\lambda^{n+1}-\lambda^n}C - \left(\frac{a^{n+1}-a^n}{\lambda^{n+1}-\lambda^n}L_u + \frac{a^n(a-\lambda)}{\lambda-1}L_s\right)\right] - G(t)\right\} \\ &\quad +\eta g_b \cdot (L_u + L_s) + (\psi - \eta)\,\phi L_u^{1-\kappa} \cdot G(t)^{\kappa} \end{aligned}$$

First order conditions:

$$\frac{\partial H}{\partial C} = 0 \Rightarrow C(t)^{-\sigma}e^{-\rho t} = \widehat{\lambda}_{n,n+1}\delta\frac{a^{n+1}-a^n}{\lambda^{n+1}-\lambda^n} \Rightarrow -\sigma\frac{\dot{C}}{C} - \rho = \frac{\dot{\widehat{\lambda}}_{n,n+1}}{\widehat{\lambda}_{n,n+1}} \quad (39)$$

$$\frac{\partial H}{\partial G} = 0 \Rightarrow [\psi - \eta]\,\kappa\phi L_u^{1-\kappa} \cdot G(t)^{\kappa-1} = \widehat{\lambda}_{n,n+1} \quad (40)$$

$$\dot{\widehat{\lambda}}_{n,n+1} = -\frac{\partial H}{\partial K} \Rightarrow \frac{\dot{\widehat{\lambda}}_{n,n+1}}{\widehat{\lambda}_{n,n+1}} = -\xi \quad (41)$$

$$\dot{\eta} = -\frac{\partial H}{\partial L_u} \Rightarrow \frac{\dot{\eta}}{\eta} = -g_b - \delta\frac{\widehat{\lambda}_{n,n+1}}{\eta}\frac{a^{n+1}-a^n}{\lambda^{n+1}-\lambda^n} + \left(1 - \frac{\psi}{\eta}\right)\phi\,(1-\kappa)\,L_u^{-\kappa} \cdot G(t)^{\kappa} \quad (42)$$

$$\dot{\psi} = -\frac{\partial H}{\partial L_s} \Rightarrow \frac{\dot{\psi}}{\psi} = -\left[\delta\frac{\widehat{\lambda}_{n,n+1}}{\psi}\frac{a^n(a-\lambda)}{\lambda-1} + \frac{\eta}{\psi}g_b\right] \quad (43)$$

(39) and (41) jointly imply

$$\frac{\dot{C}}{C} = \frac{\xi - \rho}{\sigma}.$$

$\frac{\dot{\eta}}{\eta}$ is a constant when $\frac{\dot{L_u}}{L_u} = \frac{\dot{G}}{G}$ and $\frac{\dot{\eta}}{\eta} = \frac{\dot{\widehat{\lambda}}_{n,n+1}}{\widehat{\lambda}_{n,n+1}} = \frac{\dot{\psi}}{\psi}$ both hold, in which case

$$\frac{\dot{\eta}}{\eta} = -g_b - \delta \frac{\widehat{\lambda}_{n,n+1}}{\eta} \frac{a^{n+1} - a^n}{\lambda^{n+1} - \lambda^n} + \left(1 - \frac{\psi}{\eta}\right) \phi \left(1 - \kappa\right) L_u^{-\kappa} \cdot G(t)^{\kappa} = \rho - \xi,$$

so that

$$L_u^{-\kappa} \cdot G(t)^{\kappa} = -\frac{\xi - g_b - \delta \frac{\widehat{\lambda}_{n,n+1}}{\eta} \frac{a^{n+1} - a^n}{\lambda^{n+1} - \lambda^n}}{\left(1 - \frac{\psi}{\eta}\right) \phi \left(1 - \kappa\right)}. \tag{44}$$

By (11), $\frac{\dot{L_s}}{L_s}$ is constant if and only if $\frac{\dot{L_s}}{L_s} = \frac{\dot{L_u}}{L_u} = \frac{\dot{G}}{G}$. Let $g_G \equiv \frac{\dot{G}}{G}$. (12) implies

$$\begin{aligned} \frac{\dot{L_u}}{L_u} &= g_b \cdot \left(1 + \frac{L_s}{L_u}\right) - \frac{L_s}{L_u} \frac{\dot{L_s}}{L_s} \\ g_G &= g_b + \left(g_b - g_G\right) \frac{L_s}{L_u}, \end{aligned}$$

which can be true if and only if $g_G = g_b$ because $\frac{L_s}{L_u} > 0$. Thus

$$\frac{\dot{L_s}}{L_s} = \frac{\dot{L_u}}{L_u} = \frac{\dot{G}}{G} = g_b.$$

(40) can be rewritten as

$$\frac{\widehat{\lambda}_{n,n+1}}{\eta} = -\phi L_u^{1-\kappa} \cdot \kappa G(t)^{\kappa-1} \left(1 - \frac{\psi}{\eta}\right). \tag{45}$$

Substitute out $\frac{\widehat{\lambda}_{n,n+1}}{\eta}$ in (45) into equation (44), we obtain

$$\left(1 - \frac{\psi}{\eta}\right) \left(1 - \kappa\right) \phi \left[\frac{G(t)}{L_u(t)}\right]^{\kappa} = g_b - \xi - \left(1 - \frac{\psi}{\eta}\right) \phi \kappa \delta \frac{a^{n+1} - a^n}{\lambda^{n+1} - \lambda^n} \left[\frac{G(t)}{L_u(t)}\right]^{\kappa-1}. \tag{46}$$

Since $\frac{\dot{\eta}}{\eta} = \frac{\dot{\widehat{\lambda}}_{n,n+1}}{\widehat{\lambda}_{n,n+1}} = \frac{\dot{\psi}}{\psi}$, (43) and (41) jointly imply

$$\delta \frac{\widehat{\lambda}_{n,n+1}}{\eta} \frac{a^n(a - \lambda)}{\lambda - 1} + g_b = \xi \frac{\psi}{\eta} \tag{47}$$

so we have three unknowns $\frac{\widehat{\lambda}_{n,n+1}}{\eta}, \frac{\psi}{\eta}, \frac{G(t)}{L_u(t)}$ and three equations (45)-(47).

Using the brutal force, we obtain (17), that is,

$$(1-\kappa)\,\phi\left[\frac{G(t)}{L_u(t)}\right]^{\kappa}=\frac{a^n}{\lambda-1}\left\{(a-\lambda)-\frac{a-1}{\lambda^n}\right\}\phi\kappa\delta\left[\frac{G(t)}{L_u(t)}\right]^{\kappa-1}-\xi,$$

which uniquely determines $\frac{G(t)}{L_u(t)}$. Denote this solution by χ_n. By (11), we have

$$\frac{L_s}{L_u}=\frac{\phi\chi_n^{\kappa}}{g_b},$$

and

$$\begin{aligned}\frac{\psi}{\eta} &= \frac{g_b-\phi L_u^{1-\kappa}\cdot\kappa G(t)^{\kappa-1}\delta\frac{a^n(a-\lambda)}{\lambda-1}}{\xi-\phi L_u^{1-\kappa}\cdot\kappa G(t)^{\kappa-1}\delta\frac{a^n(a-\lambda)}{\lambda-1}}\\ &= 1+\frac{\xi-g_b}{(1-\kappa)\,\phi\chi_n^{\kappa}+\frac{a^n}{\lambda-1}\frac{a-1}{\lambda^n}\phi\kappa\delta\chi_n^{\kappa-1}}\end{aligned}$$

$$\begin{aligned}\frac{\widehat{\lambda}_{n,n+1}}{\eta} &= \frac{\xi\frac{\psi}{\eta}-g_b}{\delta\frac{a^n(a-\lambda)}{\lambda-1}}=\frac{\xi\left[1-\frac{g_b-\xi}{(1-\kappa)\phi\chi_n^{\kappa}+\frac{a^n}{\lambda-1}\frac{a-1}{\lambda^n}\phi\kappa\delta\chi_n^{\kappa-1}}\right]-g_b}{\delta\frac{a^n(a-\lambda)}{\lambda-1}}\\ &= (\xi-g_b)\,(\lambda-1)\,\frac{1+\frac{\xi}{(1-\kappa)\phi\chi_n^{\kappa}+\frac{a^n}{\lambda-1}\frac{a-1}{\lambda^n}\phi\kappa\delta\chi_n^{\kappa-1}}}{\delta a^n(a-\lambda)}\end{aligned}$$

$$(1-\kappa)\,\phi\left[\frac{G(t)}{L_u(t)}\right]^{\kappa}+\frac{a^n}{\lambda-1}\frac{a-1}{\lambda^n}\phi\kappa\delta\left[\frac{G(t)}{L_u(t)}\right]^{\kappa-1}=\frac{a^n(a-\lambda)}{\lambda-1}\phi\kappa\delta\left[\frac{G(t)}{L_u(t)}\right]^{\kappa-1}-\xi.$$

$$\begin{aligned}\widehat{m}_n &= \frac{\log\frac{L_u(\widehat{t}^b_{n+1})+\lambda^{n+1}L_s(\widehat{t}^b_{n+1})}{L_u(\widehat{t}^e_n)+\lambda^nL_s(\widehat{t}^e_n)}}{g_C}=\frac{\log\frac{L_u(\widehat{t}^b_{n+1})}{L_u(\widehat{t}^e_n)}\frac{1+\lambda^{n+1}\frac{L_s(\widehat{t}^b_{n+1})}{L_u(\widehat{t}^b_{n+1})}}{1+\lambda^n\frac{L_s(\widehat{t}^e_n)}{L_u(\widehat{t}^e_n)}}}{g_C}\\ &= \frac{\log e^{g_b(\widehat{t}^b_{n+1}-\widehat{t}^e_n)}\frac{1+\lambda^{n+1}\frac{\phi\chi^{\kappa}_{n+1}}{g_b}}{1+\lambda^n\frac{\phi\chi^{\kappa}_n}{g_b}}}{g_C}=\frac{g_b\widehat{m}_n+\log\frac{1+\lambda^{n+1}\frac{\phi\chi^{\kappa}_{n+1}}{g_b}}{1+\lambda^n\frac{\phi\chi^{\kappa}_n}{g_b}}}{g_C},\end{aligned}$$

which implies

$$\widehat{m}_n=\frac{\log\frac{1+\lambda^{n+1}\frac{\phi\chi^{\kappa}_{n+1}}{g_b}}{1+\lambda^n\frac{\phi\chi^{\kappa}_n}{g_b}}}{g_C-g_b}=\frac{\log\left[1+\frac{\lambda^{n+1}\frac{\phi\chi^{\kappa}_{n+1}}{g_b}-\lambda^n\frac{\phi\chi^{\kappa}_n}{g_b}}{1+\lambda^n\frac{\phi\chi^{\kappa}_n}{g_b}}\right]}{g_C-g_b}=\frac{\log\left[1+\frac{\frac{\lambda\chi^{\kappa}_{n+1}}{\chi^{\kappa}_n}-1}{\frac{1}{\lambda^n\frac{\phi\chi^{\kappa}_n}{g_b}}+1}\right]}{g_C-g_b}.$$

Now we prove that $\widehat{t}^b_{n+1}<\widehat{t}^e_{n+1}$. In particular, when $t=\widehat{t}^b_{n+1}$, by (39), we should have

$$C(\widehat{t}^b_{n+1})^{-\sigma}e^{-\rho\widehat{t}^b_{n+1}}=\widehat{\lambda}_{n,n+1}(\widehat{t}^b_{n+1})\delta\frac{a^{n+1}-a^n}{\lambda^{n+1}-\lambda^n}. \tag{48}$$

When $t = \hat{t}^e_{n+1}$, we should have

$$C(\hat{t}^e_{n+1})^{-\sigma} e^{-\rho \hat{t}^e_{n+1}} = \hat{\lambda}_{n+1,n+2}(\hat{t}^e_{n+1}) \delta \frac{a^{n+2} - a^{n+1}}{\lambda^{n+2} - \lambda^{n+1}}. \tag{49}$$

Suppose $\hat{t}^b_{n+1} = \hat{t}^e_{n+1} = \hat{t}_{n+1}$, the optimality conditions require

$$H_{0,n,n+1}(\hat{t}_{n+1}) = H_{0,n+1,n+2}(\hat{t}_{n+1}),$$

and

$$\hat{\lambda}_{n,n+1}(\hat{t}_{n+1}) = \hat{\lambda}_{n+1,n+2}(\hat{t}_{n+1}).$$

Substituting the second equation into (48) and (49), and revoking the fact that

$$C(\hat{t}^b_{n+1}) = C(\hat{t}^e_{n+1}) = \lambda^{n+1} L_s\left(\hat{t}_{n+1}\right) + L_u(\hat{t}_{n+1}),$$

we obtain

$$\frac{a^{n+1} - a^n}{\lambda^{n+1} - \lambda^n} = \frac{a^{n+2} - a^{n+1}}{\lambda^{n+2} - \lambda^{n+1}},$$

a contradiction. Therefore, we prove that $\hat{t}^b_{n+1} < \hat{t}^e_{n+1}$. We know that $C(t) = \lambda^{n+1} L_s(t) + L_u(t)$ for any $t \in [\hat{t}^b_{n+1}, \hat{t}^e_{n+1}]$. Now we pin down this duration $\hat{t}^e_{n+1} - \hat{t}^b_{n+1}$. When $t \in [\hat{t}^b_{n+1}, \hat{t}^e_{n+1}]$, $E_{0,n+1}(C(t)) = a^{n+1} L_s(t)$, and we have the following discounted-value Hamiltonian:

$$\begin{aligned} H_{0,n+1} &= \frac{C(t)^{1-\sigma} - 1}{1-\sigma} e^{-\rho t} + \hat{\lambda}_{n+1}\left[\xi K - \delta \cdot a^{n+1} L_s\right] + \zeta_{n+1}(\lambda^{n+1} L_s + L_u - C) \\ &\quad + \eta_{n+1}\left[g_b \cdot (L_u + L_s) - \phi L_u^{1-\kappa} \cdot G(t)^\kappa\right] + \psi_{n+1} \phi L_u^{1-\kappa} \cdot G(t)^\kappa \end{aligned} \tag{50}$$

$$\begin{aligned} \frac{\partial H_{0,n+1}}{\partial C} &= 0 \Rightarrow C(t)^{-\sigma} e^{-\rho t} = \zeta_{n+1} \\ -\dot{\hat{\lambda}}_{n+1} &= \frac{\partial H_{0,n+1}}{\partial K} = \xi \hat{\lambda}_{n+1} \end{aligned}$$

thus we have

$$\hat{\lambda}_{n+1}(t) = \hat{\lambda}_{n+1}(\hat{t}^b_{n+1}) e^{-\xi(t - \hat{t}^b_{n+1})}$$

for any $t \in [\hat{t}^b_{n+1}, \hat{t}^e_{n+1}]$. In particular,

$$\hat{\lambda}_{n+1}(\hat{t}^e_{n+1}) = \hat{\lambda}_{n+1}(\hat{t}^b_{n+1}) e^{-\xi(\hat{t}^e_{n+1} - \hat{t}^b_{n+1})},$$

Note that the optimality condition requires

$$\begin{aligned} \hat{\lambda}_{n,n+1}(\hat{t}^b_{n+1}) &= \hat{\lambda}_{n+1}(\hat{t}^b_{n+1}); \\ \hat{\lambda}_{n+1}(\hat{t}^e_{n+1}) &= \hat{\lambda}_{n+1,n+2}(\hat{t}^e_{n+1}), \end{aligned}$$

therefore, the previous three equations jointly imply that

$$\hat{\lambda}_{n+1,n+2}(\hat{t}^e_{n+1}) = \hat{\lambda}_{n,n+1}(\hat{t}^b_{n+1}) e^{-\xi(\hat{t}^e_{n+1} - \hat{t}^b_{n+1})}$$

which, together with (48), (49), we obtain

$$\begin{aligned}\frac{C(\widehat{t}^e_{n+1})^{-\sigma}}{C(\widehat{t}^b_{n+1})^{-\sigma}} &= \frac{a}{\lambda} e^{(\rho-\xi)(\widehat{t}^e_{n+1}-\widehat{t}^b_{n+1})} \\ -\sigma \ln \frac{C(\widehat{t}^e_{n+1})}{C(\widehat{t}^b_{n+1})} &= \ln \frac{a}{\lambda} + (\rho-\xi)\left(\widehat{t}^e_{n+1}-\widehat{t}^b_{n+1}\right)\end{aligned}$$

=====

$$C(\widehat{t}^b_{n+1})^{-\sigma} e^{-\rho \widehat{t}^b_{n+1}} = \widehat{\lambda}_{n+1}(\widehat{t}^b_{n+1}) \delta \frac{a^{n+1}-a^n}{\lambda^{n+1}-\lambda^n}. \tag{51}$$

When $t = \widehat{t}^e_{n+1}$, we should have

$$C(\widehat{t}^e_{n+1})^{-\sigma} e^{-\rho \widehat{t}^e_{n+1}} = \widehat{\lambda}_{n+1}(\widehat{t}^e_{n+1}) \delta \frac{a^{n+2}-a^{n+1}}{\lambda^{n+2}-\lambda^{n+1}}. \tag{52}$$

=====

When $t \in [\widehat{t}^b_{n+1}, \widehat{t}^e_{n+1}]$, we have $C(t) = \lambda^{n+1} L_s(t) + L_u(t)$. The corresponding discounted-value Hamiltonian is given by

$$\begin{aligned} H_{0,n+1} &= \frac{C(t)^{1-\sigma}-1}{1-\sigma} e^{-\rho t} + \widehat{\lambda}_{n+1}\left[\xi K - \delta \cdot a^{n+1} L_s\right] + \zeta_{n+1}(\lambda^{n+1} L_s + L_u - C) \\ &\quad + \eta_{n+1}\left[g_b \cdot (L_u + L_s) - \phi L_u^{1-\kappa} \cdot G(t)^\kappa\right] + \psi_{n+1} \phi L_u^{1-\kappa} \cdot G(t)^\kappa \end{aligned} \tag{53}$$

The first order conditions are

$$\frac{\partial H}{\partial C} = 0 \Rightarrow C(t)^{-\sigma} e^{-\rho t} = \zeta_{n+1} \Rightarrow -\sigma \frac{\dot{C}}{C} - \rho = \frac{\dot{\zeta}_{n+1}}{\zeta_{n+1}} \tag{54}$$

$$\frac{\partial H}{\partial G} = 0 \Rightarrow \left[\psi_{n+1} - \eta_{n+1}\right] \kappa \phi L_u^{1-\kappa} \cdot G(t)^{\kappa-1} = \widehat{\lambda}_{n+1} \tag{55}$$

$$\dot{\widehat{\lambda}}_{n+1} = -\frac{\partial H}{\partial K} \Rightarrow \frac{\dot{\widehat{\lambda}}_{n+1}}{\widehat{\lambda}_{n+1}} = -\xi \tag{56}$$

$$\dot{\eta}_{n+1} = -\frac{\partial H}{\partial L_u} = -\left[\zeta_{n+1} + \eta_{n+1}\left(g_b - (1-\kappa)\phi L_u^{-\kappa} \cdot G(t)^\kappa\right) + \psi_{n+1}(1-\kappa)\phi L_u^{-\kappa} \cdot G(t)^\kappa\right] \tag{57}$$

$$\dot{\psi}_{n+1} = -\frac{\partial H}{\partial L_s} = -\left[-\widehat{\lambda}_{n+1}\delta \cdot a^{n+1} + \zeta_{n+1}\lambda^{n+1} + \eta_{n+1} g_b\right] \tag{58}$$

(57) implies

$$\begin{aligned}\dot{\eta}_{n+1} &= -\frac{\partial H}{\partial L_u} = -\left[\zeta_{n+1} + \eta_{n+1}\left(g_b - (1-\kappa)\phi L_u^{-\kappa} \cdot G(t)^\kappa\right) + \psi_{n+1}(1-\kappa)\phi L_u^{-\kappa} \cdot G(t)^\kappa\right] \\ \dot{\eta}_{n+1} &= -\left[\zeta_{n+1} + \eta_{n+1} g_b + \left(\psi_{n+1} - \eta_{n+1}\right)(1-\kappa)\phi L_u^{-\kappa} \cdot G(t)^\kappa\right]\end{aligned}$$

which implies

$$\frac{\dot{\eta}_{n+1}}{\eta_{n+1}} = -\left[\frac{\zeta_{n+1}}{\eta_{n+1}} + g_b + \left(\frac{\psi_{n+1}}{\eta_{n+1}} - 1\right)(1-\kappa)\phi L_u^{-\kappa} \cdot G(t)^\kappa\right]$$

or

$$\begin{aligned}
\xi &= \frac{\zeta_{n+1}}{\eta_{n+1}} + g_b + \left(\frac{\psi_{n+1}}{\eta_{n+1}} - 1\right)(1-\kappa)\,\phi L_u^{-\kappa} \cdot G(t)^{\kappa} \\
\frac{\eta_{n+1}}{\psi_{n+1}}\xi &= \frac{\zeta_{n+1}}{\psi_{n+1}} + \frac{\eta_{n+1}}{\psi_{n+1}} g_b + \left(1 - \frac{\eta_{n+1}}{\psi_{n+1}}\right)(1-\kappa)\,\phi \widehat{\chi}_{n+1}^{\kappa}
\end{aligned}$$

(58) implies

$$\begin{aligned}
\frac{\dot{\psi}_{n+1}}{\psi_{n+1}} &= -\left[-\frac{\widehat{\lambda}_{n+1}}{\psi_{n+1}}\delta \cdot a^{n+1} + \frac{\zeta_{n+1}}{\psi_{n+1}}\lambda^{n+1} + \frac{\eta_{n+1}}{\psi_{n+1}} g_b\right] \\
\xi &= -\frac{\widehat{\lambda}_{n+1}}{\psi_{n+1}}\delta \cdot a^{n+1} + \frac{\zeta_{n+1}}{\psi_{n+1}}\lambda^{n+1} + \frac{\eta_{n+1}}{\psi_{n+1}} g_b
\end{aligned}$$

On the BGP, we have

$$\frac{\dot{\psi}_{n+1}}{\psi_{n+1}} = \frac{\dot{\widehat{\lambda}}_{n+1}}{\widehat{\lambda}_{n+1}} = \frac{\dot{\eta}_{n+1}}{\eta_{n+1}} = -\xi$$

so

$$-\sigma\frac{\dot{C}}{C} - \rho = \frac{\dot{\zeta}_{n+1}}{\zeta_{n+1}}$$

Since

$$\begin{aligned}
\frac{\dot{L}_u}{L_u} &= g_b \cdot \left(1 + \frac{L_s}{L_u}\right) - \frac{L_s}{L_u}\frac{\dot{L}_s}{L_s} \\
g_G &= g_b + (g_b - g_G)\frac{L_s}{L_u},
\end{aligned}$$

we must have

$$\frac{\dot{L}_u}{L_u} = \frac{\dot{G}}{G} = \frac{\dot{L}_s}{L_s} = g_b$$

so

$$\frac{\dot{C}}{C} = g_b$$

and

$$\frac{\dot{\zeta}_{n+1}}{\zeta_{n+1}} = -\sigma\frac{\dot{C}}{C} - \rho = -\sigma g_b - \rho$$

(13) implies that

$$-\sigma g_b - \rho > -\xi$$

(55) implies

$$\begin{aligned}
\left[\frac{\psi_{n+1}}{\widehat{\lambda}_{n+1}} - \frac{\eta_{n+1}}{\widehat{\lambda}_{n+1}}\right]\kappa\phi L_u^{1-\kappa} \cdot G(t)^{\kappa-1} &= 1 \\
\left[1 - \frac{\eta_{n+1}}{\psi_{n+1}}\right]\kappa\phi &= \frac{\widehat{\lambda}_{n+1}}{\psi_{n+1}}\widehat{\chi}_{n+1}^{1-\kappa}
\end{aligned}$$

Define $\frac{G(t)}{L_u(t)} = \widehat{\chi}_{n+1}$, we have

$$\frac{\dot{\psi}_{n+1}}{\psi_{n+1}} = \frac{\dot{\widehat{\lambda}}_{n+1}}{\widehat{\lambda}_{n+1}}$$

$$\dot{L}_s = \phi L_u^{1-\kappa} \cdot G(t)^\kappa$$

$$\lambda^{n+1} L_s + L_u = C$$

$$\dot{L}_u = g_b \cdot (L_u + L_s) - \phi L_u^{1-\kappa} \cdot G(t)^\kappa.$$

so

$$\begin{aligned} \dot{C} &= \lambda^{n+1} \dot{L}_s + \dot{L}_u \\ &= \left(\lambda^{n+1} - 1\right) \phi L_u^{1-\kappa} \cdot G(t)^\kappa + g_b \cdot (L_u + L_s) \end{aligned}$$

Thus the whole life cycle of industry $n+1$ consists of a rising period (t_n^e, t_{n+1}^b), a peak period $[t_{n+1}^b, t_{n+1}^e]$, and a decline period $(t_{n+1}^e, t_{n+2}^b]$. Its life span is equal to $t_{n+2}^b - t_n^e$. It is easy to see that

$$t_{n+1}^b - t_n^e = \frac{\ln \frac{\lambda^{n+1} L_s + L_u}{\lambda^n L_s + L_u}}{\frac{\xi - \rho}{\sigma}}; \forall n \geq 1.$$

Q.E.D.

alternative proof for prop 5

Proof. Substituting (11) into (12) yields

$$\dot{L}_u = g_b \cdot (L_u + L_s) - \phi L_u^{1-\kappa} \cdot G(t)^\kappa.$$

Suppose $C(t) \in (\lambda^n L_s + L_u, \lambda^{n+1} L_s + L_u)$ for any $n \geq 1$. Establish the current-value Hamiltonian as follows:

$$\begin{aligned} H &= \frac{C(t)^{1-\sigma} - 1}{1-\sigma} + \widehat{\lambda} \left[\xi K - \delta \cdot E(C(t)) - G(t)\right] + \eta \left[g_b \cdot (L_u + L_s) - \phi L_u^{1-\kappa} \cdot G(t)^\kappa\right] + \psi \phi L_u^{1-\kappa} \cdot G(t)^\kappa \\ &= \frac{C(t)^{1-\sigma} - 1}{1-\sigma} + \widehat{\lambda} \left\{\xi K - \delta \cdot \left[\frac{a^{n+1} - a^n}{\lambda^{n+1} - \lambda^n} C - \left(\frac{a^{n+1} - a^n}{\lambda^{n+1} - \lambda^n} L_u + \frac{a^n (a - \lambda)}{\lambda - 1} L_s\right)\right] - G(t)\right\} \\ &\quad + \eta g_b \cdot (L_u + L_s) + (\psi - \eta) \phi L_u^{1-\kappa} \cdot G(t)^\kappa \end{aligned}$$

First order conditions:

$$\frac{\partial H}{\partial C} = 0 \Rightarrow C(t)^{-\sigma} = \widehat{\lambda}\delta \frac{a^{n+1} - a^n}{\lambda^{n+1} - \lambda^n} \Rightarrow -\sigma \frac{\dot{C}}{C} = \frac{\dot{\widehat{\lambda}}}{\widehat{\lambda}} \tag{59}$$

$$\frac{\partial H}{\partial G} = 0 \Rightarrow [\psi - \eta] \, \kappa \phi L_u^{1-\kappa} \cdot G(t)^{\kappa-1} = \widehat{\lambda} \tag{60}$$

$$\dot{\widehat{\lambda}} = \rho \widehat{\lambda} - \frac{\partial H}{\partial K} \Rightarrow \frac{\dot{\widehat{\lambda}}}{\widehat{\lambda}} = \rho - \xi \tag{61}$$

$$\dot{\eta} = \rho \eta - \frac{\partial H}{\partial L_u}$$

$$\Rightarrow \frac{\dot{\eta}}{\eta} = \rho - g_b - \delta \frac{\widehat{\lambda}}{\eta} \frac{a^{n+1} - a^n}{\lambda^{n+1} - \lambda^n} + \left(1 - \frac{\psi}{\eta}\right) \phi \left(1 - \kappa\right) L_u^{-\kappa} \cdot G(t)^{\kappa} \tag{62}$$

$$\frac{\partial H}{\partial L_s} = \delta \lambda \frac{a^n (a - \lambda)}{\lambda - 1} + \eta g_b$$

$$\dot{\psi} = \rho \psi - \frac{\partial H}{\partial L_s} = \rho \psi - \left[\delta \widehat{\lambda} \frac{a^n (a - \lambda)}{\lambda - 1} + \eta g_b\right] \Rightarrow$$

$$\frac{\dot{\psi}}{\psi} = \rho - \left[\delta \frac{\widehat{\lambda}}{\psi} \frac{a^n (a - \lambda)}{\lambda - 1} + \frac{\eta}{\psi} g_b\right] \tag{63}$$

thus

$$\frac{\dot{C}}{C} = \frac{\xi - \rho}{\sigma}$$

and $\frac{\dot{\eta}}{\eta}$ is a constant when $\frac{\dot{L_u}}{L_u} = \frac{\dot{G}}{G}$ and $\frac{\dot{\eta}}{\eta} = \frac{\dot{\widehat{\lambda}}}{\widehat{\lambda}} = \frac{\dot{\psi}}{\psi}$ both hold, in which case

$$\frac{\dot{\eta}}{\eta} = \rho - g_b - \delta \frac{\widehat{\lambda}}{\eta} \frac{a^{n+1} - a^n}{\lambda^{n+1} - \lambda^n} + \left(1 - \frac{\psi}{\eta}\right) \phi \left(1 - \kappa\right) L_u^{-\kappa} \cdot G(t)^{\kappa} = \rho - \xi,$$

which implies

$$L_u^{-\kappa} \cdot G(t)^{\kappa} = -\frac{\xi - g_b - \delta \frac{\widehat{\lambda}}{\eta} \frac{a^{n+1} - a^n}{\lambda^{n+1} - \lambda^n}}{\left(1 - \frac{\psi}{\eta}\right) \phi \left(1 - \kappa\right)}. \tag{64}$$

By (11), $\frac{\dot{L_s}}{L_s}$ is constant if and only if $\frac{\dot{L_s}}{L_s} = \frac{\dot{L_u}}{L_u} = \frac{\dot{G}}{G}$. Let $g_G \equiv \frac{\dot{G}}{G}$. (12) implies

$$\frac{\dot{L_u}}{L_u} = g_b \cdot \left(1 + \frac{L_s}{L_u}\right) - \frac{L_s}{L_u} \frac{\dot{L_s}}{L_s}$$

$$g_G = g_b + (g_b - g_G) \frac{L_s}{L_u},$$

which can be true if and only if $g_G = g_b$ because $\frac{L_s}{L_u} > 0$. Thus

$$\frac{\dot{L_s}}{L_s} = \frac{\dot{L_u}}{L_u} = \frac{\dot{G}}{G} = g_b.$$

(40) can be rewritten as

$$\frac{\widehat{\lambda}}{\eta} = -\phi L_u^{1-\kappa} \cdot \kappa G(t)^{\kappa-1} \left(1 - \frac{\psi}{\eta}\right), \tag{65}$$

which is used to substitute out $\frac{\widehat{\lambda}}{\eta}$ in equation (44), we obtain

$$\left(1 - \frac{\psi}{\eta}\right)(1-\kappa)\,\phi \left[\frac{G(t)}{L_u(t)}\right]^{\kappa} = g_b - \xi - \left(1 - \frac{\psi}{\eta}\right)\phi\kappa\delta\frac{a^{n+1} - a^n}{\lambda^{n+1} - \lambda^n}\left[\frac{G(t)}{L_u(t)}\right]^{\kappa-1}. \tag{66}$$

Since $\frac{\dot{\eta}}{\eta} = \frac{\dot{\widehat{\lambda}}}{\widehat{\lambda}} = \frac{\dot{\psi}}{\psi}$, (43) and (41) jointly imply

$$\delta\frac{\widehat{\lambda}}{\eta}\frac{a^n(a-\lambda)}{\lambda - 1} + g_b = \xi\frac{\psi}{\eta} \tag{67}$$

so we have three unknowns $\frac{\widehat{\lambda}}{\eta}, \frac{\psi}{\eta}, \frac{G(t)}{L_u(t)}$ and three equations (45)-(47).

Using the brutal force, we obtain

$$(1-\kappa)\,\phi\left[\frac{G(t)}{L_u(t)}\right]^{\kappa} = \frac{a^n}{\lambda - 1}\left\{(a-\lambda) - \frac{a-1}{\lambda^n}\right\}\phi\kappa\delta\left[\frac{G(t)}{L_u(t)}\right]^{\kappa-1} - \xi,$$

which uniquely determines $\frac{G(t)}{L_u(t)}$.
Denote this solution by χ_n. By (11), we have

$$\frac{L_s}{L_u} = \frac{\phi\chi_n^{\kappa}}{g_b},$$

and

$$\begin{aligned}\frac{\psi}{\eta} &= \frac{g_b - \phi L_u^{1-\kappa}\cdot\kappa G(t)^{\kappa-1}\delta\frac{a^n(a-\lambda)}{\lambda-1}}{\xi - \phi L_u^{1-\kappa}\cdot\kappa G(t)^{\kappa-1}\delta\frac{a^n(a-\lambda)}{\lambda-1}} \\ &= 1 + \frac{\xi - g_b}{(1-\kappa)\,\phi\chi_n^{\kappa} + \frac{a^n}{\lambda-1}\frac{a-1}{\lambda^n}\phi\kappa\delta\chi_n^{\kappa-1}}\end{aligned}$$

$$\begin{aligned}\frac{\widehat{\lambda}}{\eta} &= \frac{\xi\frac{\psi}{\eta} - g_b}{\delta\frac{a^n(a-\lambda)}{\lambda-1}} = \frac{\xi\left[1 - \frac{g_b - \xi}{(1-\kappa)\phi\chi_n^{\kappa} + \frac{a^n}{\lambda-1}\frac{a-1}{\lambda^n}\phi\kappa\delta\chi_n^{\kappa-1}}\right] - g_b}{\delta\frac{a^n(a-\lambda)}{\lambda-1}} \\ &= (\xi - g_b)(\lambda - 1)\frac{1 + \frac{\xi}{(1-\kappa)\phi\chi_n^{\kappa} + \frac{a^n}{\lambda-1}\frac{a-1}{\lambda^n}\phi\kappa\delta\chi_n^{\kappa-1}}}{\delta a^n(a-\lambda)}\end{aligned}$$

$$(1-\kappa)\,\phi\left[\frac{G(t)}{L_u(t)}\right]^{\kappa} + \frac{a^n}{\lambda-1}\frac{a-1}{\lambda^n}\phi\kappa\delta\left[\frac{G(t)}{L_u(t)}\right]^{\kappa-1} = \frac{a^n(a-\lambda)}{\lambda-1}\phi\kappa\delta\left[\frac{G(t)}{L_u(t)}\right]^{\kappa-1} - \xi,$$

$$\begin{aligned}\widehat{m}_n &= \frac{\log \frac{L_u(\widehat{t}_{n+1})+\lambda^{n+1}L_s(\widehat{t}_{n+1})}{L_u(\widehat{t}_n)+\lambda^n L_s(\widehat{t}_n)}}{g_C} = \frac{\log \frac{L_u(\widehat{t}_{n+1})}{L_u(\widehat{t}_n)} \frac{1+\lambda^{n+1}\frac{L_s(\widehat{t}_{n+1})}{L_u(\widehat{t}_{n+1})}}{1+\lambda^n \frac{L_s(\widehat{t}_n)}{L_u(\widehat{t}_n)}}}{g_C} \\ &= \frac{\log e^{g_b(\widehat{t}_{n+1}-\widehat{t}_n)} \frac{1+\lambda^{n+1}\frac{\phi\chi_{n+1}^{\kappa}}{g_b}}{1+\lambda^n\frac{\phi\chi_n^{\kappa}}{g_b}}}{g_C} = \frac{g_b\widehat{m}_n + \log \frac{1+\lambda^{n+1}\frac{\phi\chi_{n+1}^{\kappa}}{g_b}}{1+\lambda^n\frac{\phi\chi_n^{\kappa}}{g_b}}}{g_C},\end{aligned}$$

which implies

$$\widehat{m}_n = \frac{\log \frac{1+\lambda^{n+1}\frac{\phi\chi_{n+1}^{\kappa}}{g_b}}{1+\lambda^n\frac{\phi\chi_n^{\kappa}}{g_b}}}{g_C-g_b} = \frac{\log\left[1+\frac{\lambda^{n+1}\frac{\phi\chi_{n+1}^{\kappa}}{g_b}-\lambda^n\frac{\phi\chi_n^{\kappa}}{g_b}}{1+\lambda^n\frac{\phi\chi_n^{\kappa}}{g_b}}\right]}{g_C-g_b} = \frac{\log\left[1+\frac{\frac{\lambda\chi_{n+1}^{\kappa}}{\chi_n^{\kappa}}-1}{\frac{1}{\lambda^n\frac{\phi\chi_n^{\kappa}}{g_b}}+1}\right]}{g_C-g_b}$$

When $t \in [t_{n+1}^b,\, t_{n+1}^e]$, $E_{0,n+1}(C(t)) = \lambda^{n+1}L_s$, and we have the following discounted-value Hamiltonian:

$$H_{0,n+1} = \frac{C(t)^{1-\sigma}-1}{1-\sigma}e^{-\rho t} + \widehat{\lambda}_{n+1}\left[\xi K - \delta\cdot\lambda^{n+1}L_s\right] + \eta_{n+1}(\lambda^{n+1}L_s + L_u - C) \tag{68}$$

$$\begin{aligned}\frac{\partial H_{0,n+1}}{\partial C} &= 0 \Rightarrow C(t)^{-\sigma}e^{-\rho t} = \eta_{n+1} \\ -\dot{\widehat{\lambda}}_{n+1} &= \frac{\partial H_{0,n+1}}{\partial K} = \xi\widehat{\lambda}_{n+1}\end{aligned}$$

thus we have

$$\widehat{\lambda}_{n+1}(t) = \widehat{\lambda}_{n+1}(t_{n+1}^b)e^{-\xi(t-t_{n+1}^b)}$$

for any $t \in [t_{n+1}^b,\, t_{n+1}^e]$. In particular,

$$\widehat{\lambda}_{n+1}(t_{n+1}^e) = \widehat{\lambda}_{n+1}(t_{n+1}^b)e^{-\xi(t_{n+1}^e-t_{n+1}^b)},$$

Note that the optimality condition requires

$$\begin{aligned}\widehat{\lambda}_{n,n+1}(t_{n+1}^b) &= \widehat{\lambda}_{n+1}(t_{n+1}^b); \\ \widehat{\lambda}_{n+1}(t_{n+1}^e) &= \widehat{\lambda}_{n+1,n+2}(t_{n+1}^e),\end{aligned}$$

therefore, the previous three equations jointly imply that

$$\widehat{\lambda}_{n+1,n+2}(t_{n+1}^e) = \widehat{\lambda}_{n,n+1}(t_{n+1}^b)e^{-\xi(t_{n+1}^e-t_{n+1}^b)}$$

which, together with (35), (36) and (37), yields

$$t_{n+1}^e - t_{n+1}^b = \frac{\ln\frac{a}{\lambda}}{\xi-\rho}.$$

Thus the whole life cycle of industry $n+1$ consists of a rising period (t_n^e, t_{n+1}^b), a peak period $[t_{n+1}^b,\, t_{n+1}^e]$, and a decline period $(t_{n+1}^e, t_{n+2}^b]$. Its life span is equal to $t_{n+2}^b - t_n^e$. It is easy to see that

$$t_{n+1}^b - t_n^e = \frac{\ln\frac{\lambda^{n+1}L_s+L_u}{\lambda^n L_s+L_u}}{\frac{\xi-\rho}{\sigma}};\forall n \geq 1.$$

To have sustainable consumption growth, we have to ensure that $\xi > g_b$, in which case we must have

$$\frac{\dot{E}}{E} > \frac{\dot{L}_s}{L_s} \geq \frac{\dot{L}_u}{L_u}.$$

Suppose we have $\frac{\dot{L}_s}{L_s} = \frac{\dot{L}_u}{L_u}$ in equilibrium, then (12) implies $\frac{\dot{L}_s}{L_s} = \frac{\dot{L}_u}{L_u} = g_b$. (11) further implies $g_G = g_b$. Everything becomes the same as before, and we reach a contradiction because $\xi > g_b$. Thus, it must be true that $\frac{\dot{L}_s}{L_s} \neq \frac{\dot{L}_u}{L_u}$.

Suppose $\frac{\dot{L}_s}{L_s} > \frac{\dot{L}_u}{L_u}$, and $\frac{L_u(t)}{L_s(t)} \to 0$, $L_u(t) \to 0$ and $g_G > g_b$. So all labor is eventually transformed into skilled labor, and the economy eventually grows like that in Ju, Lin and Wang (2015). (40) and (42) jointly imply

$$\frac{\dot{\eta}}{\eta} = \rho - \left[g_b - (1-\kappa)\,\phi L_u^{-\kappa} \cdot G(t)^{\kappa}\right] + \delta\phi L_u^{1-\kappa} \cdot \kappa G(t)^{\kappa-1} \frac{a^{n+1} - a^n}{\lambda^{n+1} - \lambda^n}.$$

Suppose $\frac{\dot{\eta}}{\eta}, \frac{\dot{L}_u}{L_u}$ and g_G are all constant, then

$$-\phi L_u^{1-\kappa} \cdot \kappa G(t)^{\kappa-1} = \frac{\lambda}{\eta}$$

implies

$$\begin{aligned}
(1-\kappa)(g_u - g_G) &= \rho - \xi - \frac{\dot{\eta}}{\eta} \\
&= -\xi + g_b - (1-\kappa)\,\phi L_u^{-\kappa} \cdot G(t)^{\kappa} - \delta\phi L_u^{1-\kappa} \cdot \kappa G(t)^{\kappa-1} \frac{a^{n+1} - a^n}{\lambda^{n+1} - \lambda^n}
\end{aligned}$$

which is a constant if and only if

$$\begin{aligned}
& -(1-\kappa)\,\phi L_u^{-\kappa} \cdot G(t)^{\kappa} - \delta\phi L_u^{1-\kappa} \cdot \kappa G(t)^{\kappa-1} \frac{a^{n+1} - a^n}{\lambda^{n+1} - \lambda^n} \\
= & \; \phi L_u^{1-\kappa} G(t)^{\kappa-1} \left[-(1-\kappa)\,L_u^{-1} \cdot G(t) - \delta\kappa \frac{a^{n+1} - a^n}{\lambda^{n+1} - \lambda^n}\right] \\
= & \; \phi \left[\frac{G(0)}{L_u(0)}\right]^{\kappa-1} e^{(1-\kappa)(g_u - g_G)t} \left[-(1-\kappa)\,\frac{G(0)}{L_u(0)} e^{-(g_u - g_G)t} - \delta\kappa \frac{a^{n+1} - a^n}{\lambda^{n+1} - \lambda^n}\right]
\end{aligned}$$

is a constant.

Appendix

Figure A2 shows the counterpart for Fact 0 and Table A2 establishes the counterpart for Fact 1. The evidence of the counterpart for Facts 2 to 4 is provided below.

The counterpart for fact 2 and 3:
To further investigate the cross-country empirical evidence for industry dynamics, we extend regression (1) to a cross-country regression:

$$Y_{itc} = \beta_0 + \beta_1 t + \beta_2 t^2 + \beta_3 k_{itc} \cdot t + \beta_4 T_{itc} + \beta_5 D_{ic} + \beta_6 GDPGR_{tc} + \varepsilon_{itc}. \tag{69}$$

where subscript c represents country. The results are summarized in Table A2.
We regress the peak time of a industry's share (either employment share or output share) on its capital-skilled labor ratio and capital-labor ratio. The results are in Table A3. Column (1) and column (3) show that the peak time of an industry is positively correlated with its capital-skilled labor ratio.
The findings are consistent with the results in Table 2 and Table 3, suggesting that the patterns observed in the US are also true for other countries. Next we extend regression (2) to a cross-country regression. The results for the cross-country counterpart of regression (2) are summarized in Table A4.
The findings are consistent with the results in Table 4.

A different formulation of externality

Everything is identical to the previous case excep that (11) is replaced by

$$\dot{L}_s = \phi L_s^{1-\nu} \left[L_u^{1-\kappa} \cdot G(t)^{\kappa} \right]^{\nu} \overline{G}^{\zeta}(t), \tag{70}$$

where $\nu \in [0, 1]$. When $1 - \nu > 0$, it captures the positive externality of existing skilled labor on the training/ learning of unskilled labor.

$$\max_{C(t),G(t)} \int_0^{\infty} \frac{C(t)^{1-\sigma} - 1}{1-\sigma} e^{-\rho t} dt, \text{ where } \sigma \in (0, 1),$$

subject to

$$\begin{aligned} \dot{K} &= \xi K - \delta \cdot E(C(t)) - G(t), & (71) \\ \dot{L}_s &= \phi L_s^{1-\nu} \left[L_u^{1-\kappa} \cdot G(t)^{\kappa} \right]^{\nu} \overline{G}^{\zeta}(t), & (72) \\ \dot{L}_u &= g_b \cdot (L_u + L_s) - \dot{L}_s & (73) \end{aligned}$$

Decentralized equilibrium:

$$\dot{L}_u = g_b \cdot (L_u + L_s) - \phi L_s^{1-\nu} \left[L_u^{1-\kappa} \cdot G(t)^{\kappa} \right]^{\nu} \overline{G}^{\zeta}(t). \tag{74}$$

Suppose $C(t) \in (\lambda^n L_s + L_u, \lambda^{n+1} L_s + L_u)$ for any $n \geq 1$. Establish the current-value Hamiltonian as follows:

$$H = \frac{C(t)^{1-\sigma} - 1}{1-\sigma} + \lambda\left[\xi K - \delta \cdot E(C(t)) - G(t)\right] + \eta\left[g_b \cdot (L_u + L_s) - \phi L_s^{1-\nu}\left[L_u^{1-\kappa} \cdot G(t)^{\kappa}\right]^{\nu} \overline{G}^{\zeta}(t)\right]$$

First order conditions:

$$\frac{\partial H}{\partial C} = 0 \Rightarrow C(t)^{-\sigma} = \lambda\delta\frac{\partial E(C(t))}{\partial C(t)} \Rightarrow -\sigma\frac{\dot{C}}{C} = \frac{\dot{\lambda}}{\lambda} \tag{75}$$

$$\frac{\partial H}{\partial G} = 0 \Rightarrow \eta\phi\kappa\nu L_s^{1-\nu}\left[L_u^{1-\kappa} \cdot G(t)^{\kappa}\right]^{\nu} \overline{G}^{\zeta}(t) \cdot G(t)^{-1} = \lambda \tag{76}$$

$$\dot{\lambda} = \rho\lambda - \frac{\partial H}{\partial K} = \rho\lambda - \lambda\xi \Rightarrow \frac{\dot{\lambda}}{\lambda} = \rho - \xi$$

$$\dot{\eta} = \rho\eta - \frac{\partial H}{\partial L_u} \Rightarrow \frac{\dot{\eta}}{\eta} = \rho - \left[g_b - (1-\kappa)\phi L_u^{-\kappa} \cdot G(t)^{\kappa}\overline{G}^{\zeta}(t)\right] \tag{77}$$

Part C. Human Capital Externality

Now we introduce externality in human capital investment. Everything is identical to Part B except that (11) is replaced by the following:

$$\dot{L}_s = \phi L_u^{1-\kappa} \cdot G(t)^{\kappa}\overline{G}^{\zeta}(t), \tag{78}$$

where $\overline{G}(t)$ is the average household spending on human capital investment at time t and the parameter $\zeta \geq 0$. Part B is a special case when $\zeta = 0$. When $\zeta > 0$, it captures the positive externality in human capital investment, which is our analytical focus below.

Proposition 6. *With positive externality in human capital investment ($\zeta > 0$), there exists no Balanced Growth Path. Instead, on the transitional path the following is true*

$$g_C = \frac{\xi - \rho}{\sigma}, \forall t \geq \widehat{t}_0$$

$$\frac{\dot{L}_s}{L_s} > g_b > \frac{\dot{L}_u}{L_u}, \forall t \tag{79}$$

$$\frac{L_s}{L_u} = \frac{\phi\chi_n^{\kappa}}{g_b}, \forall t \in (\widehat{t}_n, \widehat{t}_{n+1}] \tag{80}$$

$$\frac{G(t)}{L_u(t)} = \chi_n, \forall t \in (\widehat{t}_n, \widehat{t}_{n+1}) \tag{81}$$

when industry n and industry $n+1$ coexist, (that is, $t \in [\widehat{t}_n, \widehat{t}_{n+1})$), the following is true:

$$\xi + \phi(1-\kappa)L_u^{-\kappa} \cdot G(t)^{\kappa+\zeta} + (1-\kappa)\frac{\dot{L}_u}{L_u} + (\kappa - 1 + \zeta)\frac{\dot{G}}{G} = \left[\frac{a^n(a-\lambda)}{\lambda - 1} - \frac{a^{n+1} - a^n}{\lambda^{n+1} - \lambda^n}\right]\delta\phi L_u^{1-\kappa} \cdot \kappa G(t)^{\kappa-1+\zeta}$$

for $\forall n \geq 1$. In the very long run, $\frac{L_s}{L_u + L_s} \to 1$, and the economy converges to one with only skilled labor and physical capital, which is characterized in Ju,Lin and Wang (2015).

Proof: Substituting (96) into (12) yields

$$\dot{L}_u = g_b \cdot (L_u + L_s) - \phi L_u^{1-\kappa} \cdot G(t)^{\kappa} \overline{G}^{\zeta}(t). \tag{82}$$

Suppose $C(t) \in (\lambda^n L_s + L_u, \lambda^{n+1} L_s + L_u)$ for any $n \geq 1$. Establish the current-value Hamiltonian as follows:

$$\begin{aligned} H &= \frac{C(t)^{1-\sigma} - 1}{1-\sigma} + \lambda \left[\xi K - \delta \cdot E(C(t)) - G(t)\right] + \eta \left[g_b \cdot (L_u + L_s) - \phi L_u^{1-\kappa} \cdot G(t)^{\kappa} \overline{G}^{\zeta}(t)\right] \\ & \quad + \psi \phi L_u^{1-\kappa} \cdot G(t)^{\kappa} \overline{G}^{\zeta}(t) \end{aligned}$$

To characterize the decentralized equilibrium, we derive the following optimality conditions:

$$\begin{aligned} \frac{\partial H}{\partial C} &= 0 \Rightarrow C(t)^{-\sigma} = \lambda \delta \frac{\partial E(C(t))}{\partial C(t)} \Rightarrow -\sigma \frac{\dot{C}}{C} = \frac{\dot{\lambda}}{\lambda} \\ \frac{\partial H}{\partial G} &= 0 \Rightarrow (\psi - \eta) \phi L_u^{1-\kappa} \cdot \kappa G(t)^{\kappa-1} \overline{G}^{\zeta}(t) = \lambda & (83) \\ \dot{\lambda} &= \rho \lambda - \frac{\partial H}{\partial K} = \rho \lambda - \lambda \xi \Rightarrow \frac{\dot{\lambda}}{\lambda} = \rho - \xi \\ \frac{\dot{\psi}}{\psi} &= \rho - \left[\delta \frac{\widehat{\lambda}}{\psi} \frac{a^n (a - \lambda)}{\lambda - 1} + \frac{\eta}{\psi} g_b\right] & (84) \\ \dot{\eta} &= \rho \eta - \frac{\partial H}{\partial L_u} \Rightarrow \frac{\dot{\eta}}{\eta} = \rho - g_b - \delta \frac{\widehat{\lambda}}{\eta} \frac{a^{n+1} - a^n}{\lambda^{n+1} - \lambda^n} + \left(1 - \frac{\psi}{\eta}\right) \phi (1-\kappa) L_u^{-\kappa} \cdot G(t)^{\kappa} \overline{G}^{\zeta}(t). & (85) \end{aligned}$$

Thus we still have

$$\frac{\dot{C}}{C} = \frac{\xi - \rho}{\sigma}.$$

Moreover, in equilibrium, we must have $\overline{G}(t) = G(t)$, so (100) can be rewritten as

$$\frac{\dot{\eta}}{\eta} = \rho - \left[g_b - \left(1 - \frac{\psi}{\eta}\right)(1-\kappa) \phi L_u^{-\kappa} \cdot G(t)^{\kappa+\zeta}\right] - \delta \frac{\widehat{\lambda}}{\eta} \frac{a^{n+1} - a^n}{\lambda^{n+1} - \lambda^n}, \tag{86}$$

which is a constant if

$$\frac{\dot{\eta}}{\eta} = \frac{\dot{\psi}}{\psi} = \frac{\dot{\widehat{\lambda}}}{\widehat{\lambda}} = \rho - \xi \tag{87}$$

and

$$\frac{\dot{L}_u}{L_u} = \frac{(\kappa + \zeta)}{\kappa} \frac{\dot{G}}{G}. \tag{88}$$

Substituting $\overline{G}(t) = G(t)$ into (98) yields

$$\frac{\widehat{\lambda}}{\eta} = \left(\frac{\psi}{\eta} - 1\right) \phi L_u^{1-\kappa} \cdot \kappa G(t)^{\kappa - 1 + \zeta}, \tag{89}$$

Substitute the above into (101) and use (102), we obtain

$$\xi - g_b - \left(\frac{\psi}{\eta} - 1\right)\phi\left[(1-\kappa)\, L_u^{-\kappa} \cdot G(t)^{\kappa+\zeta} + \delta\frac{a^{n+1} - a^n}{\lambda^{n+1} - \lambda^n} L_u^{1-\kappa} \cdot \kappa G(t)^{\kappa-1+\zeta}\right] = 0 \tag{90}$$

By revoking (102) and (99), we obtain

$$\delta\frac{\widehat{\lambda}}{\eta}\frac{a^n(a-\lambda)}{\lambda - 1} + g_b = \frac{\psi}{\eta}\xi. \tag{91}$$

(104) and (106) jointly imply

$$\frac{\psi}{\eta} - 1 = \frac{\xi - g_b}{\delta\phi L_u^{1-\kappa} \cdot \kappa G(t)^{\kappa-1+\zeta}\frac{a^n(a-\lambda)}{\lambda-1} - \xi}.$$

Substitute the above into (105), we obtain

$$\left[\frac{a^n(a-\lambda)}{\lambda - 1} - \frac{a^{n+1} - a^n}{\lambda^{n+1} - \lambda^n}\right]\delta\phi L_u^{1-\kappa} \cdot \kappa G(t)^{\kappa-1+\zeta} = \phi\,(1-\kappa)\, L_u^{-\kappa} \cdot G(t)^{\kappa+\zeta} + \xi \tag{92}$$

$$\phi\,(1-\kappa)\, L_u^{-\kappa} \cdot G(t)^{\kappa}\overline{G}^{\zeta}(t) + (1-\kappa)\frac{\dot{L}_u}{L_u} + (\kappa - 1 + \zeta)\, g_G + \xi = \left[\frac{a^n(a-\lambda)}{\lambda - 1} - \frac{a^{n+1} - a^n}{\lambda^{n+1} - \lambda^n}\right]\delta\phi L_u^{1-\kappa} \cdot \kappa G(t)^{\kappa-1}\overline{G}^{\zeta}(t$$

Corallary: When $\kappa - 1 + \zeta = 0$, the above equation becomes

$$\phi\,(1-\kappa)\, L_u^{-\kappa} \cdot G(t) + (1-\kappa)\frac{\dot{L}_u}{L_u} + \xi = \left[\frac{a^n(a-\lambda)}{\lambda - 1} - \frac{a^{n+1} - a^n}{\lambda^{n+1} - \lambda^n}\right]\delta\phi L_u^{1-\kappa} \cdot \kappa$$

which implies

$$G(t) = \frac{\left[\frac{a^n(a-\lambda)}{\lambda-1} - \frac{a^{n+1}-a^n}{\lambda^{n+1}-\lambda^n}\right]\delta\phi L_u^{1-\kappa} \cdot \kappa - \xi - (1-\kappa)\frac{\dot{L}_u}{L_u}}{\phi\,(1-\kappa)\, L_u^{-\kappa}}$$

or equivalently,

$$\frac{G(t)}{L_u} = \frac{\delta\kappa\left[\frac{a^n(a-\lambda)}{\lambda-1} - \frac{a^{n+1}-a^n}{\lambda^{n+1}-\lambda^n}\right]}{(1-\kappa)} - \frac{\left[\xi + (1-\kappa)\frac{\dot{L}_u}{L_u}\right]}{\phi\,(1-\kappa)\, L_u^{1-\kappa}},$$

which strictly increases over time if $\frac{\dot{L}_u}{L_u} > 0$

Substituting $\overline{G}(t) = G(t)$ into (97) yields

$$\dot{L}_u = g_b \cdot (L_u + L_s) - \phi L_u^{1-\kappa} \cdot G(t)^{\kappa+\zeta},$$

so $\frac{\dot{L}_u}{L_u}$ is constant if and only if $\frac{\dot{L}_s}{L_s} = \frac{\dot{L}_u}{L_u}$ because of (103). (96) implies

$$\frac{\dot{L}_s}{L_s} = \frac{L_u}{L_s}\phi L_u^{-\kappa} \cdot G(t)^{\kappa+\zeta} \tag{93}$$

Let $g_G \equiv \frac{\dot{G}}{G}$. (97) implies

$$\frac{\dot{L}_u}{L_u} = g_b \cdot \left(1 + \frac{L_s}{L_u}\right) - \phi L_u^{-\kappa} \cdot G(t)^{\kappa+\zeta},$$

together with (93), implies

$$g_G = g_b \frac{\kappa}{\kappa + \zeta},$$

and

$$\frac{\dot{L}_s}{L_s} = \frac{\dot{L}_u}{L_u} = g_b.$$

Let $\Lambda \equiv \phi L_u^{-\kappa} \cdot G(t)^{\kappa+\zeta}$, so

$$\frac{L_u}{L_s} = \frac{\frac{\dot{L}_s}{L_s}}{\phi L_u^{-\kappa} \cdot G(t)^{\kappa+\zeta}} = \frac{g_b}{\Lambda}.$$

(107) implies

$$\frac{G(t)}{L_u} = \frac{\left[\frac{a^n(a-\lambda)}{\lambda-1} - \frac{a^{n+1}-a^n}{\lambda^{n+1}-\lambda^n}\right] \delta\kappa\Lambda}{(1-\kappa)\Lambda + \xi}$$

(98), which contradicts that Λ is a constant because $g_G = g_b \frac{\kappa}{\kappa+\zeta}$ unless $g_G = g_b = 0$.

When $g_G = g_b = 0$, $G(t)$ and L_u, L_s are all constant. In particular, when $G(t) = L_u = 0$, the economy becomes identical to the economy characterized in Ju, Lin and Wang (2015).

$$(\psi - \eta)\, \phi L_u^{1-\kappa} \cdot \kappa G(t)^{\kappa-1} \overline{G}^{\zeta} ($$

$$\frac{\dot{\psi} - \dot{\eta}}{\psi - \eta} + (1-\kappa) \frac{\dot{L}_u}{L_u} + (\kappa - 1 + \zeta)\, g$$

$$\frac{\psi}{\psi - \eta} \frac{\dot{\psi}}{\psi} - \frac{\dot{\eta}}{\eta} \frac{\eta}{\psi - \eta} + (1-\kappa) \frac{\dot{L}_u}{L_u} + (\kappa - 1 + \zeta)\, g$$

$$\left[\frac{a^{n+1} - a^n}{\lambda^{n+1} - \lambda^n} - \frac{a^n(a-\lambda)}{\lambda - 1}\right] \frac{\widehat{\lambda}}{\psi - \eta} \delta + \phi (1-\kappa) L_u^{-\kappa} \cdot G(t)^{\kappa} \overline{G}^{\zeta}(t) + (1-\kappa) \frac{\dot{L}_u}{L_u} + (\kappa - 1 + \zeta)\, g$$

$$\left[\frac{a^{n+1} - a^n}{\lambda^{n+1} - \lambda^n} - \frac{a^n(a-\lambda)}{\lambda - 1}\right] \delta\phi L_u^{1-\kappa} \cdot \kappa G(t)^{\kappa-1} \overline{G}^{\zeta}(t) + \phi (1-\kappa) L_u^{-\kappa} \cdot G(t)^{\kappa} \overline{G}^{\zeta}(t) + (1-\kappa) \frac{\dot{L}_u}{L_u} + (\kappa - 1 + \zeta)\, g$$

If

$$L_u^{1-\kappa} G(t)^{\kappa-1} \overline{G}^{\zeta}(t) = \Xi = const$$

then

$$\frac{G(t)}{L_u} = \frac{\left[\frac{a^n(a-\lambda)}{\lambda-1} - \frac{a^{n+1}-a^n}{\lambda^{n+1}-\lambda^n}\right] \delta\phi\kappa\Xi - \xi}{\phi(1-\kappa)\Xi}$$

$$g_G = \frac{-(1-\kappa)}{(\kappa - 1 + \zeta)} \frac{\dot{L}_u}{L_u} = \frac{(1-\kappa)}{(1-\kappa-\zeta)} \frac{\dot{L}_u}{L_u}$$

Suppose $\frac{\dot{L}_u}{L_u} \neq 0$, then $g_G \neq \frac{\dot{L}_u}{L_u}$, contradicting the fact that Ξ is a constant. Thus we must have $\frac{\dot{L}_u}{L_u} = 0$ and $G(t)$ must be a constant. In that case,

$$\left[\frac{a^{n+1} - a^n}{\lambda^{n+1} - \lambda^n} - \frac{a^n(a-\lambda)}{\lambda - 1}\right] \delta\phi L_u^{1-\kappa} \cdot \kappa G(t)^{\kappa-1} \overline{G}^{\zeta}(t) + \phi(1-\kappa) L_u^{-\kappa} \cdot G(t)^{\kappa} \overline{G}^{\zeta}(t) = -\xi$$

$$\begin{aligned} \dot{L}_s &= \phi L_u^{1-\kappa} \cdot G(t)^{\kappa} \overline{G}^{\zeta}(t), \\ \dot{L}_s &= g_b \cdot (L_u + L_s) \\ g_b \cdot (L_u + L_s) &= \phi L_u^{1-\kappa} \cdot G(t)^{\kappa} \overline{G}^{\zeta}(t) \end{aligned}$$

so g_b has to be zero and L_s is a constant. Moreover, $L_u = 0$ or $G = 0$.

Suppose $L_u = 0$.

]]

Suppose

$$L_u^{1-\kappa} G(t)^{\kappa-1} \overline{G}^{\zeta}(t) = \Xi(t),$$

define

$$g_\Xi(t) \equiv \frac{\dot{\Xi}(t)}{\Xi(t)},$$

then

$$\begin{aligned} (1-\kappa) \frac{\dot{L}_u}{L_u} + (\kappa - 1 + \zeta) g_G &= g_\Xi(t) \\ g_G - \frac{\dot{L}_u}{L_u} &= \frac{g_\Xi(t) - \zeta g_G}{(\kappa - 1)} \end{aligned} \tag{95}$$

substituting it into (94) yields

$$\frac{G(t)}{L_u} = \frac{\left[\frac{a^n(a-\lambda)}{\lambda-1} - \frac{a^{n+1}-a^n}{\lambda^{n+1}-\lambda^n}\right] \kappa\delta\phi\Xi(t) - [\xi + g_\Xi(t)]}{(1-\kappa)\phi\Xi(t)},$$

which increases over time if and only if $\frac{\xi + g_\Xi(t)}{\Xi(t)}$ decreases over time, or equivalently,

$$g_\Xi'(t) - [\xi + g_\Xi(t)] g_\Xi(t) < 0$$

$$g_\Xi(t) < \zeta g_G$$

because of (95).

In particular, when $\kappa - 1 + \zeta = 0$, we have $g_\Xi(t) = (1-\kappa)\frac{\dot{L_u}}{L_u}$, so $g_\Xi(t) < \zeta g_G$ is reduced to

$$\frac{\dot{L_u}}{L_u} < g_G.$$

=====================

$$L_u^{-\kappa} \cdot G(t)^{\kappa+\zeta}$$

$$\begin{aligned} \Lambda &\equiv L_u^{-\kappa} \cdot G(t)^{\kappa+\zeta} \\ \frac{\dot{\Lambda}}{\Lambda} &= g_\Xi(t) - \frac{\dot{L_u}}{L_u} + g_G \end{aligned}$$

$$\frac{\dot{L_u}}{L_u} = g_b \cdot \left(1 + \frac{L_s}{L_u}\right) - \phi\Lambda,$$

$$\frac{L_s}{L_u}\frac{\dot{L_s}}{L_s} = \phi\Lambda$$

In equilibrium, we have $\frac{\dot{L_s}}{L_s} > g_b > \frac{\dot{L_u}}{L_u}$, both $\frac{L_s}{L_u}$ and Λ increase over time, $\frac{\dot{L_u}}{L_u}$ is positive at the begining and eventually becomes negative (that is, $g_b \cdot \left(1 + \frac{L_s}{L_u}\right) > \phi\Lambda$ when t is sufficiently small and then the opposite is true afterwards)...

$$\begin{aligned} g_b \cdot \left(1 + \frac{L_s}{L_u}\right) &> \phi\Lambda \\ \phi\Lambda &> \frac{\dot{L_s}}{L_s}\left(\frac{\phi\Lambda}{g_b} - 1\right) \\ \frac{\dot{L_s}}{L_s} &> \left(\frac{\frac{\dot{L_s}}{L_s}}{g_b} - 1\right)\phi\Lambda \end{aligned}$$

When $\frac{\dot{L_u}}{L_u} = 0$, we have

$$\frac{L_s}{L_u} = \frac{\phi\Lambda}{g_b} - 1$$

and

$$\frac{\dot{L_s}}{L_s} = \frac{\phi\Lambda g_b}{\phi\Lambda - g_b} > g_b,$$

which means that $\phi\Lambda > g_b$ and

$$g_b = \frac{\dot{L}_s + \dot{L}_u}{L_s + L_u} = \frac{\dot{L}_s}{L_s}\frac{L_s}{L_s + L_u} + \frac{\dot{L}_u}{L_u}\frac{L_u}{L_s + L_u}$$

$$\frac{\dot{L}_u}{L_u} = g_b \cdot \left(1 + \frac{L_s}{L_u}\right) - \phi L_u^{-\kappa} \cdot G(t)^{\kappa+\zeta},$$

$$L_u^{1-\kappa} G(t)^{\kappa-1} \overline{G}^{\zeta}(t) = \Xi(t),$$

$$\frac{\dot{L}_s}{L_s}\frac{L_s}{L_u} = \phi L_u^{-\kappa} \cdot G(t)^{\kappa+\zeta} = \frac{\left[\frac{a^n(a-\lambda)}{\lambda-1} - \frac{a^{n+1}-a^n}{\lambda^{n+1}-\lambda^n}\right]\kappa\delta\phi\Xi(t) - [\xi + g_{\Xi}(t)]}{(1-\kappa)}$$

Benevolent social planner problem taking into account the human capital externality:
(98) is changed to

$$\frac{\partial H}{\partial G} = 0 \Rightarrow (\psi - \eta)\,\phi L_u^{1-\kappa} \cdot (\kappa + \zeta)\, G(t)^{\kappa+\zeta-1} = \lambda$$

All the growth rates on the BGP are still the same as in the case when human capital externality is not internalized in the decentralized decisions. The major difference is the level effect instead of speed effect.

Part C. Human Capital Externality

Now we introduce externality in human capital investment. Everything is identical to Part B except that (11) is replaced by the following:

$$\dot{L}_s = \phi \left[L_u^{1-\kappa} \cdot G(t)^{\kappa}\right]^{1-\zeta} \overline{G}^{\zeta}(t), \tag{96}$$

where $\overline{G}(t)$ is the average household spending on human capital investment at time t and the parameter $\zeta \geq 0$. Part B is a special case when $\zeta = 0$. When $\zeta > 0$, it captures the positive externality in human capital investment, which is our analytical focus below.

$$\dot{L}_u = g_b \cdot (L_u + L_s) - \phi \left[L_u^{1-\kappa} \cdot G(t)^{\kappa}\right]^{1-\zeta} \overline{G}^{\zeta}(t). \tag{97}$$

Suppose $C(t) \in (\lambda^n L_s + L_u, \lambda^{n+1} L_s + L_u)$ for any $n \geq 1$. Establish the current-value Hamiltonian as follows:

$$\begin{aligned} H \;=\; & \frac{C(t)^{1-\sigma} - 1}{1-\sigma} + \lambda\left[\xi K - \delta \cdot E(C(t)) - G(t)\right] + \eta\left[g_b \cdot (L_u + L_s) - \phi\left[L_u^{1-\kappa} \cdot G(t)^{\kappa}\right]^{1-\zeta} \overline{G}^{\zeta}(t))\right] \\ & + \psi\phi\left[L_u^{1-\kappa} \cdot G(t)^{\kappa}\right]^{1-\zeta} \overline{G}^{\zeta}(t) \end{aligned}$$

To characterize the decentralized equilibrium, we derive the following optimality conditions:

$$
\begin{aligned}
\frac{\partial H}{\partial C} &= 0 \Rightarrow C(t)^{-\sigma} = \widehat{\lambda}\delta \frac{\partial E(C(t))}{\partial C(t)} \Rightarrow -\sigma \frac{\dot{C}}{C} = \frac{\dot{\widehat{\lambda}}}{\widehat{\lambda}} \\
\frac{\partial H}{\partial G} &= 0 \Rightarrow (\psi - \eta)\, \phi\, (1-\zeta) \left[L_u^{1-\kappa} \cdot G(t)^{\kappa}\right]^{-\zeta} \overline{G}^{\zeta}(t) \kappa L_u^{1-\kappa} \cdot G(t)^{\kappa-1} = \widehat{\lambda} \\
\dot{\widehat{\lambda}} &= \rho \widehat{\lambda} - \frac{\partial H}{\partial K} = \rho \widehat{\lambda} - \widehat{\lambda}\xi \Rightarrow \frac{\dot{\widehat{\lambda}}}{\widehat{\lambda}} = \rho - \xi \\
\frac{\dot{\psi}}{\psi} &= \rho - \left[\delta \frac{\widehat{\lambda}}{\psi} \frac{a^n (a-\lambda)}{\lambda - 1} + \frac{\eta}{\psi} g_b \right] \\
\dot{\eta} &= \rho \eta - \frac{\partial H}{\partial L_u} \Rightarrow \frac{\dot{\eta}}{\eta} = \rho - g_b - \delta \frac{\widehat{\lambda}}{\eta} \frac{a^{n+1} - a^n}{\lambda^{n+1} - \lambda^n} + \left(1 - \frac{\psi}{\eta}\right) \phi\, (1-\kappa)\, L_u^{-\kappa}\, (1-\zeta) \left[L_u^{1-\kappa} \cdot G(t)^{\kappa}\right]^{-\zeta} \cdot G(t)
\end{aligned}
$$

Thus we still have

$$
\frac{\dot{C}}{C} = \frac{\xi - \rho}{\sigma}.
$$

Moreover, in equilibrium, we must have $\overline{G}(t) = G(t)$, so (100) can be rewritten as

$$
\frac{\dot{\eta}}{\eta} = \rho - \left[g_b - \left(1 - \frac{\psi}{\eta}\right) \phi\, (1-\kappa)\, (1-\zeta)\, L_u^{-\zeta(1-\kappa)-\kappa} \cdot G(t)^{\kappa+\zeta-\zeta\kappa}\right] - \delta \frac{\widehat{\lambda}}{\eta} \frac{a^{n+1} - a^n}{\lambda^{n+1} - \lambda^n}, \quad (101)
$$

which is a constant if

$$
\frac{\dot{\eta}}{\eta} = \frac{\dot{\psi}}{\psi} = \frac{\dot{\widehat{\lambda}}}{\widehat{\lambda}} = \rho - \xi \quad (102)
$$

and

$$
\frac{\dot{L}_u}{L_u} = \frac{\dot{G}}{G}. \quad (103)
$$

Substituting $\overline{G}(t) = G(t)$ into (98) yields

$$
\left(\frac{\psi}{\eta} - 1\right) \phi \kappa\, (1-\zeta)\, G^{\zeta+\kappa-1-\kappa\zeta} L_u^{(1-\kappa)(1-\zeta)} = \frac{\widehat{\lambda}}{\eta}, \quad (104)
$$

Substitute the above into (101) and use (102), we obtain

$$
\xi - g_b + \left(1 - \frac{\psi}{\eta}\right) \phi\, (1-\zeta) \left[(1-\kappa)\, L_u^{-\zeta(1-\kappa)-\kappa} \cdot G(t)^{\kappa+\zeta-\zeta\kappa} + \delta \kappa G^{\zeta+\kappa-1-\kappa\zeta} L_u^{(1-\kappa)(1-\zeta)} \frac{a^{n+1} - a^n}{\lambda^{n+1} - \lambda^n}\right] = 0 \quad (105)
$$

By revoking (102) and (99), we obtain

$$
\delta \frac{\widehat{\lambda}}{\eta} \frac{a^n (a-\lambda)}{\lambda - 1} + g_b = \frac{\psi}{\eta} \xi. \quad (106)
$$

(104) and (106) jointly imply

$$\frac{\psi}{\eta}-1=\frac{\xi-g_b}{\delta\frac{a^n(a-\lambda)}{\lambda-1}\phi\kappa\left(1-\zeta\right)G^{\zeta+\kappa-1-\kappa\zeta}L_u^{(1-\kappa)(1-\zeta)}-\xi}$$

Substitute the above into (105), we obtain

$$\xi-g_b-\frac{\xi-g_b}{\delta\frac{a^n(a-\lambda)}{\lambda-1}\phi\kappa\left(1-\zeta\right)G^{\zeta+\kappa-1-\kappa\zeta}L_u^{(1-\kappa)(1-\zeta)}-\xi}\phi\left(1-\zeta\right)\Big[\left(1-\kappa\right)L_u^{-\zeta(1-\kappa)-\kappa}\cdot G(t)^{\kappa+\zeta-\zeta\kappa}+\delta\kappa G^{\zeta+\kappa-1-\kappa\zeta}$$

when $\xi>g_b$, we have

$$\left[\frac{a^n(a-\lambda)}{\lambda-1}-\frac{a^{n+1}-a^n}{\lambda^{n+1}-\lambda^n}\right]\delta\kappa\left(\frac{G}{L_u}\right)^{-(1-\kappa)(1-\zeta)}=\left(1-\kappa\right)\left(\frac{G}{L_u}\right)^{\kappa+\zeta-\zeta\kappa}+\frac{\xi}{\phi\left(1-\zeta\right)}\qquad(107)$$

Observe that there exists a unique solution to the above equation, and $\frac{G}{L_u}$ has the following properties:

$$\frac{\partial}{\partial\xi}\left(\frac{G}{L_u}\right)<0;\frac{\partial}{\partial\delta}\left(\frac{G}{L_u}\right)>0;\frac{\partial}{\partial n}\left(\frac{G}{L_u}\right)>0;\frac{\partial}{\partial\phi}\left(\frac{G}{L_u}\right)>0;\frac{\partial}{\partial\zeta}\left(\frac{G}{L_u}\right)<0;\frac{\partial}{\partial a}\left(\frac{G}{L_u}\right)>0$$

$$\begin{aligned}\frac{\dot{L_u}}{L_u} &= g_b\cdot\left(1+\frac{L_s}{L_u}\right)-\phi\frac{\left[L_u^{1-\kappa}\cdot G(t)^{\kappa}\right]^{1-\zeta}\overline{G}^{\zeta}(t)}{L_u}\\ &= g_b\cdot\left(1+\frac{L_s}{L_u}\right)-\phi\left(\frac{G}{L_u}\right)^{\kappa(1-\zeta)+\zeta}\end{aligned}$$

$$\begin{aligned}\frac{\dot{L_s}}{L_s} &= \frac{L_u}{L_s}\phi\frac{\left[L_u^{1-\kappa}\cdot G(t)^{\kappa}\right]^{1-\zeta}\overline{G}^{\zeta}(t)}{L_u}\\ &= \frac{L_u}{L_s}\phi\left(\frac{G}{L_u}\right)^{\kappa(1-\zeta)+\zeta}\\ &= g_b\cdot\left(1+\frac{L_s}{L_u}\right)-\phi\left(\frac{G}{L_u}\right)^{\kappa(1-\zeta)+\zeta}\end{aligned}$$

$$\frac{\phi}{g_b}\left(\frac{G}{L_u}\right)^{\kappa(1-\zeta)+\zeta}=\frac{L_s}{L_u}$$

To characterize the social optimum, we have

$$\begin{aligned}H &= \frac{C(t)^{1-\sigma}-1}{1-\sigma}+\widehat{\lambda}\left[\xi K-\delta\cdot E(C(t))-G(t)\right]+\eta\left[g_b\cdot(L_u+L_s)-\phi\left[L_u^{1-\kappa}\cdot G(t)^{\kappa}\right]^{1-\zeta}\overline{G}^{\zeta}(t))\right]\\ &\quad+\psi\phi\left[L_u^{1-\kappa}\cdot G(t)^{\kappa}\right]^{1-\zeta}\overline{G}^{\zeta}(t)\end{aligned}$$

Substituting $\overline{G}(t) = G(t)$ into the above yields

$$\begin{aligned} H &= \frac{C(t)^{1-\sigma} - 1}{1-\sigma} + \widehat{\lambda}\left[\xi K - \delta \cdot E(C(t)) - G(t)\right] + \eta\left[g_b \cdot (L_u + L_s) - \phi L_u^{(1-\kappa)(1-\zeta)} \cdot G(t)^{\kappa(1-\zeta)+\zeta}\right] \\ &\quad + \psi \phi L_u^{(1-\kappa)(1-\zeta)} \cdot G(t)^{\kappa(1-\zeta)+\zeta}. \end{aligned}$$

All the optimal conditions remain the same as the case of laissez affaire market equilibrium except that (98) becomes the following

$$(\psi - \eta)\,\phi\left[\kappa(1-\zeta) + \zeta\right] L_u^{(1-\kappa)(1-\zeta)} G(t)^{\kappa(1-\zeta)+\zeta-1} = \widehat{\lambda} \tag{108}$$

except into (98) yields

$$\begin{aligned} \frac{\partial H}{\partial C} &= 0 \Rightarrow C(t)^{-\sigma} = \widehat{\lambda}\delta \frac{\partial E(C(t))}{\partial C(t)} \Rightarrow -\sigma \frac{\dot{C}}{C} = \frac{\dot{\widehat{\lambda}}}{\widehat{\lambda}} \\ \frac{\partial H}{\partial G} &= 0 \Rightarrow (\psi - \eta)\,\phi\left[\kappa(1-\zeta) + \zeta\right] L_u^{(1-\kappa)(1-\zeta)} G(t)^{\kappa(1-\zeta)+\zeta-1} = \widehat{\lambda} \\ \dot{\widehat{\lambda}} &= \rho\widehat{\lambda} - \frac{\partial H}{\partial K} = \rho\widehat{\lambda} - \widehat{\lambda}\xi \Rightarrow \frac{\dot{\widehat{\lambda}}}{\widehat{\lambda}} = \rho - \xi \\ \frac{\dot{\psi}}{\psi} &= \rho - \left[\delta \frac{\widehat{\lambda}}{\psi} \frac{a^n(a-\lambda)}{\lambda - 1} + \frac{\eta}{\psi} g_b\right] \\ \dot{\eta} &= \rho\eta - \frac{\partial H}{\partial L_u} \Rightarrow \frac{\dot{\eta}}{\eta} = \rho - g_b - \delta \frac{\widehat{\lambda}}{\eta} \frac{a^{n+1} - a^n}{\lambda^{n+1} - \lambda^n} + \left(1 - \frac{\psi}{\eta}\right) \phi(1-\kappa) L_u^{-\kappa} (1-\zeta) \left[L_u^{1-\kappa} \cdot G(t)^{\kappa}\right]^{-\zeta} \cdot G(t)^{\kappa} \end{aligned}$$

Thus we still have

$$\xi - g_b = \delta \frac{\widehat{\lambda}}{\eta} \frac{a^{n+1} - a^n}{\lambda^{n+1} - \lambda^n} + \left(1 - \frac{\psi}{\eta}\right) \phi(1-\kappa) L_u^{-\kappa} (1-\zeta) \left[L_u^{1-\kappa} \cdot G(t)^{\kappa}\right]^{-\zeta} \cdot G(t)^{\kappa} G^{\zeta}(t)$$

$$\left(\frac{\psi}{\eta} - 1\right) \phi\left[\kappa(1-\zeta) + \zeta\right] G^{\zeta+\kappa-1-\kappa\zeta} L_u^{(1-\kappa)(1-\zeta)} = \frac{\widehat{\lambda}}{\eta}, \tag{112}$$

substituting the above into (100) yields

$$\begin{aligned} \frac{\dot{\eta}}{\eta} &= \rho - g_b - \delta \frac{\widehat{\lambda}}{\eta} \frac{a^{n+1} - a^n}{\lambda^{n+1} - \lambda^n} + \left(1 - \frac{\psi}{\eta}\right) \phi(1-\kappa) L_u^{-\kappa} (1-\zeta) \left[L_u^{1-\kappa} \cdot G(t)^{\kappa}\right]^{-\zeta} \cdot G(t)^{\kappa} G^{\zeta}(t) \\ \frac{\dot{\eta}}{\eta} &= \rho - g_b - \delta \frac{\widehat{\lambda}}{\eta} \frac{a^{n+1} - a^n}{\lambda^{n+1} - \lambda^n} - \frac{\frac{\widehat{\lambda}}{\eta}}{\phi\left[\kappa(1-\zeta) + \zeta\right] G^{\zeta+\kappa-1-\kappa\zeta} L_u^{(1-\kappa)(1-\zeta)}} \phi(1-\kappa) L_u^{-\kappa} (1-\zeta) \left[L_u^{1-\kappa} \cdot G(t)^{\kappa}\right]^{-\zeta} \end{aligned}$$

Using (102) , we further obtain

$$
\begin{aligned}
\xi - g_b &= \frac{\widehat{\lambda}}{\eta}\left[\delta\frac{a^{n+1}-a^n}{\lambda^{n+1}-\lambda^n} + \frac{(1-\kappa)(1-\zeta)G(t)}{[\kappa(1-\zeta)+\zeta]L_u}\right] \\
\frac{\xi - g_b}{\left[\delta\frac{a^{n+1}-a^n}{\lambda^{n+1}-\lambda^n} + \frac{(1-\kappa)(1-\zeta)G(t)}{[\kappa(1-\zeta)+\zeta]L_u}\right]} &= \frac{\widehat{\lambda}}{\eta} \\
\frac{\psi}{\eta} &= \frac{\xi - g_b}{\phi[\kappa(1-\zeta)+\zeta]G^{\zeta+\kappa-1-\kappa\zeta}L_u^{(1-\kappa)(1-\zeta)}\left[\delta\frac{a^{n+1}-a^n}{\lambda^{n+1}-\lambda^n} + \frac{(1-\kappa)(1-\zeta)G(t)}{[\kappa(1-\zeta)+\zeta]L_u}\right]} + 1
\end{aligned}
$$

On the other hand,

$$
\frac{\psi}{\eta} = \frac{\left[\delta\frac{\widehat{\lambda}}{\eta}\frac{a^n(a-\lambda)}{\lambda-1} + g_b\right]}{\xi}
$$

thus

$$
\begin{aligned}
\frac{\xi - g_b}{\phi[\kappa(1-\zeta)+\zeta]G^{\zeta+\kappa-1-\kappa\zeta}L_u^{(1-\kappa)(1-\zeta)}\left[\delta\frac{a^{n+1}-a^n}{\lambda^{n+1}-\lambda^n} + \frac{(1-\kappa)(1-\zeta)G(t)}{[\kappa(1-\zeta)+\zeta]L_u}\right]} + 1 &= \frac{\left[\delta\frac{\widehat{\lambda}}{\eta}\frac{a^n(a-\lambda)}{\lambda-1} + g_b\right]}{\xi} \\
\frac{\xi - g_b}{\phi[\kappa(1-\zeta)+\zeta]G^{\zeta+\kappa-1-\kappa\zeta}L_u^{(1-\kappa)(1-\zeta)}\left[\delta\frac{a^{n+1}-a^n}{\lambda^{n+1}-\lambda^n} + \frac{(1-\kappa)(1-\zeta)G(t)}{[\kappa(1-\zeta)+\zeta]L_u}\right]} &= \frac{\delta\frac{\xi-g_b}{\left[\delta\frac{a^{n+1}-a^n}{\lambda^{n+1}-\lambda^n}+\frac{(1-\kappa)(1-\zeta)G(t)}{[\kappa(1-\zeta)+\zeta]L_u}\right]}\frac{a^n(a-\lambda)}{\lambda-1} + }{\xi} \\
\frac{\xi}{\phi[\kappa(1-\zeta)+\zeta]G^{\zeta+\kappa-1-\kappa\zeta}L_u^{(1-\kappa)(1-\zeta)}} &= \delta\frac{a^n(a-\lambda)}{\lambda-1} - \delta\frac{a^{n+1}-a^n}{\lambda^{n+1}-\lambda^n} - \frac{(1-}{[\kappa(} \\
\frac{\xi}{\phi[\kappa(1-\zeta)+\zeta]}\left[\frac{G}{L_u}\right]^{(1-\kappa)(1-\zeta)} &= \delta\left[\frac{a^n(a-\lambda)}{\lambda-1} - \frac{a^{n+1}-a^n}{\lambda^{n+1}-\lambda^n}\right] - \frac{(1}{[\kappa} \\
\frac{\xi}{\phi[\kappa(1-\zeta)+\zeta]G^{\zeta+\kappa-1-\kappa\zeta}L_u^{(1-\kappa)(1-\zeta)}} &= \delta\left[\frac{a^n(a-\lambda)}{\lambda-1} - \frac{a^{n+1}-a^n}{\lambda^{n+1}-\lambda^n}\right] - \frac{(1}{[\kappa}
\end{aligned}
$$

From (rr) and(ff), we obtain

$$
\xi - g_b + \left(1 - \frac{\psi}{\eta}\right)\phi(1-\zeta)\left[(1-\kappa)L_u^{-\zeta(1-\kappa)-\kappa}\cdot G(t)^{\kappa+\zeta-\zeta\kappa} + \delta\kappa G^{\zeta+\kappa-1-\kappa\zeta}L_u^{(1-\kappa)(1-\zeta)}\frac{a^{n+1}-a^n}{\lambda^{n+1}-\lambda^n}\right] = 0 \tag{113}
$$

which implies that

$$
\frac{\xi - g_b}{\phi(1-\zeta)\left[(1-\kappa)L_u^{-\zeta(1-\kappa)-\kappa}\cdot G(t)^{\kappa+\zeta-\zeta\kappa} + \delta\kappa G^{\zeta+\kappa-1-\kappa\zeta}L_u^{(1-\kappa)(1-\zeta)}\frac{a^{n+1}-a^n}{\lambda^{n+1}-\lambda^n}\right]} = \left(\frac{\psi}{\eta} - 1\right)
$$

thus

$$
\begin{aligned}
\frac{\widehat{\lambda}}{\eta} &= \left(\frac{\psi}{\eta} - 1\right)\phi[\kappa(1-\zeta)+\zeta]G^{\zeta+\kappa-1-\kappa\zeta}L_u^{(1-\kappa)(1-\zeta)} \\
&= \frac{[\kappa(1-\zeta)+\zeta]G^{\zeta+\kappa-1-\kappa\zeta}L_u^{(1-\kappa)(1-\zeta)}}{(1-\zeta)\left[(1-\kappa)L_u^{-\zeta(1-\kappa)-\kappa}\cdot G(t)^{\kappa+\zeta-\zeta\kappa} + \delta\kappa G^{\zeta+\kappa-1-\kappa\zeta}L_u^{(1-\kappa)(1-\zeta)}\frac{a^{n+1}-a^n}{\lambda^{n+1}-\lambda^n}\right]}(\xi - g_b)
\end{aligned}
$$

$$
\begin{aligned}
\xi - g_b &= \frac{\frac{\widehat{\lambda}}{\eta}}{\left[\kappa\left(1-\zeta\right)+\zeta\right] G^{\zeta+\kappa-1-\kappa\zeta} L_u^{(1-\kappa)(1-\zeta)}} \left(1-\zeta\right) \Bigg[\left(1-\kappa\right) L_u^{-\zeta(1-\kappa)-\kappa} \cdot G(t)^{\kappa+\zeta-\zeta\kappa} + \delta\kappa G^{\zeta+\kappa-1-\kappa\zeta} L_u^{(1-} \\
\xi - g_b &= \frac{\frac{\left[\kappa(1-\zeta)+\zeta\right] G^{\zeta+\kappa-1-\kappa\zeta} L_u^{(1-\kappa)(1-\zeta)}}{(1-\zeta)\left[(1-\kappa) L_u^{-\zeta(1-\kappa)-\kappa} \cdot G(t)^{\kappa+\zeta-\zeta\kappa} + \delta\kappa G^{\zeta+\kappa-1-\kappa\zeta} L_u^{(1-\kappa)(1-\zeta)} \frac{a^{n+1}-a^n}{\lambda^{n+1}-\lambda^n}\right]} \left(\xi - g_b\right)}{\left[\kappa\left(1-\zeta\right)+\zeta\right] G^{\zeta+\kappa-1-\kappa\zeta} L_u^{(1-\kappa)(1-\zeta)}} \left(1-\zeta\right) \Bigg[\left(1-\kappa\right) L_u^{-\zeta(1-\kappa)-\kappa} \cdot G \\
1 &= \frac{\frac{\left[\kappa(1-\zeta)+\zeta\right] G^{\zeta+\kappa-1-\kappa\zeta} L_u^{(1-\kappa)(1-\zeta)}}{(1-\zeta)\left[(1-\kappa) L_u^{-\zeta(1-\kappa)-\kappa} \cdot G(t)^{\kappa+\zeta-\zeta\kappa} + \delta\kappa G^{\zeta+\kappa-1-\kappa\zeta} L_u^{(1-\kappa)(1-\zeta)} \frac{a^{n+1}-a^n}{\lambda^{n+1}-\lambda^n}\right]}}{\left[\kappa\left(1-\zeta\right)+\zeta\right] G^{\zeta+\kappa-1-\kappa\zeta} L_u^{(1-\kappa)(1-\zeta)}} \left(1-\zeta\right) \Bigg[\left(1-\kappa\right) L_u^{-\zeta(1-\kappa)-\kappa} \cdot G(t)^{\kappa+\zeta-}
\end{aligned}
$$

Appendix

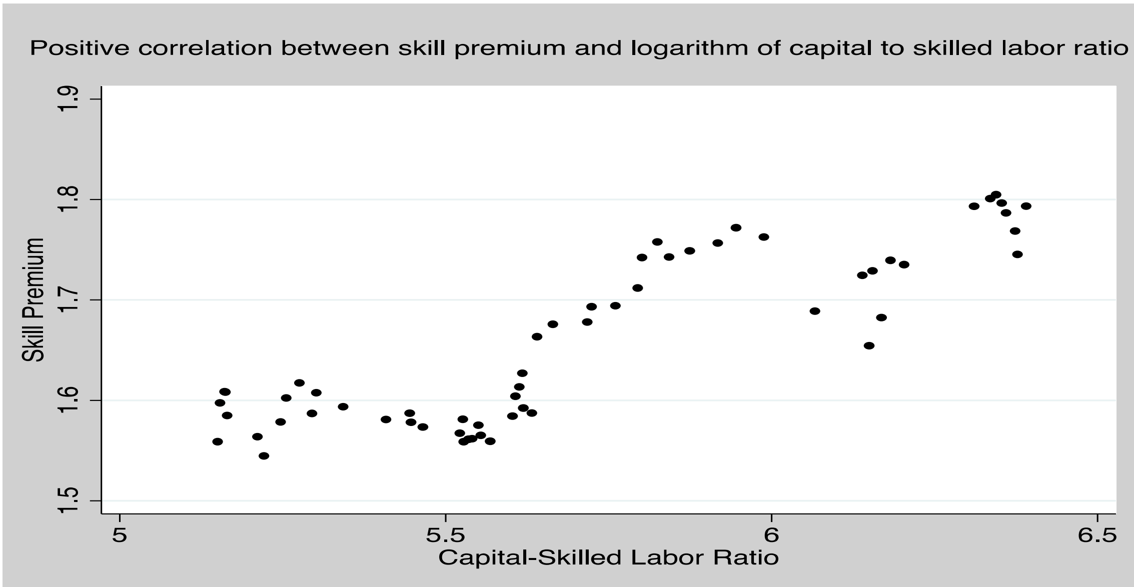


Figure 1: Positive correlation between skill premium and logarithm of capital to skilled labor ratio(USA:1958-2016). The skilled labor is measured by the non-production workers and the unskilled-labor is measured by the production workers. Capital and Labor are measured in 1$m and 1000s. Date source: NBER-CES manufacturing database.

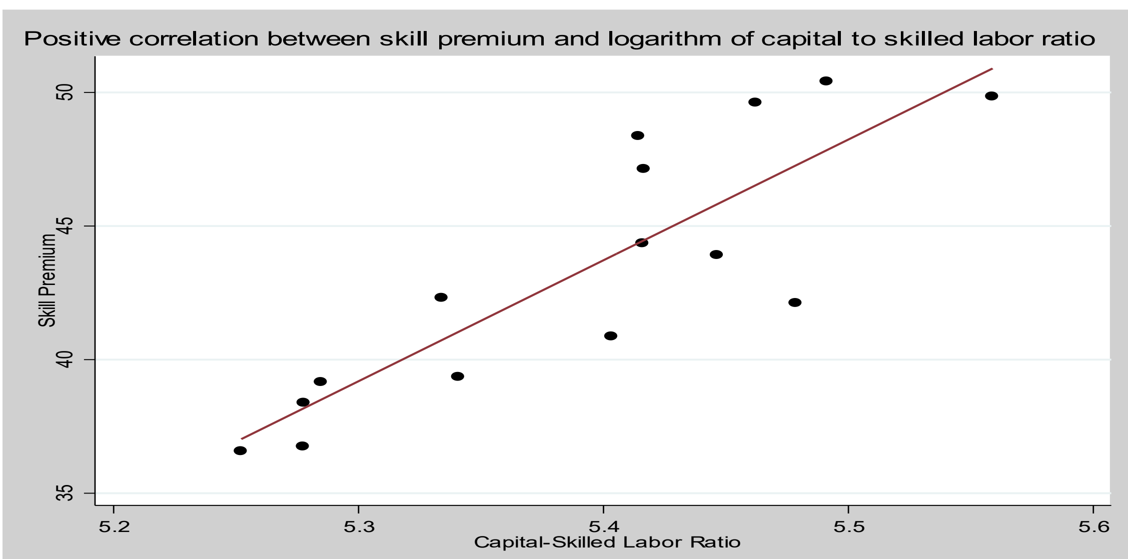


Figure 1A: Positive correlation between skill premium and logarithm of capital to skilled labor ratio(USA:1995-2009). In WIOD (SEA) skill type is defined on the basis of the level of educational attainment of the worker. We define the labor force with first stage of tertiary education or above as skilled labor, and the labor force with post-secondary non-tertiary education or below as unskilled labor. See Table 6 for more classification details. Capital is measured in millions of Dollars (1995=100), skilled labor is measured in millions of hours. Date source: WIOD Socioeconomic Accounts (SEA) manufacturing database.

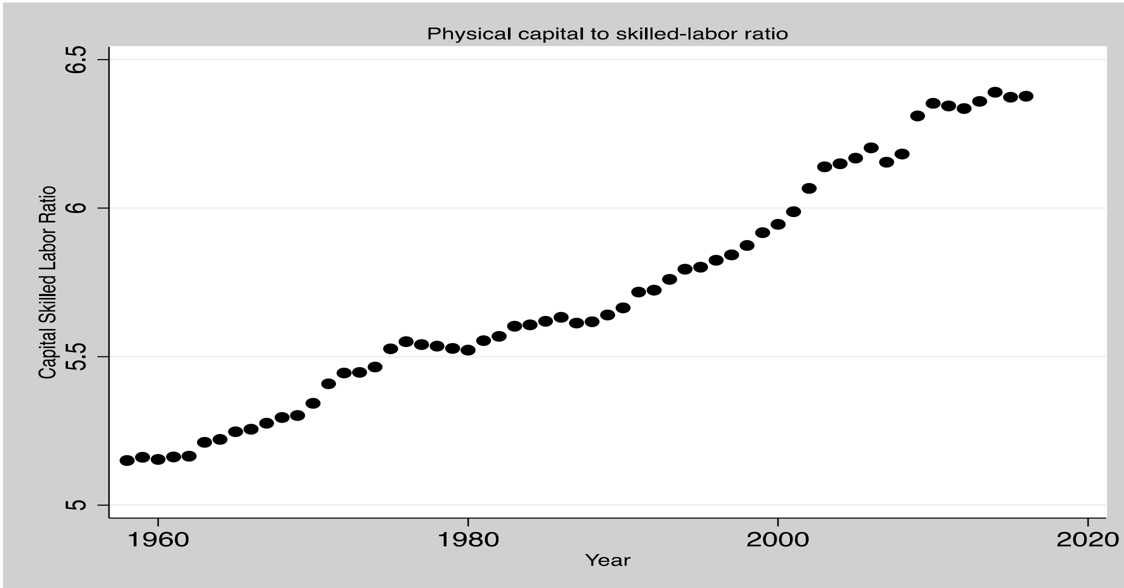


Figure 2: Physical capital to skilled labor ratio(USA:1958-2011). The skilled labor is measured by the non-production workers and the unskilled-labor is measured by the production workers. Capital and Labor are measured in 1$m and 1000s.Date source: NBER-CES manufacturing database.

Table 1. Cross-industry heterogeneity in Capital-Skilled labor Ratio in the USA

Year	**Mean**	**Std. Dev.**	**Min**	**Max**	**Max/Min**
1958	308.736	285.915	24.730	1638.108	66.240
1968	324.607	318.161	23.378	1931.087	82.603
1978	365.970	387.591	20.615	2182.396	105.865
1988	377.415	428.646	14.807	2862.787	193.340
1998	417.286	521.639	37.641	3629.678	96.429
2008	459.676	550.501	55.355	3856.926	69.676
2016	485.734	602.188	64.502	4388.876	68.043

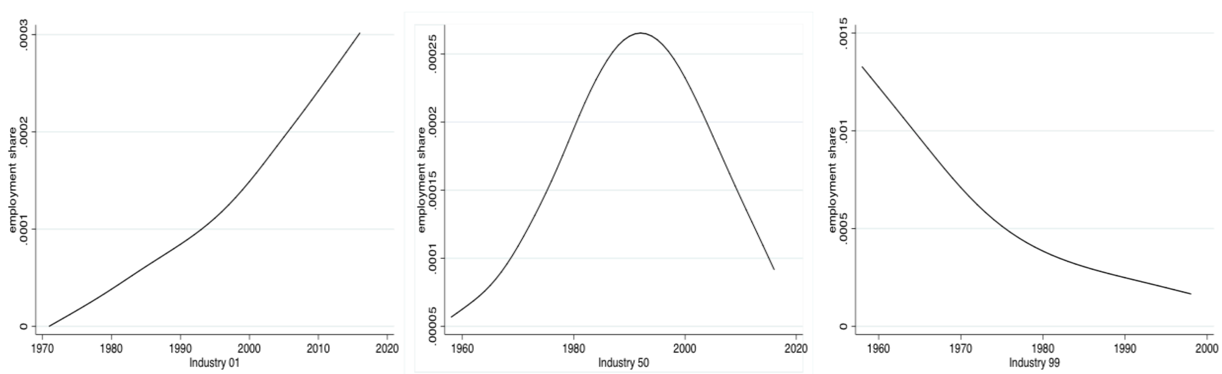
Note: Redefining the industries according to the capital-skilled labor ratios, we first rank all the 27751 observations by capital-skilled labor ratios in each firm in an increasing order, and then equally divide all these observations into 99 bins. Capital and skilled labor are measured in 1$m and 1000s. Date source: NBER-CES manufacturing database.

Table 1A. Cross-industry heterogeneity in Capital-Skilled labor Ratio in the USA (originally industries)

Year	Mean	Std. Dev.	Min	Max	Max/Min
1958	180.972	195.641	2.733	1638.108	599.381
1968	227.514	227.284	8.695	1931.087	222.092
1978	292.623	300.318	14.140	2425.79	171.556
1988	369.154	426.240	14.807	3845	259.692
1998	461.395	559.745	37.641	5657.706	150.307
2008	675.076	769.884	53.390	6301.445	118.027
2016	756.885	952.820	64.502	9294.75	144.100

Note: Capital and skilled labor are measured in 1$m and 1000s. Date source: NBER-CES manufacturing database

HP- filtered employment shares of three typical newly-defined "Industries"



Figuer 3: HP- filtered employment shares of three typical "Industries" (Newly Defined) in The USA(1958-2016). The horizontal axis is the year, and the vertical axis is the employment share of "Industry 1 (50 and 99)" in all manufacturing industries. If there is a missing value in the constructed time series for a certain year, that missing value is replaced by the simple average of the observations immediately before and after that year. The HP filter parameter λ is set to 100.

Table 2. Cross-industry heterogeneity in Capital-Skilled labor Ratio in the USA

Dependent variable	Employment share*10000		Value-added share*10000	
	(1)	(2)	(3)	(4)
t	1.547^{***}	1.547^{***}	45.711^{***}	45.710^{***}
	0.312	0.312	18.269	18.269
t^2	-0.0004^{***}	-0.0004^{***}	-0.012^{***}	-0.012^{***}
	0.0001	0.0001	0.005	0.005
t*k	0.0001^{***}	0.0001^{***}	0.003^{***}	0.003^{***}
k	-0.100^{***}	-0.099^{***}	-6.436^{***}	-6.429^{***}
	0.007	0.007	0.390	0.391
T	0.0004	0.0004	0.667^{***}	0.667^{***}
	0.001	0.001	0.029	0.029
GDPGR	0.006	0.006	-0.452	-0.452
	0.009	0.009	0.542	0.542
Constant	-1505.222^{***}	-1508.14^{***}	-42022.2^{**}	-42045.82^{**}
	309.340	309.321	18126.74	18125.65
Industry dummies	yes	yes	yes	yes
Observation	5416	5416	5416	5416
R-squared	0.16	0.16	0.34	0.34

Note: *t;k;T* and *GDPGR* represent year, average capital-skilled labor ratio, labor productivity and GDP growth rate, respectively. Standard errors in parentheses. **Indicates significance at 5% level, ***Indicates significance at 1% level.

Table 3. Peak time of industries in The USA

Dependent variable	Peak time of Employment share*10000		Peak time of Value-added share*10000	
	(1)	(2)	(3)	(4)
Capital-skilled labor ratio	0.0001^{***}		0.172^{***}	
	6.32e-13		7.77e-10	
Capital expenditure ratio		0.0001^{***}		0.003^{***}
		4.33e-08		6.22e-06
Constant	0.00002^{***}	0.294^{***}	0.052^{***}	62.794^{***}
	1.73e-10	2.85e-09	2.54e-07	4.11e-07
R-squared	1	0.9	1	0.9
Observation	99	99	99	99

In Table 3, standard errors in parentheses. ***Denotes significance at the 1% level.

Table 4. Cross-industry heterogeneity in Capital-Skilled labor Ratio in the USA

Dependent variable	Value-added share			
	(1)	(2)	(3)	(4)
Congruence term	-2.365^{***}	-2.463^{***}	-0.153^{***}	-0.218^{***}
	0.234	0.246	0.013	0.023
T	0.030^{***}	0.020^{***}	0.003^{***}	0.004^{***}
	0.002	0.002	0.0001	0.0001
Constant	53.180^{***}	23.622^{***}	0.806^{***}	1.007^{***}
	2.983	1.622	0.321	0.364
Industry dummies	yes	yes	yes	yes
Observation	5416	5335	27367	21112
R-squared	0.24	0.22	0.82	0.82

Note: Congruence term is the absolute value of a normalized difference newly-defined sector is capital-skilled labor ratio and the aggregate capital-skilled labor ratio in manufacturing sectors at year t. The skilled labor is represented by their non-production workers. T is the labor productivity of industries i of year t. In Table 4, standard errors in parentheses. ***Indicates significance at 1% level.

Table 5. Asymmetric Duration fact in The USA

θ	P Newly-defined industries	P Original 6-digit industries
5	0.829	0.810
10	0.833	0.808
15	0.828	0.806
20	0.858	0.803
25	0.909	0.803
30	0.882	0.819
35	0.858	0.790
40	0.909	0.802
45	0.875	0.814
50	0.800	0.825

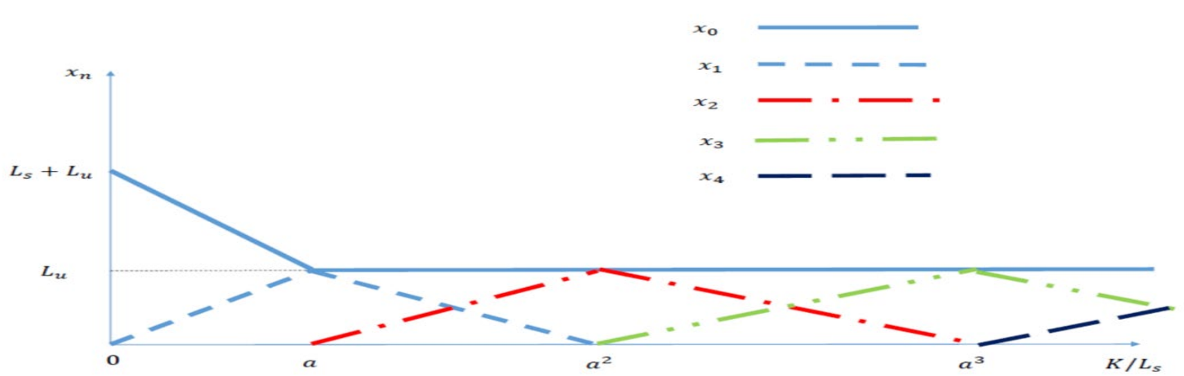
Table 6. Definition of skills

Skill Type	1997 ISCED Level	1997 ISCED Level description
Low	1	Primary education or first stage of basic education
Low	2	Lower secondary or second stage of basic education
Low	3	(Upper) secondary education
Low	4	Post secondary non-tertiary education
High	5	First stage of tertiary education
High	6	Second stage of tertiary education

Source: ISCED, http://www.uis.unesco.org/Education/Pages/

Table 7: Quantities in Static Equilibrium

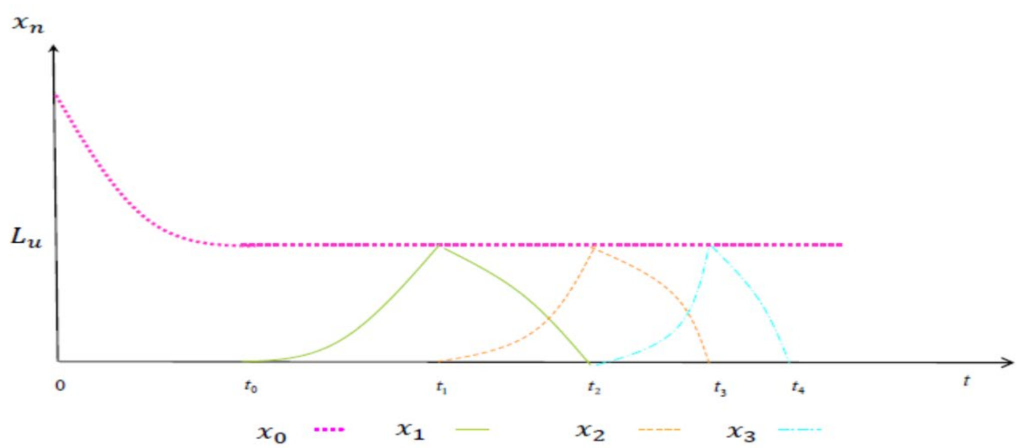
$0 \leq K < aL_s$	$a^n L_s \leq K < a^{n+1} L_s$ for $n \geq 1$
$x_0 = L_u + L_s - \frac{K}{a}$	$x_0 = L_u$
$x_1 = \frac{K}{a}$	$x_n = \frac{L_s a^{n+1} - K}{a^{n+1} - a^n}$
	$x_{n+1} = \frac{K - a^n L_s}{a^{n+1} - a^n}$
$x_j = 0$,for $\forall j \neq 0, 1$	$x_j = 0$,for $\forall j \neq 0, n, n+1$
$X = L_u + L_s + \frac{\lambda - 1}{a} K$	$X = L_u + \frac{\lambda^n (a - \lambda)}{a - 1} L_s + \frac{\lambda^{n+1} - \lambda^n}{a^{n+1} - a^n} K$



Figuer 4: Static Equilibrium

Table 8: Prices and Factor Income Shares in Static Equilibrium

$0 \le K < aL_s$	$a^n L_s \le K < a^{n+1} L_s$ for $n \ge 1$
$\frac{w_s}{w_u} = 1$	$\frac{w_s}{w_u} = \frac{\lambda^n(a-\lambda)}{a-1}$
$\frac{r}{w_s} = \frac{\lambda-1}{a}$	$\frac{r}{w_s} = \frac{\lambda-1}{a^n(a-\lambda)}$
$p_0 = w_u = 1$	$p_0 = w_u = 1$
$p_1 = \lambda$	$p_n = \lambda^n$
	$p_{n+1} = \lambda^{n+1}$
$\theta = \frac{L_u+L_s}{L_u+L_s+\frac{\lambda-1}{a}K}$	$\theta = \frac{L_u+\frac{\lambda^n(a-\lambda)}{a-1}L_s}{L_u+\frac{\lambda^n(a-\lambda)}{a-1}L_s+\frac{\lambda^{n+1}-\lambda^n}{a^{n+1}-a^n}K}$

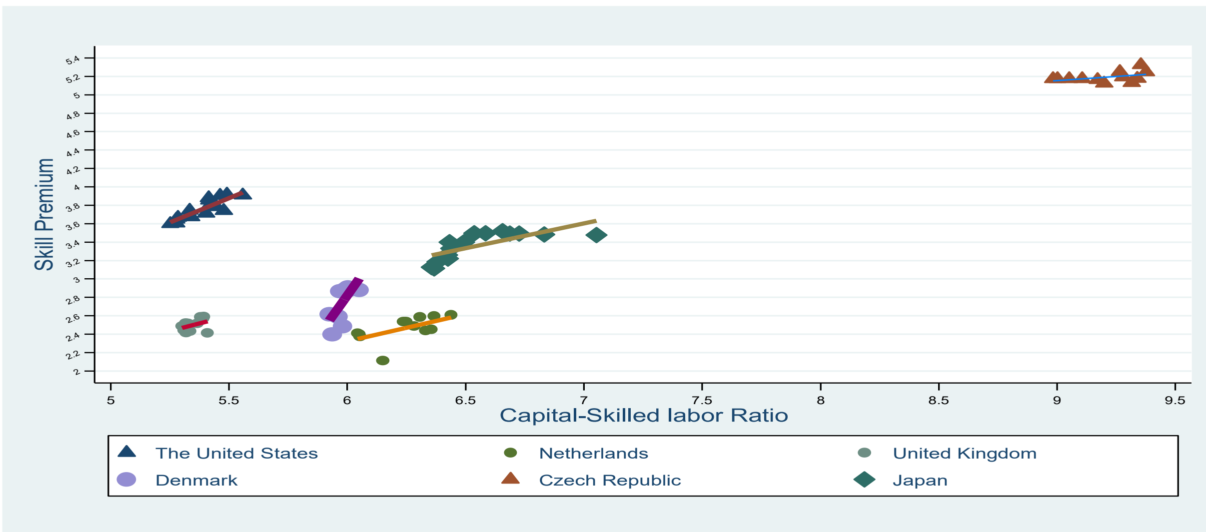


Figuer 5: Dynamic Equilibrium

Evidence from cross-country data

Figure 1B shows the relationship between capital-skilled labor ratio and skill premium in six countries. A positive correlation between skill premium and the ratio of capital to skilled labor is discernible.

Figure1B: Capital-Skilled Labor Ratio and Skill Premium in Six Countries: 1995-2009



Note: Log Capital-Skilled Labor Ratio vs. log Skill Premium. Capital is measured in millions of Dollars (1995=100), skilled Labor is measured in millions of hours. Countries: Denmark(1995-2007), Czech Republic(1995-2007), United Kingdom(1995-2007), Japan(1995-2009), United States(1995-2009) and Netherlands(1995-2007).Data source: WIOD Socio-Economic Accounts (SEA) manufacturing database.

Table 1B shows that there exists tremendous cross-industry heterogeneity in the capital-skilled labor ratio in six countries.

Country	Year	Mean	Std. Dev.	Min	Max	Max/Min
Czech Republic	1995	6200.321	4670.369	1636.397	17870.76	10.921
	2007	11721.6	15534.95	3546.543	63806.07	17.991
United Kingdom	1995	148.262	58.32756	88.45324	284.9299	3.22
	2007	136.0983	39.85463	96.34043	224.1298	2.326
Denmark	1995	4788.203	9139.143	1111.355	36293.59	32.747
	2007	4488.754	7451.816	1496.27	30203.3	20.186
Japan	1995	128253.7	280000	27705.99	1097037	39.600
	2009	210491.3	404641.6	54544.62	1609794	29.513
	1995	620.011	698.2273	164.7473	2876.205	17.458

Netherlands	2007	413.3582	412.5819	97.98441	1715.209	17.505
United States	1995	246.5035	152.9504	131.3463	723.6995	5.05
	2009	328.9353	201.7134	192.9896	992.5255	5.143

Table 1B Cross-industry heterogeneity in Capital-Skilled labor Ratio Cross countries: 1995-2009

Note: Capital is measured in millions of national currency (1995=100), and the skilled labor is measured in millions of hours. 14 industries: Food, beverages and tobacco (15&16);Textiles and textile (17&18);Leather, leather and footwear (19);Wood and of wood and Cork (20);Pulp, Paper, Printing and Publishing (21&22);Coke, refined petroleum and nuclear fuel (23); Chemicals and chemical (24);Rubber and plastics (25); Other non-Metallic Mineral (26);Basic metals and fabricated metal (27&28);Machinery, Nec (29);Electrical and optical equipment (30&33);Transport Equipment (34&35);Manufacturing Nec; Recycling (36&37).Data source: WIOD Socio-Economic Accounts (SEA) manufacturing database.

To further investigate the cross-country empirical evidence for industry dynamics, we extend regression (1) to a cross-country regression:

$$Y_{itc}=\beta_0+\beta_1 t+\beta_2 t^2+\beta_3.k_{itc}.t+\beta_4.T_{itc}+\beta_5.D_{ic}+\beta_6 GDPGR_{tc}+\varepsilon_{itc}$$

where subscript c represents a country index. The results are summarized in Table2A.

Table 2A. Hump-shaped pattern of industrial dynamics cross countries: 1995-2009

Dependent variable	Employment share(1) *1000		Employment share（2）*1000	
	(1)	(2)	(3)	(4)
	(5.597)	(5.595)	(5.597)	(5.204)
t^2	-0.003**	-0.003**	-0.001**	-0.003**
	(0.001)	(0.001)	(0.001)	(0.001)
t*k	3.92e-07***	3.92e-07***	1.00e-07***	3.92e-07***
k	-0.001***	-0.001***	-0.001***	0.001***
T	0.004**	0.004**	0.004**	0.004**
	(0.002)	(0.002)	(0.002)	(0.002)
GDPGR	0.042***	0.065***	0.042***	0.042***
	(0.012)	(0.011)	(0.012)	(0.012)
Constant	-13622.85**	-13622.85**	-13718.93***	-13622.85**
	(5600.983)	(5598.351)	(5600.983)	(5598.351)
Country*Industry dummies	yes	yes	yes	yes
Observation	992	992	992	992
R-squared	0.973	0.973	0.973	0.973

Note: t; k; T and GDPGR represent Year, average capital-skilled labor ratio or capital labor ratio of an industry, Labor productivity of an industry, and GDP growth rate for each specific country, respectively. Employment share (1) is the share of hours worked by persons engaged, employment share (2) is the share of persons engaged. Countries: Denmark(1995-2007), Czech Republic(1995-2007), United Kingdom(1995-3007), Japan(1995-2009), United States(1995-2009) and Netherlands(1995-2007).Data source: WIOD Socio-Economic Accounts (SEA) (1995–2009).* p < 0.5, ** p < 0.05, *** p < 0.005.

We regress the peak time of a industry's share (either employment share or output share) on its capital-skilled labor ratio and capital-labor ratio. The results are in Table 3A. Column (1) and column (3) show that the peak time of an industry is positively correlated with its capital-skilled labor ratio. Column (2) and column (4) confirm the finding in Ju, Lin and Wang (2015).

Table3A. Hump-shaped pattern of industrial dynamics cross countries: 1995-2009

Dependent variable	Peak time of Employment Share(1)		Peak time of Employment Share(2)	
	(1)	(2)	(3)	(4)
Capital-skilled labor ratio	7.14e-12***		7.15e-12***	
	(1.43e-14)		(5.53e-15)	
Capital-labor ratio		6.70e-11***		6.67e-12***
		(7.84e-15)		(1.31e-14)
Constant	0.056***	0.055***	0.056***	0.055***
	(1.49e-10)	(1.28e-10)	(9.86e-11)	(1.22e-10)
R-squared	0.999	1.000	0.999	0.999
Observation	84	84	84	84

Note: Employment share (1) is the share of hours worked by persons engaged, employment share (2) is the share of persons engaged.Countries: Denmark(1995-2007), Czech Republic(1995-2007), United Kingdom(1995-3007), Japan(1995-2009), United States(1995-2009) and Netherlands(1995-2007).Data source: WIOD Socio-Economic Accounts (SEA) (1995–2009).Standard errors in parentheses * p < 0.05, ** p < 0.01, *** p < 0.001

The findings are consistent with the results in Table2 and Table 3, which suggests that the industry dynamics patterns observed in the US are actually quite general and also true for other countries.

Next we extend regression (2) to a cross-country regression:

$LS_{itc}=\beta_0+\beta_1\left|\frac{K_{itc}/L_{sitc}-K_{tc}/L_{stc}}{K_{tc}/L_{stc}}\right|+\beta_2T_{itc}+Country*industry_{dummy}+\varepsilon_{it}$, (4) where subscript c represents a country index. The results are summarized in Table 4A.

Table4A. Congruence facts cross countries: 1995-2009

Dependent variable	Value-added share *1000	Employment share*1000
	（1）	（2）
	(0.044)	(0.015)
T	0.040***	0.001*
	(0.001)	(0.0004)
Constant	5.429***	8.203***
	(0.687)	(0.232)
Country*Industry	yes	yes
Observation	1160	1160
R-squared	0.840	0.981

Note: Congruence term is the absolute value of a normalized difference sector *i*'s capital-skilled labor ratio and the aggregate capital-skilled labor ratio in manufacturing sectors in country c at year *t*. T_{it} is the labor productivity of

industries *i* of year *T*.Countries: Denmark(1995-2007), Czech Republic(1995-2007), United Kingdom(1995-3007), Japan(1995-2009), United States(1995-2009) and Netherlands(1995-2007).
Data source: WIOD Socio-Economic Accounts (SEA) (1995–2009).
Standard errors in parentheses* p < 0.5, ** p < 0.05, *** p < 0.001

The findings are consistent with the results in Table4, which suggests that the congruence facts observed in the US are actually quite general and also true for other countries.