PEKING UNIVERSITY
Institute of New Structural Economics

Working Paper Series of New Structural Economics

No. E2019009 **2019-12-02**

Human Capital, Industrial Dynamics and Skill Premium*

Yong Wang† Xuemin Tang‡

Institute of New Structural Economics, Peking University

November 2019

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 intensities. Closed-form solutions are obtained to fully characterize the endogenous dynamic changes of factor endowment structure, industrial composition, life-cycle pattern of each industry and the skill premium along the whole growth path. We highlight that (1) optimal human capital investment should be stage dependent and match the skill demand from the endogenously switching industrial structures and that (2) the aggregate skill premium and its dynamics are endogenously determined in the process of industrial dynamics at the disaggregated level. Our model features endogenous structural differences across different stages of development, which is essentially ignored in the existing 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, Yao Wen, Daniel Xu, Akos Valentiyi and Xiaodong Zhu for their valuable suggestions and comments. All potential deficiencies are our own.

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

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

1. Introduction

This paper has two interrelated primary goals. The first goal is to explore how the level and dynamics of skill premium at the aggregate level is logically related to the industrial structures and industrial dynamics at different stages of economic development. By industrial structures, we refer to the endogenous composition of different industries at the disaggregated level. Industrial dynamics means the life-cycle dynamics of an industry: when and how an industry enters the market, booms, reaches its peak, declines and eventually exits the market (see Klepper,1996, 1997). To guide and discipline our theoretical exploration, we document several stylized facts about skill premium, industrial structures and industrial dynamics using both US and cross-country data (Section 2).

There are in general two main approaches to study skill premium in the pertinent macro development 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, Krusell *et al.*, 2000, Stokey, 1996). 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 this 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 due to lack of micro foundations. First, both approaches make the *ad hoc* assumption that the mathematical form of the aggregate production function is exogenously given and time-invariant. However, the functional form of the aggregate production function could be endogenous to the composition of underlying industries at the disaggregated level, which in turn could be endogenously different at different levels of development (Ju, Lin and Wang, 2015). Second, these two approaches both attribute skill premium differences entirely to the *quantitative* differences in aggregate factors or aggregate technology, while keeping agnostic about whether and how skill premium (and its dynamics) at the aggregate level is related to the *structural* differences in terms of composition (and shifts) of underlying industries. Consequently, those aggregate models are unable to distinguish, quantitatively or qualitatively, roles of a wide array of micro-level frictions of different natures such as industry-specific policies, because all these micro-level structural differences are lumped together as quantitative differences in one homogeneous exogenous aggregate variable. Moreover, these aggregate models cannot (and are not designed to) explain the stylized facts about skill premium and industrial dynamics at the disaggregated level.

The second goal is to explicitly explore optimal human capital investment when industrial structures and industrial dynamics are endogenously different at different levels of economic development. The pertinent growth literature mostly studies human capital investment in structureless frameworks by assuming a time-invariant functional form of aggregate production function and time-invariant sectorial compositions (mostly one-sector models) for all levels of development (see, for example, Lucas, 1988, Mankiw, Romer and Weil, 1992, Barro and Sala-i-Martin, 1996, *etc.*). 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. Notable exceptions include Buera, Kaboski and Rogerson (2018), where skill intensities are asymmetric across sectors. However, they do not investigate the interactions between physical and human capital investment or the link between human capital investment and life-cycle dynamics of endogenously switching industries. As a result, the existing literature fails to tell us explicitly how factor endowment structures could shape the composition of different industries and hence determine the functional form of aggregate production function, how the investment in human capital and physical capital could drive the life-cycle industrial dynamics, changes in skill premium, and aggregate economic growth, and how human capital investment decisions should be optimally made at different stages of economic development when industrial structures potentially change over time. In our model, expectations for future skill premium, as determined by the evolution of industrial structures, would affect decentralized decisions on human capital investment, which in turn feedbacks on the future endowment structures and hence industrial structures and skill premium. Meanwhile, human capital investment should be also synchronized with the physical capital accumulation due to capital-skill complementarity. Our model highlights that human capital investment should dynamically match the industrial structures as demand for skilled and unskilled labor is heterogeneous across industries. Consequently, optimal human capital investment is stage-dependent as industrial structures endogenously evolves as the economy grows.

Using both the US and cross-country disaggregate industry level data, we first document six stylized facts about skill premium, factor endowment, industrial heterogeneity, and life-cycle dynamics of industries (Section 2). 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 and skill intensities and 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 the static model how the given endowment structure determines the optimal industrial structure, endogenous aggregate production function, and all the factor prices including skill premium simultaneously (Section 3). Then based on this static setting, we further develop a dynamic framework in which both physical capital and human capital investment are endogenously made so that the factor endowment structures evolves endogenously over time. 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. We examine how the change in endowment structures drive the dynamic changes in industrial compositions, life-cycle dynamic patterns of each underlying industries as well as the changes of the skill premium along the aggregate growth path (Section 4). In particular, we show how the dynamic model could explain all the stylized facts documented in Section 4 simultaneously. More specifically, we first only allow for endogenous accumulation of physical capital without human capital investment, so the composition of skilled labor and unskilled labor remains time-invariant (Section 4.1). Then we allow for both endogenous accumulation of physical capital and endogenous human capital investment, so that each of the three production factors changes endogenously (Section 4.2). We show that the optimal human capital investment should be stage-dependent in the sense that it should be consistent with the demand for skills from the underlying industries, which are in turn endogenously switching and evolving in response to the changes in factor endowment structures. Skill premium is shown to change at the same frequency as the industrial compositions. Section

5 concludes.

This paper generalizes the two-factor model in Ju, Lin and Wang (2015) to a three-factor case. In Ju, Lin and Wang (2015), there is one type of labor and one type of capital (physical capital), whereas here we divide labor into skilled labor and unskilled labor and we introduce both physical capital and human capital, which enables us to examine skill premium and human capital investment in the context of industrial dynamics and economic growth. Ju, Lin and Wang (2015) is a special case of the current model when all labor is skilled labor. Another important difference is that Ju, Lin and Wang (2015) 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 well documented 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. Similar to Ju, Lin, Wang (2015), the driving force of structural change in this paper is also changes in factor endowment, so we call it endowment-driven structural change, which is different from the other mechanisms of structural change such as non-homothetic preferences, unbalanced productivity growth or international trade (see Kongasmut, Rebelo and Xie, 2001; Ngai and Pissarides, 2007; Matsuyama, 2009; Uy, Yi and Zhang, 2013; Lin and Wang, 2019).

2. Stylized Facts

In this section, we document the following stylized facts about skill premium, factor endowments and industrial dynamics that are found in the US and other countries.

- **Fact 0 (positive correlation):** There exists a positive correlation between skill premium and physical capital to skilled-labor ratio for the aggregate manufacturing sector.
- **Fact 1 (cross-industry heterogeneity):** There exists tremendous cross-industry heterogeneity in capital to skilled-labor 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 share is the industry.
- **Fact 5 (shakeout):**The booming period of an industry 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 475 sub-industries within the manufacturing sector from 1958 to 2011. At this disaggregated level, the rank in terms of capital-skilled labor ratio between two industries is frequently reversed over time, which creates challenges for us to confront data with models that typically 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 25,386 observations consisting of 469 industries for 54 years (7 industries with 156 observations are dropped due to missing values in employment) by the capital–skilled labor 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 257 observations. By construction, the capital to skilled-labor ratio is the lowest in the first bin, called "industry 1" and the ratio is highest in bin "industry 99". Moreover, the rank of capital to skilled-labor 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 2011. 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.[1]

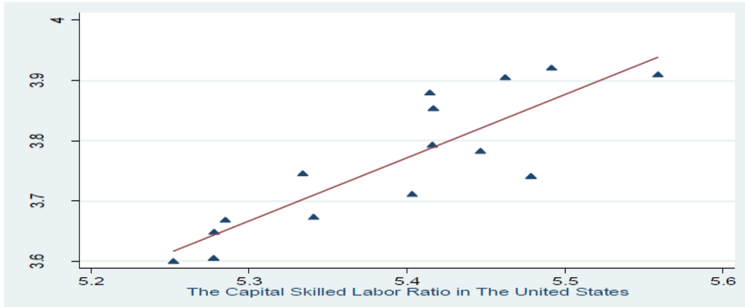
Log of (Capital to Skilled Labor Ratio) and Skill Premium in the USA: 1958-2011

Note: The skilled labor is measured by the production workers and the unskilled-labor is measured by the non-production workers. Capital and Labor are measured in 1$m and 1000s.
Data source: NBER-CES manufacturing database.

[1] 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 1958-2011. 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 1A 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.

Log of (Capital to Skilled-Labor Ratio) and Skill Premium in the USA: 1995-2009

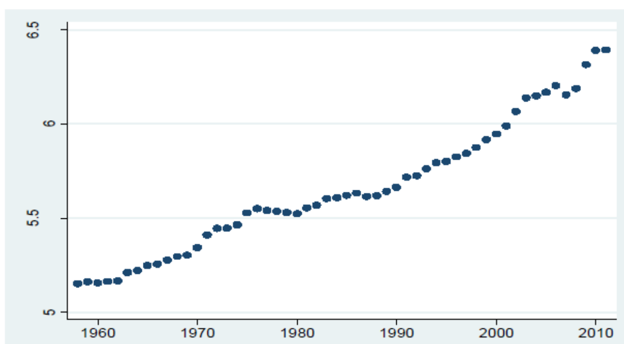


Note: Capital is measured in millions of Dollars (1995=100), Skilled Labor is measured in millions of hours.

Data source: WIOD Socio-Economic Accounts (SEA) manufacturing database.

Figure 2 shows that the logarithm of physical capital to skilled-labor ratio increases over the period from 1958 to 2011.

Log of (Capital to Skilled-Labor Ratio) within Manufacturing Sector in the USA: 1958-2011



Note: The Skilled Labor is measured by the production workers and the unskilled-labor is measured by the non-production workers. Capital and Skilled Labor are measured in 1\$m and 1000s.

Data Source: NBER-CES Database

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 1638.108 US dollars per worker, which is 67.326 times larger than the lowest one in the same year. In 2011, the highest capital to skilled-labor ratio is still about 44 times higher than the lowest one. The standard deviation across industries is about 285 in 1958, and monotonically increases in each decade, reaching 542 in 2011.[2]

[2] If we do the same exercises based on the 469 originally defined industries, the cross-industry heterogeneity is even larger. See Table A1 in the Appendix.

Table 1

Cross-industry heterogeneity in Capital-Skilled labor Ratio in The USA: 1958-2011

Year	Mean	Std. Dev.	Min	Max	$\frac{\text{Max}}{\text{Min}}$
1958	301.7397	285.4502	24.33098	1638.108	67.326
1968	322.7822	328.1377	23.37773	1813.382	75.569
1978	337.5161	346.5433	20.61539	2167.917	105.160
1988	359.8799	412.1848	14.80645	2862.787	193.347
1998	392.5532	501.0803	37.6413	3629.678	96.428
2008	436.9066	506.8563	54.51049	3431.02	62.612
2011	477.8149	542.2032	81.69276	3628.515	44.417

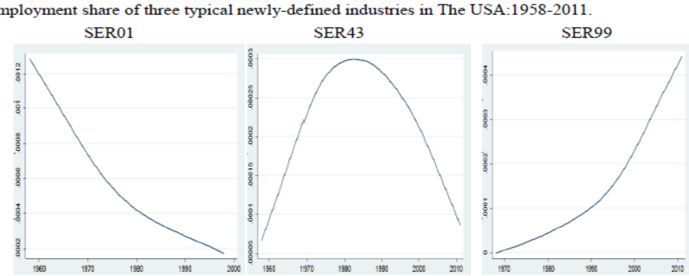
Note: Redefined industries according to the capital-skilled labor ratios. Data source: NBER-CES manufacturing database. (Capital and Labor are measured in 1\$m and 1000s)

Hump-shaped Life Cycle Pattern and Timing Fact

We document a non-monotonic development pattern of an industry. Figure 1 plots the time series of the HP-filtered employment shares of three newly-defined industries in the total manufacturing sector from 1958 to 2011.

Fig 1

Employment share of three typical newly-defined industries in The USA:1958-2011.



Note: HP-filtered employment shares of three typical newly-defined "Industries" (Newly Defined) in The USA from 1958 to 2011. Note: The horizontal axis is the year (from 1958 to 2011), and the vertical axis is the employment share of "Industry 1 (43 and 99)" in all manufacturing industries. An industry with a higher index is more capital intensive relative to skilled labor by construction. 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. Data source: NBER-CES manufacturing database.

Figure 1 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:

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

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 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.

Hump-shaped pattern of industrial dynamics in The USA: 1958-2011

Dependent variable	Employment share*1000		Value-added share *1000	
	GLS	OLS	GLS	OLS
t	0.187***	0.187***	8.426***	8.426***
	(0.041)	(0.042)	(2.242)	(2.266)
t^2	-0.001***	-0.001***	-0.002***	-0.002***
	(0.001)	(0.001)	(0.001)	(0.001)
t*k	5.91e-06 ***	5.91e-06 ***	0.001***	0.001***
	(4.21e-07)	(4.25e-07)	(0.001)	(0.001)
k	-0.018***	-0.018***	-0.696***	-0.696***
	(0.001)	(0.001)	(0.046)	(0.046)
T	0.001	0.001	0.082***	0.082***
	(0.001)	(0.001)	(0.004)	(0.004)
GDPGR	0.001	0.001	-0.050**	-0.050
	(0.001)	(0.001)	(0.053)	(0.054)
Constant	-182.453***	-221.181***	-7996.886***	-7996.82***
	(40.886)	(41.361)	(2221.524)	(2244.834)
Industry dummies	yes	yes	yes	yes
Observation	5032	5032	5032	5032
R-squared		0.150		0.332

Standard errors in parentheses

* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Note:t; k; T and GDPGR represent year, average capital-skilled labor ratio, labor productivity and GDP growth rate, respectively. Data source: NBER-CES manufacturing database.

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.

Industries with Higher Capital Skill Ratios Reach Peaks Later in US Manufacturing: 1958-2011

Dependent variable	Peak time of employment share		Peak time of value-added share	
	(1)	(2)	(3)	(4)
Capital-skilled labor ratio	1.41e-10***		3.89e-07***	
	(1.85e-17)		(4.58e-14)	
Capital expenditure share		4.57e-07***		4.29e-10***
		(6.44e-08)		(6.03e-11)
Constant	4.45e-06***	0.009***	0.009***	4.82e-06***
	(1.38e-14)	(4.17e-09)	(3.41e-11)	(3.90e-12)
R-squared	1.000	0.335	1.000	0.343
Observation	99	99	99	99

Note: Standard errors in parentheses

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Data source: NBER-CES manufacturing database.

Congruence fact

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

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

where Y_{it} is the ratio of newly defined industry i's employment share in total manufacturing 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 whole manufacturing sector at year t, T_{it} is the labor productivity of industry i at year t, D_i is industry dummy. The results are reported in Column (1) in Table 4. It shows that β_1 is negative and significant, indicating that an industry's employment share is smaller if the capital skilled labor ratio of the industry deviates further from the aggregate capital–skilled labor ratio, which is called as the congruence fact. Column (2) shows that this result is robust if the employment share is replaced by the value-added share. (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)

Congruence fact in The USA: 1958-2011

Dependent variable	Employment share *1000	Value-added share*1000
	(1)	(2)
Congruence term	-2.693***	-2.285***
	(0.213)	(0.364)
T	0.014***	0.096***
	(0.005)	(0.011)
Constant	30.142***	29.932***
	(3.091)	(7.364)
Industry dummies	yes	yes
Observation	5032	5032
R-squared	0.124	0.313

Standard errors in parentheses

* p < 0.1, ** p < 0.05, *** p < 0.01

Note: Congruence term is the absolute value of a normalized difference newly-defined sector i's 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_{it} is the labor productivity of industries i of year t. Data source: NBER-CES manufacturing database.

INSERT TABLE 4

To check the robustness of the congruence fact, we use the original NAICS industry classification and run regression (2) again. The results are as follows (with standard errors in parentheses):

$$Y_{it} = \underset{(0.100)}{2.756} - \underset{(0.028)}{0.156} \left| \frac{K_{it}/L_{sit} - K_t/L_{st}}{K_t/L_{st}} \right| + \underset{(0.0002)}{0.003}\, T_{it} + \beta_3 D_i,$$

where Y_{it} is the ratio of an originally-defined industry i's employment share at year t. The coefficient β_1 is again negative and significant at the 99% level, confirming the congruence fact. This result remains significant and robust when controlling for employment share and year dummies.

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 at. In WIOD (SEA) skill type is defined on the basis of the level of educational attainment of the worker. 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 5.

The 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

Table 5. Definitions of Skills

We replicate the same exercise to check all the five facts for other countries. The results show that all of these facts are still valid, suggesting that observed in the US are actually quite general and also true for other countries. For example, Figure xx shows the counterpart for Fact 0 and Table xx establishes the counterpart for Fact 1. The evidence of the counterpart for Facts 2 to 4 is provided in the apppendix.

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(1995-3007), Japan(1995-2009), United States(1995-2009) and Netherlands(1995-2007).
Data source: WIDO Socio-Economic Accounts (SEA) manufacturing database.

Table 2

Cross-industry heterogeneity both in Capital-Skilled labor Ratio cross countries: 1995-2009

Country	Year	Mean	Std. Dev.	Min	Max	$\frac{\text{Max}}{\text{Min}}$
Czech	1995	415.5291	314.6355	113.125	1205.153	10.653
Republic	2007	927.8111	1229.653	280.723	5050.504	17.991
United	1995	38.82903	13.93903	25.27321	71.73659	2.838
Kingdom	2007	50.38292	14.75399	35.66476	82.97176	2.326
Denmark	1995	1168.195	2664.897	233.3862	10399.98	44.561
	2007	1428.889	2372.11	476.3022	9614.506	20.186
Japan	1995	30400.99	77810.08	3963.731	300112.8	75.715
	2009	52951.91	113016.2	6790.862	443203	65.264
Netherlands	1995	121.9292	151.1525	34.63337	614.9808	17.757
	2007	127.3225	127.0834	30.18114	528.3182	15.505
United	1995	55.80374	57.01398	24.24362	245.745	10.136
States	2009	91.68769	78.01277	42.00431	348.3629	8.294

Data source: WIOD Socio-Economic Accounts (SEA). Capital and Skilled-labor are measured in millions (1995=100) and thousands.

[[[[put into the appendix: 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:

$$Y_{itc} = \beta_0 + \beta_1 t + \beta_2 t^2 + \beta_3 k_{itc} \cdot t + \beta_4 T_{itc} + \beta_5 D_{ic} + \beta_6 GDPGR_{tc} + \varepsilon_{itc}. \tag{3}$$

where subscript c represents country. The results are summarized in Table 6.

Table 6

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

Dependent variable	Output share*1000		Employment share *1000	
	(1)	(2)	(3)	(4)
t	21.155***	21.016***	21.155***	21.016***
	(5.211)	(5.204)	(5.211)	(5.204)
t²	-0.005***	-0.005***	-0.005***	-0.005***
	(0.001)	(0.001)	(0.001)	(0.001)
t*k	2.69e-09*	2.54e-09*	2.69e-09*	2.54e-09*
	(2.59e-09)	(2.75e-09)	(2.59e-08)	(2.75e-09)
T	0.002*	0.002*	0.002*	0.002*
	(0.002)	(0.002)	(0.002)	(0.002)
GDPGR	0.069***	0.065***	0.069***	0.069***
	(0.011)	(0.011)	(0.111)	(0.011)
Constant	-19007***	-20926.21***	-21065.42***	-20926.21***
	(6603.297)	(5206.338)	(5213.201)	(5206.338)
Country*Industry dummies	yes	yes	yes	yes
Observation	992	992	992	992
R-squared	0.978	0.978	0.978	0.978

Note: t; k; T and GDPGR represent Year, 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. Countries include Denmark, Czech Republic, United Kingdom, Japan, United States, and Netherlands.
Data source: WIOD Socio-Economic Accounts (SEA) (1995–2009).
* $p < 0.5$, ** $p < 0.01$, *** $p < 0.001$

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 7. Column (1) and column (3) show that the peak time of an industry is positively correlated with its capital-skilled labor ratio.

Table 7

Peak time of industries cross countries: 1995-2009

Dependent variable	Peak time of Output share		Peak time of Employment share	
	(1)	(2)	(3)	(4)
Capital-skilled labor ratio	5.72e-12***		5.75e-12***	
	(1.10e-13)		(1.16e-13)	
Capital-labor ratio		5.15e-11***		5.08e-12***
		(1.26e-12)		(1.09e-13)
Constant	0.056***	0.055***	0.056***	0.055***
	(3.36e-10)	2.13e-09	(3.03e-10)	(2.51e-10)
R-squared	0.986	0.985	0.988	0.982
Observation	84	84	84	84

Countries include Denmark, Czech Republic, United Kingdom, Japan, United States, and Netherlands.
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 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 8.

Table 8
Congruence facts cross countries: 1995-2009

Dependent variable	Value-added share *1000		Employment share*1000	
	(1)	(2)	(3)	(4)
Coherence term 1	-0.425***		-0.029**	
	(0.042)		(0.014)	
Coherence term 2		-0.394***		-0.031**
		(0.042)		(0.014)
T	0.040***	0.039***	0.001**	0.0004**
	(0.001)	(0.001)	(0.0004)	(0.001)
Constant	5.442***	5.107***	8.203***	6.416***
	(0.852)	(0.700)	(0.232)	(0.232)
Country*Industry dummies	yes	yes	yes	yes
Observation	1160	1160	1160	1160
R-squared	0.836	0.835	0.981	0.981

Note: Coherence term 1 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. Coherence term 2 is the absolute value of a normalized difference sector i's skilled labor to employment ratio and the aggregate skilled labor to total employment ratio in manufacturing sectors in country c at year t. T_{it} is the labor productivity of industries i of year T. Countries include Denmark, Czech Republic, United Kingdom, Japan, United States and Netherlands, Data source: WIOD Socio-Economic Accounts (SEA) (1995–2009).
Standard errors in parentheses
* $p < 0.01$, ** $p < 0.05$, *** $p < 0.001$

The findings are consistent with the results in Table 4.

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 but can endogenously generate Facts 0, 2, 3, 4 and 5 simultaneously. We start with the static model, in which all production factors are exogenously given endowment.

3. Static Model

The model setting extends the model in Ju, Lin and Wang (2015) 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 K, skilled labor L_s and unskilled labor L_u. The total labor $(L_s + L_u)$ can be interpreted as the total family size of each household. The model in Ju, Lin and Wang (2015) is a special case of this model when $L_u = 0$. The production function of the final commodity is

$$X = \sum_{n=0}^{\infty} \lambda^n x_n, \tag{4}$$

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 $CRRA$:

$$U(C) = \frac{C^{1-\sigma} - 1}{1 - \sigma}, \text{ where } \sigma \in (0, 1], \tag{5}$$

where C denotes consumption per household.
All technologies exhibit constant returns to scale. Let $F_n(k, l_s, l_u)$ denote the production function for industry $n \geq 0$, where k , 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 $F_0(k, l_s, l_u) = l_u + l_s$. 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 $n \geq 1$, $F_n(k, l_s, l_u)$ is equal to the following value:

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

subject to

$$\begin{aligned} l_{s1} + l_{s2} &\leq l_s, \\ l_{s1} &\geq 0, \\ l_{s2} &\geq 0. \end{aligned}$$

It implies that

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

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. It means that capital not only substitutes unskilled labor and also substitutes "unequipped" skilled labor.

To make the analysis non-trivial, we assume[3]

$$\min\{a - 1, b\} > \lambda > 1. \tag{6}$$

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

$$x_n = F_n(k, l_s, l_u) = \begin{cases} l_u + l_s, & \text{if } n = 0 \\ \frac{l_u}{b^n} + \min\{\frac{k}{a^n}, l_s\}, & \text{if } n \geq 1 \end{cases}. \tag{7}$$

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 (gross interest rate) and wage rates for skilled and unskilled labor, respectively.

[3] Observe that if $\lambda \leq 1$, (4) , (6) and (7) 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.

All the markets are perfectly competitive. As in the standard perfectly competitive general equilibrium model, all firms maximize their profits by taking all prices as given, and each household maximizes her utility function (5) subject to the following budget constraint

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

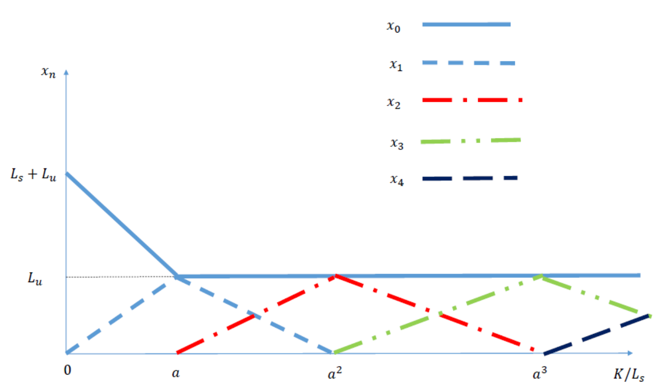
By resorting to the Second Welfare Theorem, we obtain the following proposition.
Proposition 1. *There exists a unique decentralized perfectly competitive market equilibrium, in which industrial output* $\{x_n\}_{n=0}^{\infty}$ *and the final output* X *are characterized in the following table:*

Table 9: 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 $\forall j \neq 0, n, n+1$
$X = L_u + L_s + \frac{\lambda - 1}{a} K$	$X = L_u + \frac{\lambda^n (a - \lambda)}{a - 1} L_s + \frac{\lambda^{n+1} - \lambda^n}{a^{n+1} - a^n} K$

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 $\frac{K}{L_s}$. More precisely, when $\frac{K}{L_s} \in (a^n, a^{n+1})$, 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 $\frac{K}{L_s} \in (0, a)$, only industry 0 and industry 1 coexist, and capital to skilled labor ratios of these two sectors are also closed to the endowment structure. These equilibrium results are consistent with Fact 4 (the congruence fact) documented in Section 2. Table xx is graphically illustrated in the following diagram.



Observe that the output of industry 0 first decreases with $\frac{K}{L_s}$ 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 $\frac{K}{L_s}$ when $\frac{K}{L_s} \in (a^{n-1}, a^n)$ and then decreases with $\frac{K}{L_s}$ when $\frac{K}{L_s} \in (a^n, a^{n+1})$, 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 $\frac{K}{L_s} = a^n$, which means that higher-indexed industries reach their peaks of output at higher levels of $\frac{K}{L_s}$, consistent with Fact 3 (the timing fact).

Table xx shows that the aggregate production function $X(K, L_u, L_s)$ has endogenously different functional forms , depending on which industries are active, which is in turn dictated by the endowment structure, $\frac{K}{L_s}$. In other words, endowment structures determine industrial structures, which endogenously generate the functional form of the aggregate production function. This feature is different from standard macro models, where the functional form of the aggregate production function is exogenous and assumed to be time invarient.

In addition, note that the rental price of capital (gross interest rate) r, skilled-labor wage w_s and unskilled-labor wage w_u are equal to the marginal products of capital, skilled labor and unskilled labor, respectively. They can be directly derived from the endogenous aggregate production function shown in Table XX. For instance, when $a^n L_s \leq K < a^{n+1} L_s$ for some $n \geq 1$, the factor prices are, respectively, given by

$$\begin{aligned} w_u &= \frac{\partial X}{\partial L_u} = 1, \\ w_s &= \frac{\partial X}{\partial L_s} = \frac{\lambda^n (a-\lambda)}{a-1}, \\ r &= \frac{\partial X}{\partial K} = \frac{\lambda^{n+1} - \lambda^n}{a^{n+1} - a^n}. \end{aligned}$$

Let p_n denotes the market price for good n. We have $p_n = \lambda^n$ for any $n \geq 0$, and the skill premium is given by

$$\frac{w_s}{w_u} = \begin{cases} 1, & \text{when } 0 \leq K < aL_s \\ \frac{\lambda^n(a-\lambda)}{a-1}, & \text{when } a^n L_s \leq K < a^{n+1} L_s \text{ for } n \geq 1 \end{cases},$$

which shows that $\frac{w_s}{w_u}$ is positively correlated with $\frac{K}{L_s}$ because both are weakly increasing in n. It is consistent with Fact 0 documented earlier. To summarize, we obtain the following table, where θ denotes the labor income share in total GDP:

Table xxx: Static Equilibrium

$0 \leq K < aL_s$	$a^n L_s \leq K < a^{n+1} L_s$ for $n \geq 1$
$X = L_u + L_s + \frac{\lambda-1}{a} K$	$X = L_u + \frac{\lambda^n(a-\lambda)}{a-1} L_s + \frac{\lambda^{n+1}-\lambda^n}{a^{n+1}-a^n} K$
$\frac{w_s}{w_u} = 1$	$\frac{w_s}{w_u} = \frac{\lambda^n(a-\lambda)}{a-1}$
$\frac{r}{w_s} = \frac{\lambda-1}{a}$	$\frac{r}{w_s} = \frac{\lambda-1}{a^n(a-\lambda)}$
$\theta = \frac{L_u+L_s}{L_u+L_s+\frac{\lambda-1}{a}K}$	$\theta = \frac{L_u+\frac{\lambda^n(a-\lambda)}{a-1}L_s}{L_u+\frac{\lambda^n(a-\lambda)}{a-1}L_s+\frac{\lambda^{n+1}-\lambda^n}{a^{n+1}-a^n}K}$

$$E(C) = \frac{a^{n+1} - a^n}{\lambda^{n+1} - \lambda^n} C - \left[\frac{a^{n+1} - a^n}{\lambda^{n+1} - \lambda^n} L_u + \frac{a^n(a-\lambda)}{\lambda - 1} L_s \right]$$

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 examine the case with endogenous changes in physical capital while shutting down human capital investment, or alternatively speaking, we will take the amount of skilled labor and unskilled as exogenous and time-invariant. In Part B, we further allow endogenous human capital investment, which transforms unskilled labor into skilled labor, and we explore the interaction between endogenous physical and human capital accumulation in determining industrial upgrading and skill premium. In the appendix, we also generalize the model in Part B by allowing for the possibility of the existence of positive externalities in human capital investment.

Part A: No Human Capital Investment

In this part, we let K grow endogenously but keep L_s and L_u fixed over time. Time is continuous and households are infinitely lived. Following Ju, Lin and Wang (2015), we assume that 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 also all the intermediate goods that are required to produce the final commodity. The final commodity is for consumption and is not storable. This sector is characterized in the previous section. To produce consumption good $C(t)$ at time point t, it requires $E(C(t))$ physical capital, where function $E(\cdot)$ is derived from the last row in Table xx. When $C(t)$ falls on different intervals, the underlying industrial composition is different. More specifically,

$$E(C) = \begin{cases} E_{0,1}(C), & \text{if } L_u + L_s \leq C < \lambda L_s + L_u \\ E_{0,n,n+1}(C), & \text{if } \lambda^n L_s + L_u \leq C < \lambda^{n+1} L_s + L_u \text{ for } n \geq 1 \end{cases}, \tag{8}$$

where

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

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

$$\dot{K} = \xi K - \delta \cdot E(C(t)), \tag{9}$$

where ρ is the time discount rate, δ is the depreciation rate, $E(C(t))$ is given by (8), and $K(0) = K_0$ is given.

(9) 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 denote the time point when output of good n reaches the highest level $x_n = \lambda^n L_s$ for $n \geq 1$. Let t_0 denote the last 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

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

subject to

$$\dot{K} = \begin{cases} \xi K & \text{when} \quad 0 \leq t \leq t_0 \\ \xi K - \delta E_{0,1}(C), & \text{when} \quad t_0 \leq t \leq t_1 \\ \xi K - \delta E_{0,n,n+1}(C), & \text{when} \quad t_n \leq t \leq t_{n+1}, \text{ for } n \geq 1 \end{cases},$$

K_0 is given.

When K_0 is sufficiently small, no intermediate good other than good 0 will be produced in the consumption sector for a while until t_0, after which capital is used for producing consumption good. Therefore,

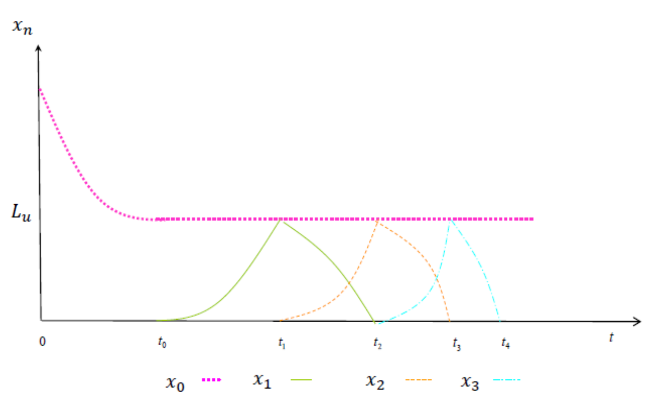
$$\begin{aligned} C(t) &= L_u + L_s, \forall t \in [0, t_0]; \\ C(t_n) &= L_u + \lambda^n L_s, \text{for } n \geq 1. \end{aligned}$$

To characterize life cycle dynamics of industries, we define $m_n \equiv t_{n+1} - t_n$, $\forall n \geq 0$.
Proposition 2. *The country's consumption remains constant at level $L_u + L_s$ and then grows at the constant rate $g_C = \frac{\xi - \rho}{\sigma}$ after time t_0. For each industry $n \geq 1$, its output exhibits a hump-shaped life-cycle dynamic pattern: it appears at time t_{n-1}, its output rises for a period of m_{n-1}, and reaches the peak at time t_n, after which its output declines for a period of m_n and disappears after t_{n+1}, so its whole life span is $m_{n-1} + m_n$, where*

$$m_n = \frac{\log \frac{L_u + \lambda^{n+1} L_s}{L_u + \lambda^n L_s}}{g_C}, \forall n \geq 0. \tag{10}$$

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.



It shows that the model predictions are 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 ratio reach the peak 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).

Observe that $\frac{\partial m_n}{\partial n} < 0$, which implies that the life span of an industry decreases with the capital-skill ratio of the industry. Moreover, for each industry, the booming period is longer than the decline period ($m_{n-1} > m_n$), consistent with the fact of "shakeout" documented in the literature (see, for example, Jovanovic and MacDonald()). As $t \to \infty$, $m_n \to \frac{\log \lambda}{g_C}$. By comparison, the booming and decline periods of any industry are equal ($m_{n-1} = m_n = \frac{\log \lambda}{g_C}$) in Ju, Lin and Wang (2015), so the life span of every industry is equal to $\frac{2 \log \lambda}{g_C}$, which is case when substituting $L_u = 0$ into (10).

Combining the static results in Table xxx and the dynamic results summarized in Proposition 2, we obtain the following proposition about how skill premium interacts with industrial upgrading over time.

Proposition 3. *The skill premium $\frac{w_s}{w_u}$ is equal to one before time t_0. For each industry $n \geq 1$, the skill premium $\frac{w_s}{w_u}$ is equal to $\frac{\lambda^n (a-\lambda)}{a-1}$ when $t \in [t_n, t_{n+1})$, during which period industry n declines whereas industry $n+1$ booms. Alternatively speaking, during the whole life span $(m_{n-1} + m_n)$ of industry n, skill premium is $\frac{\lambda^{n-1}(a-\lambda)}{a-1}$ when industry n is booming and the skill premium jumps to $\frac{\lambda^n (a-\lambda)}{a-1}$ when industry n declines till it disappears, where m_n* is given by (10) for all $n \geq 0$.

This proposition shows that skill premium adjusts at the same frequency as the industrial structure. More specifically, skill premium will change more and more frequently as the capital-skill ratios of industries increase over time (consistent with Figure 2), and eventually converges 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. We explore the situation when both physical capital investment and human capital investment are endogenously decided by private agents. Again, based on the Second Welfare Theorem, we can 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

$$\begin{aligned} \dot{K} &= \xi K - \delta \cdot E(C(t)) - G(t), & (11) \\ \dot{L}_s &= \phi L_u^{1-\kappa} \cdot G(t)^\kappa, & (12) \\ \dot{L}_u &= g_b \cdot (L_u + L_s) - \dot{L}_s & (13) \end{aligned}$$

and that $K(0), L_u(0), L_s(0)$ are all given, where $G(t)$ is human capital investment (per household) at time t and g_b is the exogenous birth rate. (11) states that human capital investment is costly and in terms of physical capital. (12) 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 transformative process, capturing all relevant institutional features that affect the rate of skill transformation such as the quality of the training program. (13) states that the family size $(L_u + L_s)$ grows at an exogenous rate g_b.[4] Moreover, we assume

$$g_b < \xi < \frac{\xi - \rho}{\sigma}, \tag{14}$$

which will be explained soon. Another interpretation of (13) is 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 $\widehat{t}_n$ denote the time point when output of good n reaches the highest level $x_n = \lambda^n L_s(\widehat{t}_n)$ for $n \geq 1$. Let $\widehat{t}_0$ denote the last time point when only good 0 is produced.

Proposition 4. *When both physical capital investment and human capital investment are endogenous, the consumption growth rate remains unchanged:*

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

and *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*

$$\frac{\dot{L}_s}{L_s} = \frac{\dot{L}_u}{L_u} = \frac{\dot{G}}{G} = g_b \tag{15}$$

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

$$\frac{L_s}{L_u} = \frac{\phi \chi_n^\kappa}{g_b}, \tag{16}$$

$$\frac{G(t)}{L_u(t)} = \chi_n, \tag{17}$$

where χ_n is uniquely determined by

$$(1-\kappa)\,\phi\chi_n^\kappa = \frac{a^n}{\lambda - 1}\left\{(a-\lambda) - \frac{a-1}{\lambda^n}\right\}\phi\kappa\delta{\chi_n}^{\kappa-1} - \xi, \tag{18}$$

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).

[4] 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 $\frac{C(t)}{L_u(t)+L_s(t)}$.

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 (14). 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 growth rate g_C of the consumption goods. (15) 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, that is, during the time interval $[\widehat{t}_n, \widehat{t}_{n+1})$. Moreover, (16) and (17) 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: This temporary BGP with the coexistence of industries $0, n$ and $n+1$ is asymptotically reached when t gets sufficiently close to $\widehat{t}_{n+1}$, and then at time point $\widehat{t}_{n+1}$, industry n exactly disappears and industry $n+2$ is about to enter, 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. The following is true on the temporary BGP when industries $0, n$ and $n+1$coexist for any $n \geq 1$:

$$\frac{\partial \chi_n}{\partial n} > 0; \frac{\partial \chi_n}{\partial \delta} > 0; \frac{\partial \chi_n}{\partial \phi} > 0; \frac{\partial \chi_n}{\partial a} > 0; \frac{\partial \chi_n}{\partial \xi} < 0. \tag{19}$$

Proof. Immediately implied by (18). **Q.E.D**

(19) 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, (16) and (18) 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).

(19) 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 ladder of neighboring industries a becomes larger. The reason is that a higher δ weakens the desirability of using physical capital to produce consumption goods, or alternatively speaking, skilled labor and hence human capital investment will be more favored, so $\frac{G}{L_u}$ increases and $\frac{L_s}{L_u}$ increases. Similarly, an increase in the efficiency of skill transformation 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 rely more on 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 the physical capital production is more efficient and hence physical capital becomes increasingly cheaper relative to skilled

labor, and to produce more consumption goods does not have to increase skilled labor, so the incentive to invest human capital is weakened as the opportunity cost rises, as a result, both $\frac{G(t)}{L_u(t)}$ and $\frac{L_s}{L_u}$ decreases 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 (18) 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 (17) 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, (16) implies that skilled to unskilled labor ratio becomes a^{κ} times higher on the new BGP than the previous BGP.
Define $\widehat{m}_n \equiv \widehat{t}_{n+1} - \widehat{t}_n$, $\forall n \geq 0$.
Proposition 5. *For each industry $n \geq 1$, its output exhibits a hump-shaped life-cycle dynamic pattern: it appears at time $\widehat{t}_{n-1}$, its output rises for a period of $\widehat{m}_{n-1}$, and reaches the peak at time $\widehat{t}_n$, after which its output declines for a period of $\widehat{m}_n$ and disappears after $\widehat{t}_{n+1}$. During industry n's whole life span $\widehat{m}_{n-1} + \widehat{m}_n$, skill premium is $\frac{\lambda^{n-1}(a-\lambda)}{a-1}$ when industry n is booming (i.e.,* when $t \in [\widehat{t}_{n-1}, \widehat{t}_n)$) *and the skill premium jumps to $\frac{\lambda^{n}(a-\lambda)}{a-1}$ when industry n* declines (*i.e., when $t \in [\widehat{t}_n, \widehat{t}_{n+1})$), where*

$$\widehat{m}_n \approx \frac{\log \frac{g_b+\phi\lambda^{n+1}\chi_n^{\kappa}}{g_b+\phi\lambda^{n}\chi_{n-1}^{\kappa}}}{\frac{\xi-\rho}{\sigma}-g_b}, \forall \text{ sufficiently large } n > 1, \tag{20}$$

where χ_n is determined by (18). Moreover, $\widehat{m}_n > m_n$ for all $n > 1$, where m_n is given by (10).
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 (20). The third semi-equation comes from (15) and (16). 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. For example, during the transition period between two different 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 not exactly equal to, but rather 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 same in Dynamic Model A and Dynamic Model B. Second, both L_u and L_s remain constant in Part A but grow at some 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 (20). 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 $\lim_{n\to\infty} m_n = \frac{\log\lambda}{g_C - g_b}$, which means that life spans of industries, hence the frequencies of skill premium adjustment, eventually will be identical.

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. In particular, our model explains explicitly how skill premium dynamics at the aggregate level is logically connected to the life-cycle dynamics of underlying industries at the disaggregated level, which may serve as a micro-foundation for the skill-biased technological progress typically assumed as exogenous in the existing macro development literature. The model also implies that the optimal human capital investment is stage-dependent and varies with the underlying industrial structures, which are in turn affected by physical capital accumulation. We also highlight that the driving force for the structural change at the disaggregated industry level (industrial dynamics) in our model is the endogenous change in factor endowment structures (capital-skill ratios), different from the mechanisms highlighted in the literature of structural transformation such as unbalanced productivity growth, non-homothetic preferences or international trade.

In this model, there is no role of government as the first welfare theorem applies. However, if we deviate from the first-best environment by introducing relevant frictions such as financial frictions, labor market frictions, or externalities, we would expect that the endogenous skill premium dynamics will be presumably different and there would be scope for discussing potentially welfare-enhancing roles of government policies. 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).

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

2. Aghion, Philippe. 2002. "Schumpeterian growth theory and the dynamics of income inequality." *Econometrica* 70(3): 855-882.

3. 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.

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

5. Buera, Francesco, Joseph Kaboski and Richard Rogerson, 2018. "Skill-biased Structural Change", unpublished working paper

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

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

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

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

12. Klepper, Steve. 1996. "Entry, exit, growth, and innovation over the product life cycle", *American Economic Review*.

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

14. 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

15. 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.

16. Kongsamut, Piyabha, Rebelo, Sergio and Danyang Xie. 2001. "Beyond Balanced Growth." Review of Economic Studies 68: 869-882

17. 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

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

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

20. 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

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

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

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

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

25. Uy, Timothy, Kei-Mu Yi and Jing Zhang. 2013. "Structural Transformation in an Open Economy." Journal of Monetary Economics

26. 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 4.

Proof. Substituting (12) into (13) 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:

$$\frac{\partial H}{\partial C} = 0 \Rightarrow C(t)^{-\sigma} = \widehat{\lambda}\delta \frac{a^{n+1}-a^n}{\lambda^{n+1}-\lambda^n} \Rightarrow -\sigma \frac{\dot{C}}{C} = \frac{\dot{\widehat{\lambda}}}{\widehat{\lambda}} \tag{21}$$

$$\frac{\partial H}{\partial G} = 0 \Rightarrow [\psi - \eta]\, \kappa \phi L_u^{1-\kappa} \cdot G(t)^{\kappa - 1} = \widehat{\lambda} \tag{22}$$

$$\dot{\widehat{\lambda}} = \rho \widehat{\lambda} - \frac{\partial H}{\partial K} \Rightarrow \frac{\dot{\widehat{\lambda}}}{\widehat{\lambda}} = \rho - \xi \tag{23}$$

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

$$\Rightarrow \frac{\dot{\eta}}{\eta} = \rho - g_b - \delta \frac{\widehat{\lambda}}{\eta} \frac{a^{n+1}-a^n}{\lambda^{n+1}-\lambda^n} + \left(1 - \frac{\psi}{\eta}\right) \phi\,(1-\kappa)\, L_u^{-\kappa} \cdot G(t)^\kappa \tag{24}$$

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

$$\frac{\dot{\psi}}{\psi} = \rho - \left[\delta \frac{\widehat{\lambda}}{\psi} \frac{a^n(a-\lambda)}{\lambda - 1} + \frac{\eta}{\psi} g_b\right] \tag{25}$$

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

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

which implies

$$L_u^{-\kappa} \cdot G(t)^{\kappa} = -\frac{\xi - g_b - \delta \frac{\widehat{\lambda}}{\eta} \frac{a^{n+1}-a^n}{\lambda^{n+1}-\lambda^n}}{\left(1 - \frac{\psi}{\eta}\right) \phi \left(1 - \kappa\right)}. \tag{26}$$

By (12), $\frac{\dot{L}_s}{L_s}$ is constant if and only if $\frac{\dot{L}_s}{L_s} = \frac{\dot{L}_u}{L_u} = \frac{\dot{G}}{G}$. Let $g_G \equiv \frac{\dot{G}}{G}$. (13) 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

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

(22) can be rewritten as

$$\frac{\widehat{\lambda}}{\eta} = -\phi L_u^{1-\kappa} \cdot \kappa G(t)^{\kappa-1} \left(1 - \frac{\psi}{\eta}\right), \tag{27}$$

which is used to substitute out $\frac{\widehat{\lambda}}{\eta}$ in equation (26), we obtain

$$\left(1 - \frac{\psi}{\eta}\right) \left(1 - \kappa\right) \phi \left[\frac{G(t)}{L_u(t)}\right]^{\kappa} = g_b - \xi - \left(1 - \frac{\psi}{\eta}\right) \phi \kappa \delta \frac{a^{n+1} - a^n}{\lambda^{n+1} - \lambda^n} \left[\frac{G(t)}{L_u(t)}\right]^{\kappa-1}. \tag{28}$$

Since $\frac{\dot{\eta}}{\eta} = \frac{\dot{\widehat{\lambda}}}{\widehat{\lambda}} = \frac{\dot{\psi}}{\psi}$, (25) and (23) jointly imply

$$\delta \frac{\widehat{\lambda}}{\eta} \frac{a^n (a - \lambda)}{\lambda - 1} + g_b = \xi \frac{\psi}{\eta} \tag{29}$$

so we have three unknowns $\frac{\widehat{\lambda}}{\eta}, \frac{\psi}{\eta}, \frac{G(t)}{L_u(t)}$ and three equations (27)-(29).

Using the brutal force, we obtain

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

which uniquely determines $\frac{G(t)}{L_u(t)}$. Denote this solution by χ_n. By (12), we have

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

and

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

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

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

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

which implies

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

Q.E.D.

[[[[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 (13) implies $\frac{\dot{L}_s}{L_s}=\frac{\dot{L}_u}{L_u}=g_b$. And (12) 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 $\frac{L_u(t)}{L_s(t)}\to0$, $L_u(t)\to0$ 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). (22) and (24) 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 iff

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

is a constant.]]]]]

Appendix

Cross-industry heterogeneity in Capital-Skilled labor Ratio in the USA:1958-2011

Year	Mean	Std. Dev.	Min	Max	Max/Min
1958	180.9721	195.6405	2.733333	1638.108	599.308
1968	227.5144	227.2842	8.695267	1931.087	222.085
1978	292.623	300.3177	14.14045	2425.79	171.550
1988	369.1544	426.2395	14.80645	3845	259.684
1998	461.3952	559.7446	37.6413	5657.706	150.306
2008	685.0558	807.5926	54.51049	7139	130.970
2011	805.0773	903.4446	81.69277	7647.8	93.617

Note: The original 469 industries within the manufacturing sector.

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

[Insert Table 1A here]

A different formulation of externality

Everything is identical to the previous case excep that (12) is replaced by

$$\dot{L}_s = \phi L_s^{1-\nu}\left[L_u^{1-\kappa}\cdot G(t)^{\kappa}\right]^{\nu}\overline{G}^{\zeta}(t), \tag{30}$$

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

$$\begin{aligned}
\dot{K} &= \xi K - \delta\cdot E(C(t)) - G(t), && (31)\\
\dot{L}_s &= \phi L_s^{1-\nu}\left[L_u^{1-\kappa}\cdot G(t)^{\kappa}\right]^{\nu}\overline{G}^{\zeta}(t), && (32)\\
\dot{L}_u &= g_b\cdot(L_u+L_s) - \dot{L}_s && (33)
\end{aligned}$$

Decentralized equilibrium:

$$\dot{L}_u = g_b\cdot(L_u+L_s) - \phi L_s^{1-\nu}\left[L_u^{1-\kappa}\cdot G(t)^{\kappa}\right]^{\nu}\overline{G}^{\zeta}(t). \tag{34}$$

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:

$$\begin{aligned}
\frac{\partial H}{\partial C} &= 0 \Rightarrow C(t)^{-\sigma} = \lambda\delta\frac{\partial E(C(t))}{\partial C(t)} \Rightarrow -\sigma\frac{\dot{C}}{C} = \frac{\dot{\lambda}}{\lambda} && (35)\\
\frac{\partial H}{\partial G} &= 0 \Rightarrow \eta\phi\kappa\nu L_s^{1-\nu}\left[L_u^{1-\kappa}\cdot G(t)^{\kappa}\right]^{\nu}\overline{G}^{\zeta}(t)\cdot G(t)^{-1} = \lambda && (36)\\
\dot{\lambda} &= \rho\lambda - \frac{\partial H}{\partial K} = \rho\lambda - \lambda\xi \Rightarrow \frac{\dot{\lambda}}{\lambda} = \rho - \xi\\
\dot{\eta} &= \rho\eta - \frac{\partial H}{\partial L_u} \Rightarrow \frac{\dot{\eta}}{\eta} = \rho - \left[g_b - (1-\kappa)\,\phi L_u^{-\kappa}\cdot G(t)^{\kappa}\overline{G}^{\zeta}(t)\right] && (37)
\end{aligned}$$

Part C. Human Capital Externality

Now we introduce externality in human capital investment. Everything is identical to Part B except that (12) is replaced by the following:

$$\dot{L}_s = \phi L_u^{1-\kappa}\cdot G(t)^{\kappa}\overline{G}^{\zeta}(t), \tag{38}$$

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*

$$\begin{aligned}
g_C &= \frac{\xi-\rho}{\sigma}, \forall t \geq \widehat{t}_0\\
\frac{\dot{L}_s}{L_s} &> g_b > \frac{\dot{L}_u}{L_u}, \forall t && (39)\\
\frac{L_s}{L_u} &= \frac{\phi\chi_n^{\kappa}}{g_b}, \forall t \in (\widehat{t}_n, \widehat{t}_{n+1}] && (40)\\
\frac{G(t)}{L_u(t)} &= \chi_n, \forall t \in (\widehat{t}_n, \widehat{t}_{n+1}) && (41)
\end{aligned}$$

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 (38) into (13) yields

$$\dot{L}_u = g_b\cdot(L_u+L_s) - \phi L_u^{1-\kappa}\cdot G(t)^{\kappa}\overline{G}^{\zeta}(t). \tag{42}$$

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:

$$
\begin{aligned}
\frac{\partial H}{\partial C} &= 0 \Rightarrow C(t)^{-\sigma} = \lambda\delta \frac{\partial E(C(t))}{\partial C(t)} \Rightarrow -\sigma \frac{\dot{C}}{C} = \frac{\dot{\lambda}}{\lambda} \\
\frac{\partial H}{\partial G} &= 0 \Rightarrow (\psi - \eta)\, \phi L_u^{1-\kappa} \cdot \kappa G(t)^{\kappa-1} \overline{G}^{\zeta}(t) = \lambda \qquad (43) \\
\dot{\lambda} &= \rho\lambda - \frac{\partial H}{\partial K} = \rho\lambda - \lambda\xi \Rightarrow \frac{\dot{\lambda}}{\lambda} = \rho - \xi \\
\frac{\dot{\psi}}{\psi} &= \rho - \left[\delta \frac{\widehat{\lambda}}{\psi} \frac{a^n(a-\lambda)}{\lambda - 1} + \frac{\eta}{\psi} g_b\right] \qquad (44) \\
\dot{\eta} &= \rho\eta - \frac{\partial H}{\partial L_u} \Rightarrow \frac{\dot{\eta}}{\eta} = \rho - g_b - \delta \frac{\widehat{\lambda}}{\eta} \frac{a^{n+1} - a^n}{\lambda^{n+1} - \lambda^n} + \left(1 - \frac{\psi}{\eta}\right) \phi\, (1-\kappa)\, L_u^{-\kappa} \cdot G(t)^{\kappa} \overline{G}^{\zeta}(t) \qquad (45)
\end{aligned}
$$

Thus we still have

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

Moreover, in equilibrium, we must have $\overline{G}(t) = G(t)$, so (45) can be rewritten as

$$
\frac{\dot{\eta}}{\eta} = \rho - \left[g_b - \left(1 - \frac{\psi}{\eta}\right)(1-\kappa)\, \phi L_u^{-\kappa} \cdot G(t)^{\kappa+\zeta}\right] - \delta \frac{\widehat{\lambda}}{\eta} \frac{a^{n+1} - a^n}{\lambda^{n+1} - \lambda^n}, \qquad (46)
$$

which is a constant if

$$
\frac{\dot{\eta}}{\eta} = \frac{\dot{\psi}}{\psi} = \frac{\dot{\widehat{\lambda}}}{\widehat{\lambda}} = \rho - \xi \qquad (47)
$$

and

$$
\frac{\dot{L}_u}{L_u} = \frac{(\kappa + \zeta)}{\kappa} \frac{\dot{G}}{G}. \qquad (48)
$$

Substituting $\overline{G}(t) = G(t)$ into (43) yields

$$
\frac{\widehat{\lambda}}{\eta} = \left(\frac{\psi}{\eta} - 1\right) \phi L_u^{1-\kappa} \cdot \kappa G(t)^{\kappa - 1 + \zeta}, \qquad (49)
$$

Substitute the above into (46) and use (47), we obtain

$$
\xi - g_b - \left(\frac{\psi}{\eta} - 1\right) \phi \left[(1-\kappa)\, L_u^{-\kappa} \cdot G(t)^{\kappa+\zeta} + \delta \frac{a^{n+1} - a^n}{\lambda^{n+1} - \lambda^n} L_u^{1-\kappa} \cdot \kappa G(t)^{\kappa-1+\zeta}\right] = 0 \qquad (50)
$$

By revoking (47) and (44), we obtain

$$\delta\frac{\widehat{\lambda}}{\eta}\frac{a^n(a-\lambda)}{\lambda-1}+g_b=\frac{\psi}{\eta}\xi. \tag{51}$$

(49) and (51) 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 (50), we obtain

$$\left[\frac{a^n(a-\lambda)}{\lambda-1}-\frac{a^{n+1}-a^n}{\lambda^{n+1}-\lambda^n}\right]\delta\phi L_u^{1-\kappa}\cdot\kappa G(t)^{\kappa-1+\zeta}=\phi\left(1-\kappa\right)L_u^{-\kappa}\cdot G(t)^{\kappa+\zeta}+\xi \tag{52}$$

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

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

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

which implies

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

or equivalently,

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

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

Substituting $\overline{G}(t)=G(t)$ into (42) 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 (48). (38) implies

$$\frac{\dot{L_s}}{L_s}=\frac{L_u}{L_s}\phi L_u^{-\kappa}\cdot G(t)^{\kappa+\zeta} \tag{53}$$

Let $g_G\equiv\frac{\dot{G}}{G}$. (42) 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 (53), 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

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

(52) implies

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

(43), 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).

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

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

$$\begin{aligned}
(1-\kappa)\,\frac{\dot{L_u}}{L_u} + (\kappa - 1 + \zeta)\, g_G &= g_{\Xi}(t) \\
g_G - \frac{\dot{L_u}}{L_u} &= \frac{g_{\Xi}(t) - \zeta g_G}{(\kappa - 1)}
\end{aligned} \tag{55}$$

substituting it into (54) 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 (55).

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:
(43) 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.