Skill Mismatch, Structural Unemployment and Industry Dynamics

Yong Wang and Siyu Chen*

April 2021

Abstract

This paper studies industy dynamics, aggregate growth and dynamics of the aggregate unemployment rate in a general equilibrium model with infinite industries that are heterogeneous in capital intensities. We analytically characterize these dynamics in equilibrium. We show that the aggregate unemployment rate exhibits a cyclical moverment because high-skilled workers in an old industry become low-skilled in the new, more capital-intensive industry and suffers from skill mismatch and low job finding rates, but they may become high-skilled through on-job learning in the new industry and their job finding rate increases. Capital accumulation drives the incessant structural change process of capital-intensive industries replacing less capital-intensive ones, resulting in endless cross-industry labor reallocation and the repeated cyclicality of the aggregate unemployment rate despite the absence of aggregate shocks. When skill mismatch becomes more severe, the unemployment rate increases and the life span of an industry becomes longer.

Keywords: Structural Change, Industry Dynamics, Skill Mismatch, Structural Unemployment

JEL Classification: E24, E37, O14, O41

*Yong Wang: Peking University. Email: yongwang@nsd.pku.edu.cn. Siyu Chen: Peking University. Email: csynsd@pku.edu.cn.

1 Introduction

The underlying composition of industries, or industrial structure, changes over time as the whole economy grows. Huge progress has been made in our understanding about the process of structural transformation at aggregated sector levels such as the first, second and tertiary sectors, which exhibits Kuznets facts, namely, the employment share (or value added share) of the agriculture sector declines, that of manufacturing sector follows a hump shape, and that of the service sector increases, as GDP per capita increases (Kuznets, 1980; Rogerson et al., 2010). However, resource reallocation also takes place across industries at a much more disaggregated sectoral level such as subindustries within the manufacturing sector. It is observed that labor-intensive industries are gradually replaced by more capital-intensive industries as capital accumulates over time (Chenery, 1969; Ju, Lin and Wang, 2015; Acemoglu and Guerierri, 2008). Old jobs in the sunset industries and novel jobs in the sunrise industries typically require different sets of skills, so workers from sunset industries may lack the skills required by jobs in sunrise industries. The mismatch between workers' current skill and the skill requirement in the newly available jobs is referred to as skill mismatch throughout this paper. Skill mismatch never disappears, as new sunrise industries keep emerging and replacing sunset industries over and over again.

There is ample empirical evidence showing that mismatch is asymmetric across sectors and that it is an important source of unemployment. For example, Şahin *et al.* (2014) find that that mismatch across industries and occupations accounts for 0.6 to 1.7 percentage points of rise in the U.S. unemployment rate. They estimate the industry-specific matching efficiency at the 2-digit industry level and find that traditional industries like mining, retail and construction have much larger matching efficiency than relatively modern industries like information and finance. Their cross-industry results show that skill mismatch is more severe in emerging industries than in traditional industries. This finding is further confirmed by Herz and Rens (2014), who propose an accounting framework to estimate the overall magnitude of mismatch-induced unemployment and the quantitive contribution of each labor market friction to the unemployment. They find that there also exists a large dispersion in job finding rates across segments. They conclude that mismatch is responsible for 84% of unemployment rate. Like Şahin *et al* (2014), they find that industry-level mismatch is most serious, which explains 14% of the increase in total unemployment rate during the Great Recession, and that mismatch would contribute to 29% of the increase if disaggregated further to the three-digit occupational level.

In light of these theoretical consideration and empirical findings, the primary goal of this paper is to theoretically investigate how skill mismatch endogenously emerges in the process of structural change and how it would impact structural unemployment, industry dynamics and aggregate growth. To this end, we develop a tractable dynamic general equilibrium model, in which there are multiple industries heterogeneous in capital intensities and the key driving force for the life-cycle industry dynamics and structural change is endogenous capital accumulation. These features echo those in Ju, Lin and Wang (2015). However, different from their assumptions of homogeneous labor and frictionless labor markets, we introduce skill heterogeneity across industries and workers as well as skill mismatch into the model. We model this part by following the similar approach as in Restrepo (2015), who study skill mismatch across symmetric industries. Our paper differs mainly in that industries are asymmetric in capital intensities and that the driving force of non-monotonic industrial dynamics and continual shifts in industrial structures is the endogenous accumulation in physical capital rather than sectoral productivity changes. In our model, workers newly relocated from old (less capital-intensive) industries are assumed to be initially low-skilled in the new (more capital-intensive) industry and facing a low job finding rate due to skill mismatch. However, these workers, once employed, could become high-skilled workers at a Poisson rate through on-job learning and therefore enjoy a relatively high job finding rate in the new industry when they exogenously lose jobs. As capital accumulates, the current industry also becomes a sunset industry later on and an even more capital-intensive new industry emerges and expands, so workers have to leave again for the new industry and become low-skilled workers there, so on and so forth.

A major finding is that the aggregate unemployment rate in equilibrium exhibits a cyclical pattern along with the perpetual changes in underlying industrial composition. The intuition is the following. When all workers are in the same industry, the number of employment and that of unemployment both reach the peak. Then low-skilled workers in the old industry start to move to the new industry, earlier than high-skilled ones, because their opportunity cost of switching industries is lower, and the aggregate unemployment rate starts to decline because low-skilled workers in both industries gradually become high-skilled due to on-job learning and therefore face higher job finding rates. The unemployment rate reaches the bottom when all unskilled workers from the old industry have moved to the sunrise industry. The unemployment rate goes up again as the remaining (high-skilled) workers in the sunset industries start to move to the sunrise industries, because they become low-skilled workers in the new industry and encounter skill mismatch. This process continues till all high-skilled workers in the old industry have moved to the new industry, at which point the old industry completely exits and all workers are in the new industry. The unemployment rate reaches the maximum value and thus completes one cycle. As capital accumulates further, an even more capital-intensive industry emerges as a new sunrise industry and it starts to expand, attracting (unskilled) workers from the old sunset industry, so the equilibrium unemployment rate declines again, ad infinitum. As long as new industries keep emerging sequentially and replaces old ones, the aggregate unemployment rate keeps repeating this cyclical pattern. Note that the cyclicality of the unemployment rate is not caused by any exogenous aggregate shocks, but rather due to the endogenously fluctuating structural unemployment created as new industries replace old ones. Consequently, the fluctuation in unemployment rates is not at the business cycle frequency, but rather at a lower frequency consonant with the speed of more capital-intensive industries replacing the old ones.

Three parameters play important roles in the model. One is the rate of investment-specific technological progress, denoted by A, which governs the speed of capital accumulation. The second is the Poisson rate of skill upgrading due to on-job learning, denoted by ξ. The third parameter is π, which measures the degree of skill mismatch as reflected in the job finding rate. The model is tractable and analytical solutions are obtained to characterize the life-cycle dynamics of each of the infinite industries and the associated labor market performance along the aggregate growth, we are able to conduct clean comparative static analyses with respect to the above three parameters. We show that, a higher A (faster investment-specific technological progress) translates into a faster aggregate consumption growth, a shorter life span of each industry, a faster process of sunset industries being replaced by sunrise industries, and a universally higher unemployment rate. It is mainly because industry dynamics is driven faster as capital accumulates faster, so workers must move to new industries more frequently and hence mismatch occurs more frequently. Morever, a higher ξ (faster skill learning) implies a universally lower unemployment rate and a shorter life span of each industry, because on-job learning dampens the negative impact of skill mismatch and provides more incentives for workers to move to new industries. Not surprisingly, a lower π (more severe skill mismatch) implies a universally higher unemployment rate and a longer industry life span because workers become more reluctant to move to new industries.

Our paper is closely related to two strands of literature. Firstly, it contributes to the literature on structural transformation and industry dynamics. In contrast to widely discussed mechanisms such as income effect due to non-homothetic preferences (Kongsmut, Rebelo and Xie, 2001; Buera and Kaboski, 2012; Boppart, 2014; Comin, Lashkari and Mestieri, 2021), or substitution effect due to unbalanced productivity growth across sectors (Ngai and Pissaridies, 2007; Acemoglu and Guerierri, 2008), the driving force of structural transformation (and related industry dynamics) in our paper is capital accumulation, which changes relative factor prices, resulting in labor-intensive industries gradually replaced by capital-intensive ones in the closed economy. This mechanism is first articulated in Ju, Lin and Wang (2015), which is a special case of our model in the following sense: when the job finding rate becomes sufficiently larger than the separation rate (and hence unemployment disappears), the model in this paper will degenerate to Ju, Lin and Wang (2015). Whereas full employment is typically assumed in the existing literature with or without labor market frictions, our paper is to the best of our knowledge the first to study how mismatch and resulting structural unemployment endogenously generate cyclical movement of aggregate unemployment rates in the process of structural change and how mismatch affects structural change and life-cycle dynamics of underlying industries. For example, Pissarides (2007) develops a three-sector model with search and match in the labor market. However, the labor market for manufacturing and market service is assumed to be integrated, so structural transformation from manufacturing to service itself does not generate structural unemployment as the two sectors are assumed to require the same skills in his model. Apart from this difference, another important difference from our model is that the driving force for structural transformation is the exogenous unbalanced productivity growth across sectors as labor is the only input. Zagler (2009) derives a hump-shaped pattern of unemployment rate by introducing costly job vacancies in the R&D sector following Romer's horizontal innovation model, so the driving force for sectoral reallocation is endogenous innovation instead of physical capital accumulation as we highlight.

Secondly, our paper is related to mismatch and fluctuation of unemployment rates in the macro labor literature, which typically focuses on mismatch between job-seekers and vacancies at the business cycle frequency (Shimer, 2007; Alvarez and Shimer, 2012). For example, mismatch in Shimer (2007) results from exogenous and random assignment of identical workers to ex ante symmetric segregated local labor markets, so there are more workers than vacancies in some markets and the opposite is true in others. The main purpose is to resolve the Shimer puzzle, that is, volatility of unemployment rate in the data is higher than what is typically implied by the Diamond-Mortensen-Pissarides model (Shimer, 2005). In contrast, mismatch in our paper is not quantity mismatch, bur rather skill mismatch between workers from old industries and job requirement in new industries. In our model, workers endogenously choose which sectors to seek jobs rather than being randomly assigned. More importantly, the motivations are different. Our model is not aiming to explain the Shimer puzzle at the business cycle frequency but rather to explain how the unemployment rate may exhibit cyclical movement at the lower frequency consistent with structural changes in industries. Restrepo (2015) show that the impact of great recessions on labor markets can be amplified with skill mismatch and generate higher volatility than the canonical Mortensen-Pissarides model, but they assume that the old jobs become obsolete exogenously and structural shift is also exogenous, whereas in our model jobs become obsolete endogenously and structural change is also endogenous. Observe that in our model the industries are heterogeneous in capital intensities and hence endogenously asymmetric in their sizes and dynamics. The cycle of unemployment rate occurs endogenously and repetitively in our model due to the recurrent cyclical change in the number of low-skilled unmeployment, which in turn stems from skill mismatch and on-job learning in the determinstic process of labor moving from old industries to new and more capital-intensive industries. In marked constrast, the existing pertinent literature on cyclical unemployment typically assumes that industries are ex ante symmetric and that cyclical movement of unemployment is caused by the stochastic sectoral shifts due to idiocyncratic exogenous industry-specific shocks (also see, Lucas and Prescott, 1974; Lilien, 1982).

The rest of the paper is structured as follows: Section 2 presents a benchmark model, in which the job finding rate is exogenous and constant and there are finite industries. Section 3 shows that the constant job finding rate can be rationalized as an equilibrium outcome in the standard Diamond-Mortensen-Pissaridies model and that the cyclical unemployment rate obtained in Sector 2 remains valid when there are infinite number of industries. Section 4 concludes. Technical proofs are mostly delegated to the Appendix.

2 Benchmark Model

In this section, we develop a multi-industry dynamic model with imperfect labor markets. The analysis will focus on analytical characterization of the endogenous aggregate umemployment and industrial dynamics along the growth path. More specifically, in the model industries are heterogeneous in capital intensities and structural change (sectoral reallocation across industries) is mainly driven by capital accumulation, which is essentially same as Ju, Lin and Wang (2015). But the model has two new features: (1) skills are industry-specific in the sense that all workers newly relocated from other industries are initially unskilled workers, who may become skilled workers at a poisson rate; (2) labor markets are frictional in the sense that the job finding rate of skilled workers is higher than that of unskilled workers in any industry.

2.1 Environment

Consider a continuous-time economy with unit mass of identical households. Each household is initially endowed with physical capital K, low-skilled labor L_l and high-skilled labor L_h. There are two sectors in the economy. One sector produces capital goods and the other produces consumption goods. Capital goods and consumption goods are distinct in nature and not substitutable. Capital goods are produced using an AK technology. Let $K(t)$ denotes capital stock available at the beginning of time t, then the output flow coming out of the capital-good sector is $AK(t)$, where parameter A is positive, capturing the rate of the investment-specific technological progress net of depreciation (if any). A larger A implies a more efficient capital goods production. The newly produced capital flow is split between two different usages:

$$AK(t) = I(t) + E(t), \tag{1}$$

where $I(t)$ denotes investment in capital and $E(t)$ denotes the flow of capital devoted to the producton of consumption goods at t. $E(t)$ fully depreciates, so $K(t)$ evolves as follows

$$\dot{K}(t) = AK(t) - E(t). \tag{2}$$

The consumption good, denoted by X, is produced by linearly combining the output from four industries:

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

where x_n denotes the intermediate input from industry n and is non-negative for any $n = 0, 1, 2, 3$. Only the final good X can be used for consumption. Consumption goods and all the intermediate inputs are non-storable. All technologies of intermediate inputs exhibit constant returns to scale as follows:

$$F_n(k, l) = \begin{cases} l & \text{if} \quad n = 0 \\ \lambda^n \min\{\frac{k}{a^n}, l\} & \text{if} \quad n = 1, 2 \\ \lambda^3 \frac{k}{a^3} & \text{if} \quad n = 3 \end{cases}, \tag{4}$$

where we impose $a > \lambda > 1$. Observe that higher-indexed industries are not only more capital intensive for $n = 0, 1, 2, 3$ (because $a > 1$) but also yields higher labor productivity (because $\lambda > 1$) for $n = 0, 1, 2$. The assumption $a > \lambda$ rules out the trivial case that only the most capital intensive industry (industry 3) would produce if capital is used to produce consumption goods. Later on, we will extend the model into infinite industries.

Households are infinitely-lived and their preferences over consumption streams can be ordered by

$$\int_0^\infty \frac{C(t)^{1-\sigma}-1}{1-\sigma} e^{-\rho t} dt, \tag{5}$$

where σ is the reciprocal of the intertemporal elasticity of substitution and ρ is time discount rate. $\sigma > 0$ and $\rho > 0$. Following conventions, we assume $0 < A - \rho < \sigma A$ to ensure positive but non-explosive consumption growth.

Labor markets are frictional and separate for different industries ($n = 0, 1, 2$). Labor markets for industries 0 and 1 are well functioning, so no mismatch occurs and the job finding follows a Poisson process with constant rate f for all workers in those two industries[1]. However, in industry 2 low-skilled workers suffer from mismatch so their job finding rate is πf, lower than that for high-skilled workers f, where $\pi \in (0, 1)$. Smaller π means more severe mismatch. This is the only difference mismatch makes. All employed workers in the same industry have identical productivity, independent of their skill levels. Jobs separate at an exogenous rate δ for all workers in all industries. Low-skilled workers in industry 2, when hired, would benefit from on-job learning and become high-skilled workers permanently at an exogenous Poisson rate ξ in that industry. A higher ξ implies higher efficiency of on-job learning. The following figure summarizes the model environment for the consumption good sector.

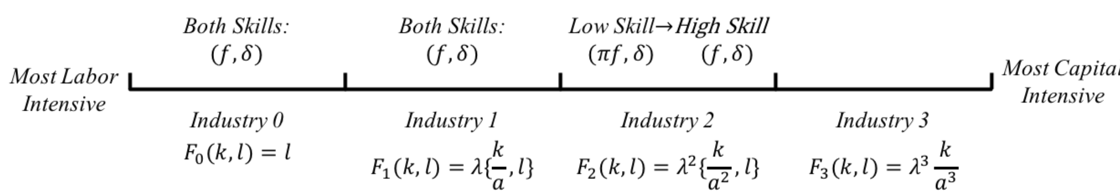

Fig. 1 *Environment in the consumption good sector with multiple industries*

Skills are industry-specific, so when a worker moves to a new industry, she automatically becomes a low-skilled worker for the new industry no matter whether she was high-skilled or low-skilled in the old industry.

Let $U_h(t)$ and $U_l(t)$ denote, respectively, the number of unemployed high-skilled workers and unemployed low-skilled workers in the whole economy at time t. Let $E_j^i(t), U_j^i(t)$ and $L_j^i(t)$ denote, respectively, the numbers of employed workers, unemployed workers, and all workers (employed and unemployed),

[1]The job finding rate is assumed constant not just for simplicity. We show in the next section that it is consistent with the socially efficient allocation in the model of Mortensen and Pissarides (1994) with a matching function.

with skill level i in the labor market for industry j at time t , for $i \in \{l, h\}$, $j \in \{0, 1, 2\}$. Obviously, $L_j^i(t) = U_j^i(t) + E_j^i(t)$.

Consider the benevolent social planner problem. The planner decides the optimal flow of consumption $C(t)$,investment $I(t)$, capital for consumption production $E(t)$, and how to allocate unemployed workers $U_h(t)$ and $U_l(t)$ into four industries to maximize the utility of a representative household as shown in (**??**).

The social planner problem becomes

$$\max_{C,} \int_0^\infty e^{-\rho t} \frac{C^{1-\sigma} - 1}{1 - \sigma} dt \tag{6}$$

subject to

$$\dot{K} = AK - E(C, E_0^l + E_0^h, E_1^l + E_1^h, E_2^l + E_2^h), \tag{7}$$

$$\dot{E}_0^l = fU_l(1 - \lambda_1^l - \lambda_2^l) - \delta E_0^l, \tag{8}$$

$$\dot{E}_0^h = fU_h(1 - \lambda_1^h - \lambda_2^h) - \delta E_0^h, \tag{9}$$

$$\dot{E}_1^l = fU_l\lambda_1^l - \delta E_1^l, \tag{10}$$

$$\dot{E}_1^h = fU_h\lambda_1^h - \delta E_1^h, \tag{11}$$

$$\dot{E}_2^l = \pi fU_l\lambda_2^l - \delta E_2^l - \xi E_2^l, \tag{12}$$

$$\dot{E}_2^h = fU_l\lambda_2^h - \delta E_2^h + \xi E_2^l, \tag{13}$$

$$\dot{U}_h = -fU_h + \delta \sum_{j=0}^{2} E_j^h, \tag{14}$$

$$\dot{U}_l = \delta \sum_{j=0}^{2} E_j^l - fU_l(1 - \lambda_2^l) - \pi fU_l\lambda_2^l, \tag{15}$$

$$0 \leq \lambda_1^i, 0 \leq \lambda_2^i \leq 1, \lambda_1^i + \lambda_2^i \leq 1, i \in \{l, h\} \tag{16}$$

$$\text{given } K(0), U_i(0), E_j^i(0), i \in \{l, h\}, j \in \{0, 1, 2\} \tag{17}$$

where λ_j^i denote the fraction of unemployed workers with skill level $i \in \{l, h\}$ who seek jobs in industry $j \in \{0, 1, 2\}$ in the total number of unemployed workers with skill level i the whole economy. The evolutions of all state variables are shown in the above differential equations (**??**)-(**??**). More specifically, (**??**) describes how capital stock evolves over time, where $E(C, E_0^l + E_0^h, E_1^l + E_1^h, E_2^l + E_2^h)$ is the minimum amount of capital required to produce final consumption C when current total employment in the first three industries are given by $E_0^l+E_0^h$, $E_1^l+E_1^h$ and $E_2^l+E_2^h$, respectively. (**??**) and (**??**) with the assumption that $a>\lambda>1$ jointly imply the following:

$$E(.)=\begin{cases} \frac{a}{\lambda}C-\frac{a}{\lambda}(E_0^l+E_0^h), & \text{if} \quad C_0<C\leq C_1 \\ \frac{a^2}{\lambda^2}C-\frac{a^2}{\lambda^2}(E_0^h+E_0^l)-\frac{a(a-\lambda)}{\lambda}(E_1^h+E_1^l), & \text{if} \quad C_1<C\leq C_2 \\ \frac{a^3}{\lambda^3}C-\frac{a^3}{\lambda^3}(E_0^h+E_0^l)-\frac{a(a^2-\lambda^2)}{\lambda^2}(E_1^h+E_1^l)-\frac{a^2(a-\lambda)}{\lambda}(E_2^h+E_2^l), & \text{if} \quad C_2<C \end{cases}$$

where

$$C_0=(E_0^l+E_0^h), C_1=(E_0^l+E_0^h)+\lambda(E_1^l+E_1^h), C_2=(E_0^l+E_0^h)+\lambda(E_1^l+E_1^h)+\lambda^2(E_2^l+E_2^h).$$

The next six state equations are about the dynamics of employment and unemployment for both skilled labor in the three industries. For example, (**??**) states that $\dot{E}_0^l$, the change in the total number of employed low-skilled workers in industry 0, is equal to the new employment flow net of exogenous job separation δE_0^l at each time point. There are $U_l\lambda_0^l$ low-skilled unemployed workers in industry 0, where $\lambda_0^l=1-\lambda_1^l-\lambda_2^l$, so the flow of new employment flow in that industry is $fU_l(1-\lambda_1^l-\lambda_2^l)$ because of no mismatch. The next three state equations (**??**)-(**??**) can be similarly interpreted. The next two equations state how total unemployment changes for the two skill levels. For example, (**??**) states that the change in low-skilled employment in industry 2 is equal to the new employment flow $\pi fU_l\lambda_2^l$ net of exogenous job separation (δE_2^l) and skill upgrade due to on-job learning (ξE_2^l). Note that mismatch takes place in this industry for low-skilled workers. (**??**) shows that the change in aggregate low-skilled unemployment is equal to the new unemployment flow due to exogenous separation in the three industries net of new employment flow in industries 0 and 1 ($fU_l(1-\lambda_2^l)$) and new employment flow in industry 2 ($\pi fU_l\lambda_2^l$).

We show that the decentralized competitive equilibrium can be characterized by resorting to the above social planner's problem.

Proposition 1 *There exists a decentralized equilibrium allocation which is the same as the social planner's allocation.*

Proof. Please see the appendix. ■

The intuition is as follows. Hosios' condition states that the decentralized market equilibrium achieves the first best when the surplus share of workers in the Nash wage bargaining is equal to the elasticity with respect to unemployment in the matching function. It is because externality of unemployed workers' additional search on firms is exactly internalized via the wage bargaining power distribution, so workers have the same incentive as the social planner. Our model is an extreme case where the bargaining power parameter for workers and the matching elasticity with respect to unemployment are both equal to one. Note that there exists no externality of additional search because job finding rates are constant, independent of unemployment. Meanwhile, firms do not need to pay any cost to post vacancies and earn zero net profits. As a result, workers receive all the surplus and there is no distortion of incentives.

Now we try to characterize the transitional dynamics for the social planner's problem. It is well known that analytical characterization of the whole transitional dynamic path is unwieldy even in a one-sector environment, let alone a multi-sector environment such as in our case. To sharpen the analysis, however, we follow Restrepo (2015) by focusing on the scenario in which gross flows between employment and unemployment are sufficiently large so that labor markets adjust fast enough to fully catch up with changes in the factor endowment. In other words, labor allocations resemble the steady state equilibrium for any given factor endowment structure. It simplifies the analysis, which allows us to derive a tractable solution for the dynamic growth path for all industries and the aggregate economy. Moreover, Davis and Haltiwanger (1990) show that the gross flows are indeed large in U.S. markets, so it is also empirically relevant. Let $f = \kappa\widehat{f}$, $\delta = \kappa\widehat{\delta}$, and suppose $\kappa \to \infty$ so that the gross flows between employment and unemployment are enormous. The following proposition shows that the dynamic optimization problem degenerates to static optimization at each instant. First, we define the optimal steady state:

Definition 1 *In an economy with factor endowment vector (E, L^l, L^h), where E denotes the capital flow used* ***only*** *to produce consumption goods, L^l and L^h denote high-skilled and low-skilled workers respectively. The job finding rate is equal to $\widehat{f}$ for all workers in all industries, except that the job finding rate of low-skilled workers is $\pi\widehat{f}$ in industry 2. All workers have the same exogenous job separation rate $\widehat{\delta}$ in all industries. Production technologies are given by (**??**) and (**??**). The optimal instantaneous equilibrium is the instantenous equilibrium in which the output of final consumption commodity is maximized given the factor endowment (E, L^l, L^h).*

Proposition 2 *For the instantaneous equilibrium (where $f = \kappa\widehat{f}$, $\delta = \kappa\widehat{\delta}$, and $\kappa \to \infty$), Dynamic optimization requires that an optimal steady state equilibrium to produce $C(t)$ at each instant if $\xi < \bar{\xi}$, where $\bar{\xi} \equiv \frac{(a-\lambda)(a\lambda-(a+\lambda)+\frac{\pi\widehat{f}+\widehat{\delta}}{\pi\widehat{f}+\pi\widehat{\delta}}}{\bar{\mu_u^h}}$ and $\bar{\mu_u^h}$ is the value of the unemployed high-skilled worker at $t = 0$.*

Proof. *Please see the appendix.* ■

When the job finding rate f and separation rate δ are sufficiently large, the employment and unemployment reach the instantaneous steady state almost immediately. Since the dynamic optimization requires labor markets adjust efficiently to cut down the flow of capital in the consumption goods sector, it is equivalent to finding an optimal steady state equilibrium to produce the required consumption goods with minimal capital flow. $\xi < \bar{\xi}$ is imposed to ensure that the instant incentive is sufficient for the infinite-horizon incentive. Since low-skilled workers in industry 2 have extra gains as their on-job learning can change their skill type, the condition $\xi < \bar{\xi}$ states that it does not dominate the benefit of an instant reduction in capital flows. For the rest of the paper, we always assume that $f = \kappa \widehat{f}$, $\delta = \kappa \widehat{\delta}$, $\kappa \to \infty$, and $\xi < \bar{\xi}$.

To derive the optimal dynamic path, we must first characterize the optimal instantaneous equilibrium, that is, we must solve the following problem:

$$C = \max_{E_j^i, K_j} \{(E_0^l + E_0^h) + \lambda \min\{E_1^l + E_1^h, \frac{K_1}{a}\} + \lambda^2 \min\{E_1^l + E_2^l, \frac{K_2}{a^2}\} + \lambda^3 \frac{K_3}{a^3}\} \tag{18}$$

subject to

$$K_1 + K_2 + K_3 \leq E, \tag{19}$$

$$E_0^l \frac{\widehat{f} + \widehat{\delta}}{\widehat{f}} + E_1^l \frac{\widehat{f} + \widehat{\delta}}{\widehat{f}} + E_2^l \frac{\pi \widehat{f} + \widehat{\delta}}{\pi \widehat{f}} \leq L^l, \tag{20}$$

$$\left(E_0^h + E_1^h + E_2^h\right) \frac{\widehat{f} + \widehat{\delta}}{\widehat{f}} \leq L^h. \tag{21}$$

where (**??**) is the resource constraint for physical capital, which is allocated to industries 1, 2 and 3; (**??**) is the resource constraint for low-skilled workers, which are allocated to industries 0, 1 and 2. Observe that $E_0^l \frac{\widehat{f}+\widehat{\delta}}{\widehat{f}}$ low-skilled workers must be allocated to industry 0 in order to have E_0^l low-skilled workers employed in that industry in the instantaneous equilibrium. The second term on the left-hand side in (**??**) is the amount of low-skilled workers allocated to industry 1. The third term is the amount of low-skilled workers allocated to industry 2, where mismatch occurs; (**??**) is the resource constraint for high-skilled workers, which are allocated to industries 0, 1 and 2. Since ξ is sufficiently small, low-skilled workers in industry 2 cannot become high-skilled in the instantaneous equilibrium. That is why ξ does not show up in the above optimization problem. The following proposition fully characterizes the optimal instantaneous equilibrium.

Proposition 3 *In the optimal instantaneous equilibrium with given capital endowment for consumption production E, endowment of low-skilled labor L_l and high-skilled labor L_h, the optimal resource allocation is summarized in Table 1.*

$0 \le E \le \frac{\hat{f}}{\hat{f}+\hat{\delta}} a(L^l + L^h)$	$\frac{\hat{f}}{\hat{f}+\hat{\delta}} a(L^l + L^h) < E \le \frac{\hat{f}}{\hat{f}+\hat{\delta}} (a^2 L^l + aL^h)$
$C = \frac{\lambda - 1}{a} E + \frac{\hat{f}}{\hat{f}+\hat{\delta}} (L^l + L^h)$	$C = \frac{\lambda^2 - \lambda}{a^2 - a} E + \frac{\hat{f}}{\hat{f}+\hat{\delta}} \frac{\lambda(a-\lambda)}{a-1} (L^l + L^h)$
$\Leftrightarrow E_{0,1} = \frac{a}{\lambda - 1} C - \frac{a}{\lambda - 1} \frac{\hat{f}}{\hat{f}+\hat{\delta}} (L^l + L^h)$	$\Leftrightarrow E_{1,2h} = \frac{a^2 - a}{\lambda^2 - \lambda} C - \frac{\hat{f}}{\hat{f}+\hat{\delta}} \frac{a(a-\lambda)}{\lambda - 1} (L^l + L^h)$
$\frac{\hat{f}}{\hat{f}+\hat{\delta}} (a^2 L^l + aL^h) < E \le a^2 (\frac{\hat{f}}{\hat{f}+\hat{\delta}} L^h + \frac{\pi \hat{f}}{\pi \hat{f}+\hat{\delta}} L^l)$	$E > a^2 (\frac{\hat{f}}{\hat{f}+\hat{\delta}} L^h + \frac{\pi \hat{f}}{\pi \hat{f}+\hat{\delta}} L^l)$
$C = \frac{\lambda^2 \frac{\pi \hat{f}}{\pi \hat{f}+\hat{\delta}} - \lambda \frac{\hat{f}}{\hat{f}+\hat{\delta}}}{a^2 \frac{\pi \hat{f}}{\pi \hat{f}+\hat{\delta}} - a \frac{\hat{f}}{\hat{f}+\hat{\delta}}} E + \frac{\hat{f}}{\hat{f}+\hat{\delta}} \frac{\lambda(a-\lambda)}{a \frac{\pi \hat{f}}{\pi \hat{f}+\hat{\delta}} - \frac{\hat{f}}{\hat{f}+\hat{\delta}}} \left(\frac{\pi \hat{f}}{\pi \hat{f}+\hat{\delta}} L^l + \frac{\hat{f}}{\hat{f}+\hat{\delta}} L^h \right)$	$C = \frac{\lambda^3}{a^3} E + \frac{\lambda^2 (a-\lambda)}{a} \left(\frac{\pi \hat{f}}{\pi \hat{f}+\hat{\delta}} L^l + \frac{\hat{f}}{\hat{f}+\hat{\delta}} L^h \right)$
$\Leftrightarrow E_{1,2l} = \frac{a^2 \frac{\pi \hat{f}}{\pi \hat{f}+\hat{\delta}} - a \frac{\hat{f}}{\hat{f}+\hat{\delta}}}{\lambda^2 \frac{\pi \hat{f}}{\pi \hat{f}+\hat{\delta}} - \lambda \frac{\hat{f}}{\hat{f}+\hat{\delta}}} C - \frac{\hat{f}}{\hat{f}+\hat{\delta}} \frac{a(a-\lambda)}{\lambda \frac{\pi \hat{f}}{\pi \hat{f}+\hat{\delta}} - \frac{\hat{f}}{\hat{f}+\hat{\delta}}} \left(\frac{\pi \hat{f}}{\pi \hat{f}+\hat{\delta}} L^l + \frac{\hat{f}}{\hat{f}+\hat{\delta}} L^h \right)$	$\Leftrightarrow E_{23} = \frac{a^3}{\lambda^3} C - \frac{a^2 (a-\lambda)}{\lambda} \left(\frac{\pi \hat{f}}{\pi \hat{f}+\hat{\delta}} L^l + \frac{\hat{f}}{\hat{f}+\hat{\delta}} L^h \right)$

Table 1 *optimal instantaneous equilibrium*

Proof. *Please see the appendix* ■

As shown in Table 1, when $E \in [0, \frac{\widehat{f}}{\widehat{f}+\widehat{\delta}} a(L^l + L^h)]$, only industry 0 and industry 1 coexist and no mismatch occurs, so the skill structure of the labor force does not matter. When $E \in (\frac{\widehat{f}}{\widehat{f}+\widehat{\delta}} a(L^l + L^h), \frac{\widehat{f}}{\widehat{f}+\widehat{\delta}} (aL^l + a^2 L^h)]$, only industry 1 and industry 2 coexist, and all workers in industry 2 are high-skilled. When $E \in (\frac{\widehat{f}}{\widehat{f}+\widehat{\delta}} (aL^l + a^2 L^h), a^2 (\frac{\widehat{f}}{\widehat{f}+\widehat{\delta}} L^h + \frac{\pi \widehat{f}}{\pi \widehat{f}+\widehat{\delta}} L^l)]$, only industry 1 and industry 2 coexist, and all workers in industry 1 are low-skilled. When $E \in [\, a^2 (\frac{\widehat{f}}{\widehat{f}+\widehat{\delta}} L^h + \frac{\pi \widehat{f}}{\pi \widehat{f}+\widehat{\delta}} L^l), \infty)$, only industry 2 and industry 3 coexist, and all labors are in industry 2. Observe that the total consumption C as a function of E, L^l and L^h has different functional forms in the above four different scenarios, which reflects the endogenous differences in the underlying industrial structures that are in turn determined by the factor endowment structure. In particular, when $\widehat{f} \to \infty$ while keeping $\widehat{\delta} < \infty$, mismatch disappears and skill types are irrelevant, which is exactly the case in Ju, Lin and Wang (2015), so Table 1 is also degenerated to the static equilibrium in Ju, Lin Wang (2015).

Let t_1 denote the time when all workers stay at industry 1, which is also the time point when $E = \frac{\widehat{f}}{\widehat{f}+\widehat{\delta}} a(L^l + L^h)$ as implied in Table 1. Let $t_{2,h}$ denote the first time point when all high-skilled workers are in industry 2, which is also the time point when $E = \frac{\widehat{f}}{\widehat{f}+\widehat{\delta}} (aL^l + a^2 L^h)$. Let $t_{2,l}$ denote the first time when all low-skilled workers stay in industry 2, which is also the first time point when all workers in the economy are in industry 2. The social planner's problem can be rewritten as

$$\max_C \int_0^{t_1} e^{-\rho t}\frac{C^{1-\sigma}-1}{1-\sigma}dt + \int_{t_1}^{t_{2,h}} e^{-\rho t}\frac{C^{1-\sigma}-1}{1-\sigma}dt + \int_{t_{2,h}}^{t_{2,l}} e^{-\rho t}\frac{C^{1-\sigma}-1}{1-\sigma}dt + \int_{t_{2,l}}^{+\infty} e^{-\rho t}\frac{C^{1-\sigma}-1}{1-\sigma}dt, \quad (22)$$

subject to

$$\dot{K} = \begin{cases} AK - E_{0,1}(C, L_l, L_h) & \text{if} \quad t \leq t_1 \\ AK - E_{1,2h}(C, L_l, L_h) & \text{if} \quad t_1 < t \leq t_{2,h} \\ AK - E_{1,2l}(C, L_l, L_h) & \text{if} \quad t_{2,h} < t \leq t_{2,l} \\ AK - E_{2,3}(C, L_l, L_h) & \text{if} \quad t > t_{2,l} \end{cases}. \quad (23)$$

where $E_{m,n}(\cdot)$ are provided in the last row in Table 1 and t_1, $t_{2,h}$, $t_{2,l}$ are endogenously determined. The Euler equation implies the following constant consumption growth rate:

$$g_c \equiv \frac{\dot{C}}{C} = \frac{A-\rho}{\sigma}. \quad (24)$$

Before $t_{2,h}$ there are no low-skilled workers in industry 2 so there is no mismatch, therefore, the aggregate unemployment is constant and equal to $\frac{\widehat{\delta}}{\widehat{f}+\widehat{\delta}}L$. After $t_{2,h}$, low-skilled workers move from industry 1 to industry 2 and suffer mismatch. The aggregate unemployment can be written as

$$U = \frac{\widehat{\delta}}{\widehat{f}+\widehat{\delta}}L + (\frac{\widehat{\delta}}{\pi\widehat{f}+\widehat{\delta}} - \frac{\widehat{\delta}}{\widehat{f}+\widehat{\delta}})L_2^l \blacksquare \quad (25)$$

where the first term on the right hand side ($\frac{\widehat{\delta}}{\widehat{f}+\widehat{\delta}}L$), is the aggregate unemployment for the whole economy if mismatch is absent (that is, when $L_2^l = 0$), and the second term is the extra amount of unemployment due to mismatch of low-skilled workers in industry 2 (when $L_2^l > 0$) relative to the case when no mismatch exists. The first term captures the frictional unemployment due to labor market frictions in job search and job destruction. The second term is referred to as structural unemployment because it is due to due to mismatch. Correspondingly, $\frac{\widehat{\delta}}{\widehat{f}+\widehat{\delta}}$ is the steady state unemployment rate without mismatch and $\frac{\widehat{\delta}}{\pi\widehat{f}+\widehat{\delta}}$ is the steady state unemployment rate if all workers suffer from mismatch. The gap between the aggregate unemployment rate and the steady state unemployment rate without mismatch, $\frac{U}{L} - \frac{\widehat{\delta}}{\widehat{f}+\widehat{\delta}}$, is referred to as the structural unemployment rate. Let $U^* = \frac{\widehat{\delta}}{\widehat{f}+\widehat{\delta}}L + (\frac{\widehat{\delta}}{\pi\widehat{f}+\widehat{\delta}} - \frac{\widehat{\delta}}{\widehat{f}+\widehat{\delta}})L^l$ denote the hypothetical upper bound unemployment, which occurs only when all low-skilled workers are in industry 2. Correspondingly, $\frac{U*}{L}$ is the upper bound unemployment rate. Consider the simplest case when all workers are low-skilled at time 0, that is, $L^l(0) = L$, so $t_{2,h} = t_1$ as there are no high-skilled workers in the instantenous equilibrium. From Table 1 and (**??**), we obtain the time path for the number of employed low-skilled workers in industry 2 as follows (The proof is delegated to the Appendix)

$$E_2^l(t) = min\{N_1 e^{g_c(t-t_1)} - (N_1 + N_2 L) + N_2 L^l, \frac{\pi\widehat{f}}{\pi\widehat{f}+\widehat{\delta}}L^l\}, \quad (26)$$

where

$$N_1 = \frac{L\frac{\widehat{f}}{\widehat{f}+\widehat{\delta}}\frac{\pi\widehat{f}}{\pi\widehat{f}+\widehat{\delta}}}{\lambda\frac{\pi\widehat{f}}{\pi\widehat{f}+\widehat{\delta}} - \frac{\widehat{f}}{\widehat{f}+\widehat{\delta}}}, N_2 = \frac{(\lambda-1)\frac{\widehat{f}}{\widehat{f}+\widehat{\delta}}\frac{\pi\widehat{f}}{\pi\widehat{f}+\widehat{\delta}}}{\lambda\frac{\pi\widehat{f}}{\pi\widehat{f}+\widehat{\delta}} - \frac{\widehat{f}}{\widehat{f}+\widehat{\delta}}}. \tag{27}$$

As low-skilled workers in industry 2 gradually become high-skilled workers through on-job learning, so

$$\dot{L}^l = -\xi E_2^l. \tag{28}$$

Substituting (**??**) into the above equation yields

$$\dot{L}_l = -\xi \min\{N_1 e^{g_c(t-t_1)} - (N_1 + N_2 L) + N_2 L^l, \frac{\pi\widehat{f}}{\pi\widehat{f}+\widehat{\delta}} L^l\}. \tag{29}$$

Solving the above differential equation with the condition that $L^l(t_1) = L$, we obtain

$$E_2^l(t) = \begin{cases} \frac{N_1 g_c}{N_2\xi + g_c}(e^{g_c(t-t_1)} - e^{-N_2\xi(t-t_1)}) & \text{when} \quad t_1 < t \le t_{2,l} \\ e^{-\frac{\pi\widehat{f}}{\pi\widehat{f}+\widehat{\delta}}\xi(t-t_{2,l})}\frac{N_1 g_c}{N_2\xi + g_c}(e^{g_c(t_{2,l}-t_1)} - e^{-N_2\xi(t_{2,l}-t_1)}) & \text{when} \quad t > t_{2,l} \end{cases} \tag{30}$$

and

$$L^l(t) = \begin{cases} \frac{-N_1}{N_2(N_2\xi+g_c)}(\xi N_2 e^{g_c(t-t_1)} + g_c e^{-N_2\xi(t-t_1)}) + \frac{N_1}{N_2} + L & \text{if} \quad t_1 < t \le t_{2,l} \\ e^{-\frac{\pi\widehat{f}}{\pi\widehat{f}+\widehat{\delta}}\xi(t-t_{2,l})}[\frac{-N_1}{N_2(N_2\xi+g_c)}(\xi N_2 e^{g_c(t_{2,l}-t_1)} + g_c e^{-N_2\xi(t_{2,l}-t_1)}) + \frac{N_1}{N_2} + L] & \text{if} \quad t > t_{2,l} \end{cases}. \tag{31}$$

The above two equations are vital as they characterize the dynamics of the whole economy. For example, when $t_1 < t \le t_{2,l}$, (**??**) implies $\frac{dL^l(t)}{dt} < 0$ and (**??**) implies $\frac{dE_2^l(t)}{dt} > 0$. After $t_{2,l}$, they imply that both $E_2^l(t)$ and L^l decrease exponentially. These findings are summarized in the following proposition.

Proposition 4 *The aggregate unemployment rate is constant at $\frac{\widehat{\delta}}{\widehat{f}+\widehat{\delta}}$ till time t_1, after which the aggregate unemployment exhibits a hump-shaped time path: it first rises due to skill mismatch and then gradually declines because more and more low-skilled workers become high-skilled due to on-job learning.*

The unemployment rate before t_1 is flat at level $\frac{\widehat{\delta}}{\widehat{f}+\widehat{\delta}}$ because industry 2 does not exist in equilibrium and hence no mismatch occurs. After t_1, the unemployment rate changes non-monotonically, not because of any exogenous aggregate shocks, but rather because the number of low-skilled workers in industry 2 changes endogenously. Recall that low-skilled workers suffer from mismatch and hence a lower job finding rate than high-skilled workers in industry 2, but low-skilled workers could become high-skilled ones through on-job learning. The number of low-skilled workers in industry 2 (hence unemployment rate) first increases because workers move from industry 1 to industry 2 as capital increases. Later on, the number of low-skilled workers in industry 2 (hence the unemployment rate) declines because of on-job learning.

To see the results more intuitively, we plot the time path for the aggregate unemployment rate $\frac{U}{L}$ in Figure 2 (the solid curve). The downward dotted curve starting from $\frac{\widehat{\delta}}{\pi \widehat{f}+\widehat{\delta}}$ on the vertical axis is $\frac{U*}{L}$, the upper bound unemployment rate. When all low-skilled workers move into industry 2, the aggregate unemployment rate reaches its peak, and then declines monotonically and eventually converges to $\frac{\widehat{\delta}}{\widehat{f}+\widehat{\delta}}$ due to on-job learning.

There are three economic forces that jointly shape the labor market performance. The first force is mismatch measured by π. When $\pi \to 1$, the aggregate unemployment rate u^* converges to $\frac{\widehat{\delta}}{\widehat{f}+\widehat{\delta}}$. When mismatch is more serious (that is, π becomes smaller), the upper bound unemployment rate $\frac{U*}{L}$ goes up. The second force is capital goods production, the efficiency of which is measured by A. A more efficient production of capital goods (larger A) would imply a faster process of the sunrise industry replacing the sunset one. The third force is on-job learning, the speed of which is measured by ξ. A faster on-job learning (higher ξ) would imply that low-skilled workers become high-skilled workers more quickly in industry 2. The following proposition summarizes the results of comparative static analyses about the three forces.

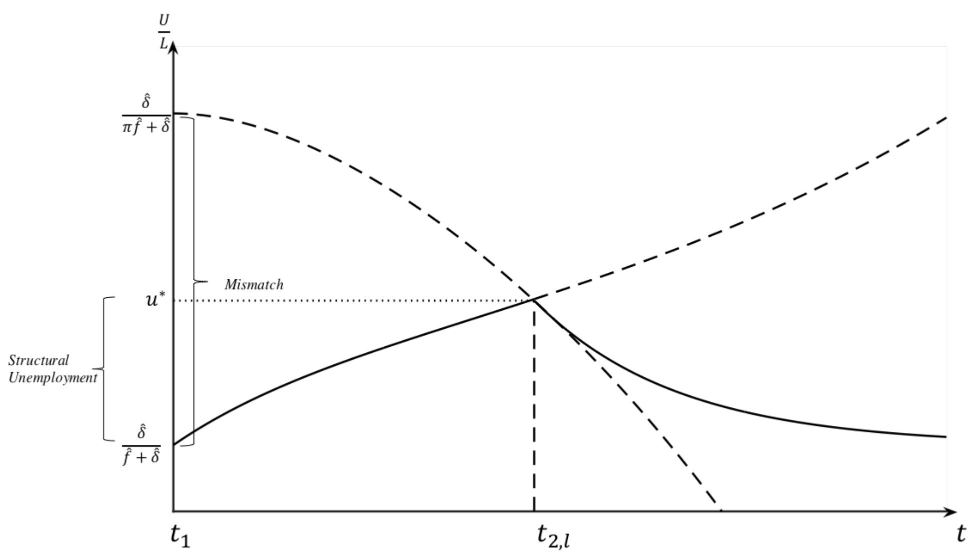

Fig. 2 *aggregate unemployment rate*

Proposition 5 *The following are true after t_1: (1) Learning efficiency: when ξ increases, the aggregate unemployment rate u shifts downward and reaches its peak later. (2) capital-goods productivity: when A increases, u first shifts upward and the peak value becomes higher, but the peak value is reached earlier. (3) Mismatch: When π increases, u decreases and reaches its peak earlier.*

Proof. Please see the appendix. ■

The intuition for the above proposition is the following:

(1) (learning efficiency) When the learning efficiency ξ increases, low-skilled workers turn into high-skilled workers more quickly, and therefore more workers will have a higher job finding rate, which results in a lower unemployment rate. To understand why the turning point for the unemployment rate is delayed, we should note that there are two competing forces when ξ increases. One the one hand, more labors are effectively employed in industry 2 as the number of high-skilled workers increase, which can sustain a longer period of expansion of industry 2 to support consumption growth before resorting to industry 3. This force tends to delay $t_{2,l}$. On the other hand, the total supply of low-skilled workers decreases faster, which tends to expedite the decline of the unemployment rate. It turns out that the first force dominates the second force. Figure 3 plots how the time path for the aggregate unemployment rate changes after on-job learning efficiency increases.

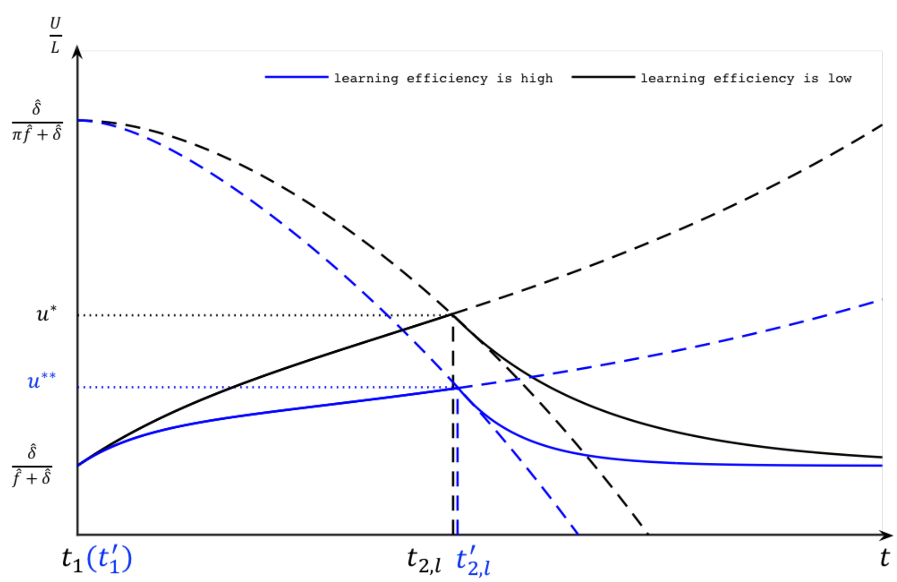

Fig. 3 *aggregate unemployment rate with respect to different learning efficiencies*

(2) (capital-goods productivity) When capital-goods productivity A increases, capital accumulates faster, which drives a more rapid industrial upgrading from industry 1 to the more capital-intensive industry 2. So low-skilled workers rush into industry 2 more quickly at the beginning, resulting in more mismatch and higher unemployment rate. But the peak value is reached earlier because the consumption growth rate is higher, so industry 3 emerges earlier to support the consumption growth, which also means that

employment in industry 2 reaches the peak at an earlier time. The unemployment rate reaches the highest value when the total employment in industry 2 reaches the peak.(see Figure 4).

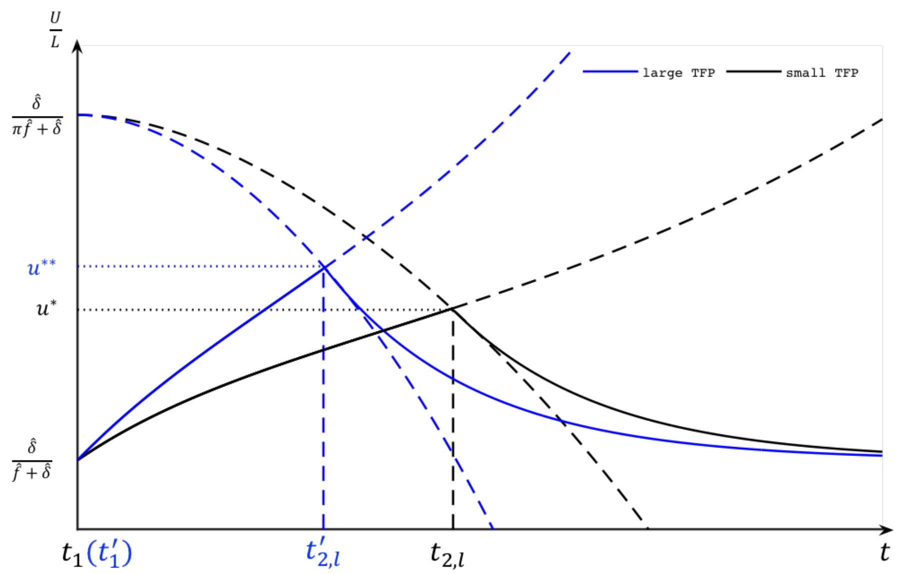

Fig. 4 *aggregate unemployment rate with respect to different TFP*

(3) (mismatch) When mismatch is more serious in industry 2 (that is, π decreases), low-skilled workers have a lower job finding rate in industry 2. As a result, the aggregate unemployment rate first shifts upward and the peak value of unemploymenr rate is also higher. As the consumption growth rate remains unchanged, so industry 3 emerges earlier to support consumption growth as the maximum amount of effective total employment in industry 2 decreases due to more severe mismatch. It implies that the unemployment rate starts to decline at an earlier time (see Figure 5).

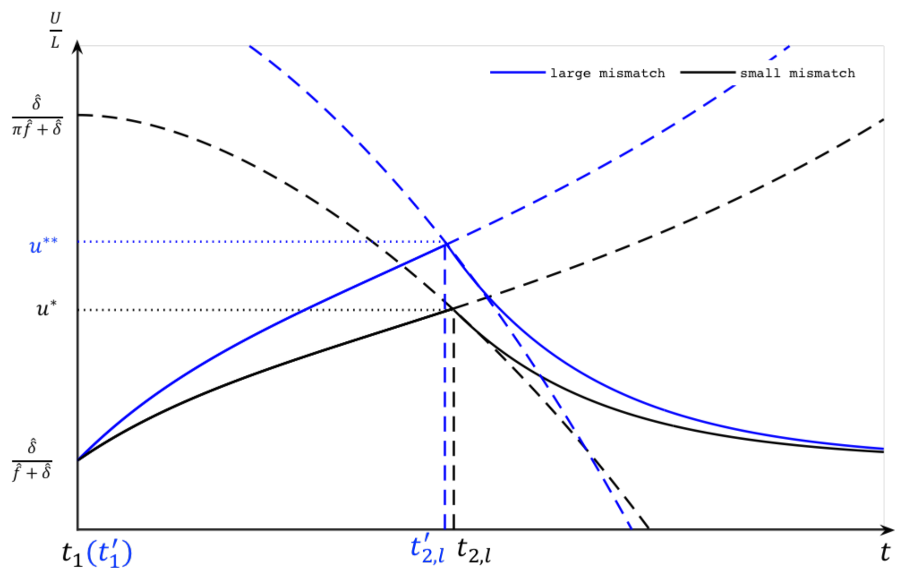

Fig. 5 *aggregate unemployment rate with respect to different mismatch*

3 Extensions

3.1 Matching function

In this section, we show that the assumption of constant job finding rate can be rationalized as an equilibirum outcome in the more general Diamond-Mortensen-Pissarides model setting. Suppose that the matching function is concave and constant returns to scale in each industry, we can prove that the job finding rate is unique and constant in the process of industry upgrading. To be specific, consider an economy with unit mass of identical households. Each household is initially endowed with capital K, low-skilled labor L_l and high-skilled labor L_h. Assume there are two sectors in the economy: one producing capital goods and another producing consumption goods. Capital goods are produced using an AK technology. The production function of the final commodity in consumption goods sector is $\sum_{n=0}^{3} x_n$, and the production function of intermediate goods n is $x_n = F_n(k, l)$. Labor markets are frictional: new matches between i skilled unemployed workers U_j^i in industry j, and job vacancies V_j^i in that industry are determined by the matching function $A_j^i m(U_j^i, V_j^i)$. $m(\cdot)$ is strictly increasing and strictly concave in both arguments, and it is constant return to scale. Suppose $A_j^i = A$ if $i \neq l$ or $j \neq 2$ and $A_2^l = \pi A$. That is, matching efficiency is lower for low-skilled workers in industry 2 due to mismatch. Let $q(\theta_j^i) \equiv \frac{m(U_j^i, V_j^i)}{U_j^i} = m(1, \theta_j^i)$ where $\theta_j^i = \frac{V_j^i}{U_j^i}$ denotes the market tightness for workers with skill level i in industry j. Matches are destroyed exogenously at rate δ in all industries. Posting vacancies is costly, and costs c (in terms of final consumption good) in any industry for all workers independent of skill types. The social planner's problem becomes

$$\max_C \int_{t=0}^{\infty} e^{-\rho t} \frac{C^{1-\sigma}-1}{1-\sigma} dt \tag{32}$$

subject to

$$\dot{K} = AK - E(C + \sum_{i\in\{l,h\}} \sum_{j=0}^{2} V_j^i c_j^i, E_0^l + E_0^h, E_1^l + E_1^h, E_2^l + E_2^h), \tag{33}$$

The evolution of labor markets is similar to the benchmark model except that the job finding rate now is $A_j^i q(\theta_j^i)$ rather than f or πf. Consider the limit case where $A_j^i = \kappa \widehat{A_j^i}$, $\delta = \kappa\widehat{\delta}$ and $\kappa \to \infty$. We can show that it is equivalent to finding an optimal steady state equilibrium which maximizes C given E, L_l and L_h if $\xi < \overline{\xi^*}$. The static optimization becomes

$$\max_{U_j^i, V_j^i, K_j} \{(E_0^l + E_0^h) + \lambda \min\{E_1^l + E_1^h, \frac{K_1}{a}\} + \lambda^2 \min\{E_1^l + E_2^h, \frac{K_2}{a^2}\} + \lambda^3 \frac{K_3}{a^3} - \sum_{i\in\{l,h\}} \sum_{j=0}^{2} V_j^i c\}, \tag{34}$$

subject to

$$\widehat{\delta} E_j^i = \widehat{A}_j^i q(\theta_j^i) U_j^i, i \in \{l,h\}, j \in \{0,1,2\} \tag{35}$$

$$\sum_{j=0}^{2} (E_j^i + U_j^i) \le L_i, i \in \{l,h\} \tag{36}$$

$$K_1 + K_2 + K_3 \le E. \tag{37}$$

We characterize properties of the optimal steady state equilibrium in the following proposition.

Proposition 6 *If the matching function $m(U_j^i, V_j^i)$ satisfies constant return to scale, strictly increasing and strictly concave in both arguments, the job finding rates $\widehat{A}_j^i q(\theta_j^i)$ and $\widehat{A}_{j+1}^i q(\theta_{j+1}^i)$ are uniquely determined and constant for any $i \in \{l,h\}$ in the optimal steady state equilibrium where industry j and industry $j+1$ coexist.*

Proof. Please see the appendix. ■

The intuition is as follows. The social planner optimally assigns the number of vacancies for each worker who moves from industry j to industry $j+1$ to maximize the net surplus of consumption goods. Since the matching functions and production functions are all constant return to scale, the market tightness is constant during thhe industrial upgrading process. Therefore, the job finding rate is endogneously constant, consistent with our benchmark model. What varies is the share of workers in industry j and in industry $j+1$. With a larger E, workers move from industry j to industry $j+1$ to absorb extra capital investments and produce more consumption goods.

3.2 Infinite-industry model

We extend our benchmark model to infinite industries. Everything is identical to the benchmark model except that now there are infinite number of intermediate goods in the consumption good sector, and the final consumption good is produced with the following technology:

$$X = \sum_{n=0}^{\infty} x_n, \tag{38}$$

where intermediate good x_n is produced with the following Leontief technology:

$$F_n(k,l) = \begin{cases} l & \text{if} \quad n = 0 \\ \lambda^n \min\{\frac{k}{a^n}, l\} & \text{if} \quad n \geq 1 \end{cases}. \tag{39}$$

Following Ju, Lin and Wang (2015), we impose $a-1 > \lambda > 1$ to rule out the trivial case that only most capital intensive goods are produced. Like in the benchmark model, when workers seek jobs in industry 0 or industry 1, they do not suffer skill mismatch and their job finding rate is f. Mismatch takes place only in industries $j \geq 2$. Skill is industry-specific. Workers can only learn the requisite skill of industry j for any $j \geq 2$ after they move into that industry and get employed. If a worker has acquired the skill in industry j, she becomes a high-skilled worker in that industry. High-skilled workers in industry j do not suffer from skill mismatch and their job finding rate is f. However, if a worker lacks the requisite skill in any industry $j \geq 2$, she is a low-skilled worker and has a lower job finding rate πf in that industry. She can turn into a high-skilled worker via on-job learning after she gets employed in that industry. No matter what skill type a worker has in industry j, when she first moves into industry $j+1$, she is low-skilled and has a lower job finding rate πf. Job separation rates are δ across industries independent of skill types of workers. Furthermore, we assume that mismatch is mild enough so that $\frac{\pi\widehat{f}+\pi\widehat{\delta}}{\pi\widehat{f}+\widehat{\delta}} > \frac{a+\lambda-1}{a\lambda}$ holds, which ensures that workers do not jump from industry j to industry $j+2$. Each household is initially endowed with capital $K(0)$ and labor L, and all labor is low-skilled initially for any industry j, $j \geq 2$. The household's preference is still given by (**??**). Figure 6 summarizes the model environment for the consumption good sector with infinite industries.

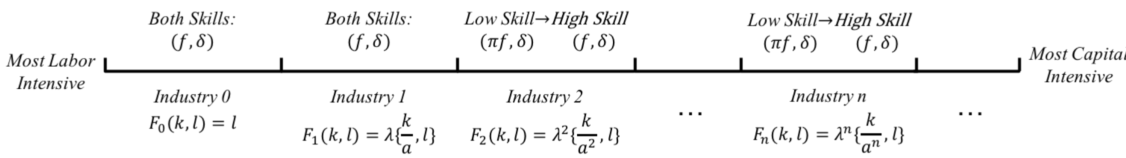

Fig. 6 *Environment in consumption good sector with infinite industries*

Like before, we focus on the limit case where $f = \kappa\widehat{f}$, $\delta = \kappa\widehat{\delta}$, $\kappa \to \infty$ and assume that $\xi < \bar{\xi}$. Following the same logic as in the benchmark model, we can easily establish the following results: (1) when industry j has only high-skilled (low-skilled) workers while industry $j+1$ has workers of both skill types, high-skilled (low-skilled) workers in industry j will first move into industry $j+1$; (2) when industry j has both low-skilled and high-skilled workers, low-skilled workers will first move into industry $j+1$ because their opportunity cost of switching industries is lower. We characterize the dynamics of aggregate unemployment rate and employment share by applying the forward induction method. For the upgrading process from industry 0 to industry 1, and then to industry 2, everything is identical to the bench model till all workers are in industry 2. After that, low-skilled workers in industry 2 first move into industry 3, followed by high-skilled workers. When all workers are in industry 3, low-skilled ones first move to industry 4, followed by high-skilled workers. This process continues repetitively. Let t^l_{j+1} denote the time point when all workers stay at industry j. It coincides with the time point when low-skilled workers in industry j begin to move into industry $j+1$. Let $\bar{L}^l_j$ and $\bar{L}^h_j$ denote respectively the number of low-skilled workers and the number of high-skilled workers in industry j at time t^l_{j+1}. Let t^h_{j+1} denote the time point when all low-skilled workers stay in industry $j+1$, which is also the time point when all workers in industry j are high-skilled. It coincides with the time point when high-skilled workers in industry j begin to move into industry $j+1$. Let $\bar{\bar{L}}^h_j$, $\bar{\bar{L}}^l_{j+1}$ and $\bar{\bar{L}}^h_{j+1}$ denote, respectively, the number of high-skilled workers in industry j, the number of low-skilled workers and the number of high-skilled workers in industry $j+1$ at time t^h_{j+1}. Technical proofs for this section are delegated into the appendix. The time path for the number of high-skilled workers in industry j is given by

$$L^h_j(t) = \begin{cases} \frac{\xi M_{j-1}}{g_c + \xi\frac{\widehat{f}}{\widehat{f}+\widehat{\delta}}} e^{g_c(t-t^l_j)} + \frac{\bar{L}^l_{j-1}}{\lambda-1} e^{-\xi\frac{\pi\widehat{f}}{\pi\widehat{f}+\widehat{\delta}}(t-t^l_j)} + (\frac{\bar{L}^h_{j-1}}{\lambda-1} - \frac{\xi M_{j-1}}{g_c+\xi\frac{\widehat{f}}{\widehat{f}+\widehat{\delta}}}) e^{-\xi\frac{\widehat{f}}{\widehat{f}+\widehat{\delta}}(t-t^l_j)} - \frac{L}{\lambda-1} & t \in [t^l_j, t^h_j) \\ (\bar{\bar{L}}^h_j + \frac{N_1}{N_2} - \frac{Q_{j-1}\xi}{N_2\xi+g_c}) e^{-N_2\xi(t-t^h_j)} + \frac{Q_{j-1}\xi}{N_2\xi+g_c} e^{g_c(t-t^h_j)} - \frac{N_1}{N_2} & t \in [t^h_j, t^l_{j+1}) \\ L - e^{-\frac{\pi\widehat{f}}{\pi\widehat{f}+\widehat{\delta}}\xi(t-t^l_{j+1})} \bar{L}^l_j - L^h_{j+1} & t \in [t^l_{j+1}, t^h_{j+1}) \\ L - L^l_{j+1} - L^h_{j+1} & t \in [t^h_{j+1}, t^l_{j+2}) \end{cases} \tag{40}$$

and the time path for the number of low-skilled workers in industry j is given by

$$L_j^l(t) = \begin{cases} \frac{\pi\widehat{f}+\widehat{\delta}}{\pi\widehat{f}}[\frac{M_{j-1}g_c}{g_c+\xi\frac{\widehat{f}}{\widehat{f}+\widehat{\delta}}}e^{g_c(t-t_j^l)} - \frac{\frac{\pi\widehat{f}}{\pi\widehat{f}+\widehat{\delta}}\bar{L}_{j-1}^l}{\lambda-1}e^{-\xi\frac{\pi\widehat{f}}{\pi\widehat{f}+\widehat{\delta}}(t-t_j^l)} - (\frac{\frac{\widehat{f}}{\widehat{f}+\widehat{\delta}}\bar{L}_{j-1}^h}{\lambda-1} - \frac{\xi M_{j-1}\frac{\widehat{f}}{\widehat{f}+\widehat{\delta}}}{g_c+\xi\frac{\widehat{f}}{\widehat{f}+\widehat{\delta}}})e^{-\xi\frac{\widehat{f}}{\widehat{f}+\widehat{\delta}}(t-t_j^l)}] & t \in [t_j^l, t_j^h) \\ \frac{\pi\widehat{f}+\widehat{\delta}}{\pi\widehat{f}}[\frac{Q_{j-1}g_c}{N_2\xi+g_c}e^{g_c(t-t_j^h)} - (N_1 + N_2\bar{\bar{L}}_j^h - \frac{Q_{j-1}\xi N_2}{g_c+\xi N_2})e^{-N_2\xi(t-t_j^h)}] & t \in [t_j^h, t_{j+1}^l) \\ e^{-\frac{\pi\widehat{f}}{\pi\widehat{f}+\widehat{\delta}}\xi(t-t_{j+1}^l)}\bar{L}_j^l - L_{j+1}^l & t \in [t_{j+1}^l, t_{j+1}^h) \\ 0 & t \in [t_{j+1}^h, t_{j+2}^l) \end{cases}, \tag{41}$$

where M_j and Q_j are given by

$$M_j \equiv \frac{\frac{\pi\widehat{f}}{\pi\widehat{f}+\widehat{\delta}}\bar{L}_j^l + \frac{\widehat{f}}{\widehat{f}+\widehat{\delta}}\bar{L}_j^h}{\lambda - 1}; Q_j \equiv \frac{(\bar{L}_j^h\frac{\widehat{f}}{\widehat{f}+\widehat{\delta}} + \lambda\bar{L}_{j+1}^l\frac{\pi\widehat{f}}{\pi\widehat{f}+\widehat{\delta}} + \lambda\bar{L}_{j+1}^h\frac{\widehat{f}}{\widehat{f}+\widehat{\delta}})\frac{\pi\widehat{f}}{\pi\widehat{f}+\widehat{\delta}}}{\lambda\frac{\pi\widehat{f}}{\pi\widehat{f}+\widehat{\delta}} - \frac{\widehat{f}}{\widehat{f}+\widehat{\delta}}}. \tag{42}$$

Since initial values $\bar{L}_1^l = 0$ and $\bar{L}_1^h = L$ are known, we use above equations to uniquely determine the values of $\bar{L}_j^l$ and $\bar{L}_j^h$ for any $j \geq 2$ by forward induction method. Given t_1^h, the turning points t_j^l and t_j^h for any $j \geq 2$ can also be uniquely determined simultaneously. Figure 7 shows the life cycle of high-skilled and low-skilled workers in each industry. We prove that both L_j^h and L_j^l are increasing from t_j^l to t_j^h because low-skilled workers in industry $j-1$ move into industry j and some of them become high-skilled workers. From t_j^h to t_{j+1}^l, L_j^h and L_j^l continue to increase because high-skilled workers in industry $j-1$ move into industry j and become low-skilled workers in the new industry, but some of them become high-skilled due to on-job learning. However, from t_{j+1}^l to t_{j+1}^h, low-skilled workers in industry j move into industry $j+1$ and some of the remaining low-skilled workers become high-skilled in industry j within this period. From t_{j+1}^h to t_{j+2}^l, high-skilled workers in industry j move into industry $j+1$ and industry j gradually declines and eventually disappears. The life span of an industry $j \geq 2$ is equal to $t_{j+2}^l - t_j^l$, the length between the time point when the first high-skilled worker appears in that industry (t_j^l) and the time point when the last high-skilled worker leaves that industry (t_{j+2}^l). [[[WE SHOULD CHARACTERIZE THE LIFE SPAN OF EACH INDUSTRY AS THOROUGHLY AS POSSIBLE AND COMPARE IT WITH THE INDUSTRY LIFE SPAN IN JLW 2015. THESE RESULTS SHOULD BE SUMMARIZED IN A PROPOSITION]]] Figure 8 shows the dynamics of sectoral employments and unemployments for each industry. The employment of an industry reaches its peak when all workers, employed or unemployed, are in that industry.

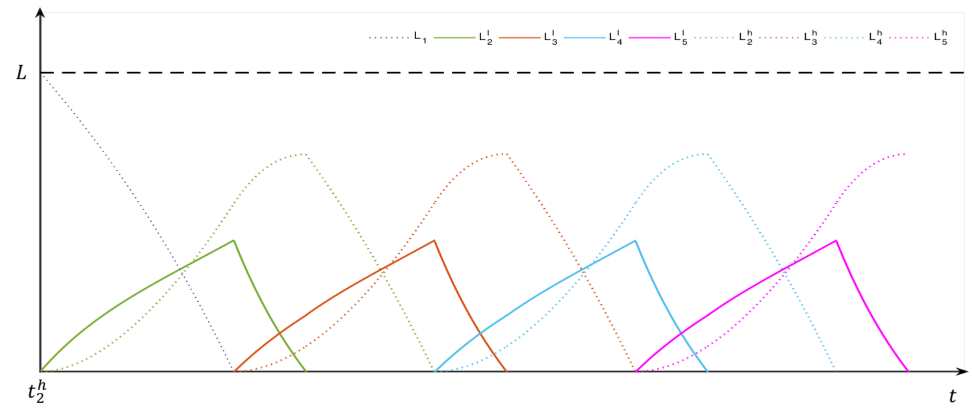

Fig. 7 *low-skilled and high-skilled workers in each industry*

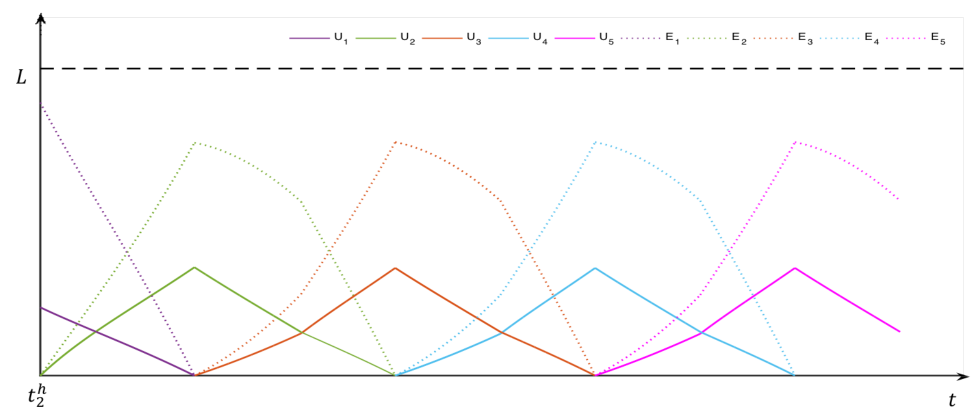

Fig. 8 *employed and unemployed workers in each industry*

The time path of the aggregate unemployment is given by

$$U(t) = \begin{cases} (\frac{\widehat{\delta}}{\pi\widehat{f}+\widehat{\delta}} - \frac{\widehat{\delta}}{\widehat{f}+\widehat{\delta}})e^{-\frac{\pi\widehat{f}}{\pi\widehat{f}+\widehat{\delta}}\xi(t-t^l_j)}\bar{L}^l_{j-1} + \frac{\widehat{\delta}}{\widehat{f}+\widehat{\delta}}L & \text{if} \quad t^l_j \leq t < t^h_j \\ (\frac{\widehat{\delta}}{\pi\widehat{f}} - \frac{\pi\widehat{f}+\widehat{\delta}}{\pi\widehat{f}}\frac{\widehat{\delta}}{\widehat{f}+\widehat{\delta}})[\frac{Q_{j-1}g_c}{N_2\xi+g_c}e^{g_c(t-t^h_j)} - (N_1 + N_2\bar{\bar{L}}^h_j - \frac{Q_{j-1}\xi N_2}{g_c+\xi N_2})e^{-N_2\xi(t-t^h_j)}] + \frac{\widehat{\delta}}{\widehat{f}+\widehat{\delta}}L & \text{if} \quad t^h_j \leq t < t^l_{j+1} \end{cases} . \tag{43}$$

Obviously, the aggregate unemployment decreases during period $[t^l_j, t^h_j)$. The intuition is that when low-skilled workers in industry $j-1$ move into industry j, they are always low-skilled type and suffer from skill mismatch. The aggregate unemployment declines because low-skilled workers in both industries gradually become high-skilled ones via on-job learning from t^l_j to t^h_j. It rises from t^h_j to t^l_{j+1} because high-skilled workers in industry $j-1$ begin to move into industry j and they become low-skilled and hence suffer from mismatch. The unemployment rate reaches the peak at t^l_{j+1}, the time point when all workers are all industry j. After that, the aggregate unmeployment declines again during the period $[t^l_{j+1}, t^h_{j+1})$, so on and so forth. As a result, the aggregate unemployment rate exhibits a cyclical pattern along with industrial upgrading on the aggregate growth path. The cycle of unemployment rate is not caused by any exogenous aggregate shocks, but rather due to the cyclicality in the number of low-skilled workers, which results from two processes: the labor reallocation process during which the number of low skilled workers increases as workers move from sunset industries to sunrise industries and the on-job learning process during which the number of low-skilled workers decreases. Figure 9 plots this cyclical pattern of aggregate unemployment rate.

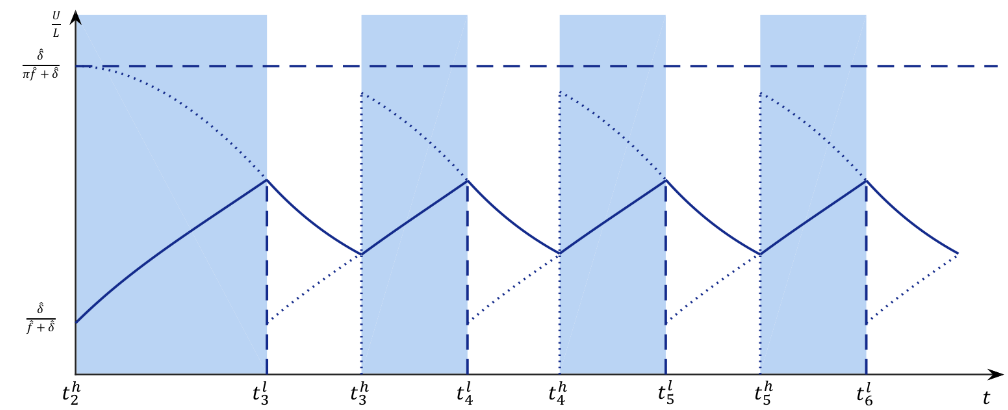

Fig. 9 *cyclical hump-shaped pattern of aggregate unemployment rate*

4 Conclusion

In this paper, we develop a highly tractable dynamic model with infinite industries to explore how frictional labor market affects industry dynamics and aggregate economic growth and how labor market performs in the context of industry dynamics. We are able to obtain a closed-form solution to fully characterize the aggregate growth, life-cycle dynamics of each of the infinite industries, as well as equilibrium unemployment rates. We show that in equilibrium the aggregate unemployment rate exhibits a cyclical pattern as the economy repeatedly kundertakes structural changes driven by endogenous capital accumulation. The aggregate unemployment rate exhibits a hump-shaped pattern: it rises at the beginning when high-skilled workers in a sunset industry move into a sunrise industry and suffers from skill mismatch as low-skilled workers in the new industry. The unemploymen rate declines later on when low-skilled workers become the high-skilled workers through on-the-job learning in the new industry. The unemployment rate goes up again as the current new industry gradually declines and is replaced by an even more capital-intensive industry, ad infinitum. We also find that there exist three critical forces: investment-specific technological progress, on-job learning efficiency and skill mismatch. When the rate of investment-specific technological progress is larger, the consumption growth rate is higher, the speed of old industries being replaced by new ones is higher, industry life span is shorter, and the aggregate unemployment rate is universally higher. When learning efficiency increases, the aggregate unemployment rate shifts downward and the life span of an industry becomes shorter. When mismatch gets more severe, the aggregate unemployment rate shifts upward and the life span of an industry becomes longer.

Appendix

A1. Proof of proposition 1:

Social Planner Problem: We start by setting the Hamiltonian for the social planner problem

$$H = e^{-\rho t}\frac{C^{1-\sigma}-1}{1-\sigma} + \mu_k[AK - E(C, E_0^l + E_0^h, E_1^l + E_1^h, E_2^l + E_2^h)] + \sum_{i=l,h}\mu_0^i[fU_i(1-\lambda_1^i-\lambda_2^i) - \delta E_0^i]$$
$$+ \sum_{i=l,h}\mu_1^i(fU_i\lambda_1^i - \delta E_1^i) + \mu_2^l(\pi f U_l \lambda_2^l - \delta E_2^l - \xi E_2^l) + \mu_2^h(fU_h\lambda_2^h - \delta E_2^h + \xi E_2^l)$$
$$+ \mu_u^l[-fU_l(1-\lambda_2^l) - \pi f U_l \lambda_2^l + \delta(E_0^l + E_1^l + E_2^l)] + \mu_u^h[-fU_h + \delta(E_0^h + E_1^h + E_2^h)] \quad (44)$$

where μ_k is the multiplier associated with capital constraint, μ_j^i is the multiplier associated with i skilled employees in industry j, and μ_u^i is the multiplier associated with unemployment of i skilled workers. The first-order conditions are

$$\dot{\mu}_k = -A\mu_k \quad (45)$$

$$-\dot{\mu}_0^i = -\delta\mu_0^i + \delta\mu_u^i - \mu_k\frac{\partial E}{\partial E_0^i}, i \in \{l \ , \ h\} \quad (46)$$

$$-\dot{\mu}_1^i = -\delta\mu_1^i + \delta\mu_u^i - \mu_k\frac{\partial E}{\partial E_1^i}, i \in \{l \ , \ h\} \quad (47)$$

$$-\dot{\mu}_2^l = -(\delta+\xi)\mu_2^l + \delta\mu_u^l + \xi\mu_2^h - \mu_k\frac{\partial E}{\partial E_2^l}, \quad (48)$$

$$-\dot{\mu}_2^h = -\delta\mu_2^h + \delta\mu_u^h - \mu_k\frac{\partial E}{\partial E_2^h}, \quad (49)$$

$$-\dot{\mu}_u^l = f(1-\lambda_1^l-\lambda_2^l)(\mu_0^l-\mu_u^l) + f\lambda_1^l(\mu_1^l-\mu_u^l) + \pi f\lambda_2^l(\mu_2^l-\mu_u^l), \quad (50)$$

$$-\dot{\mu}_u^h = f(1-\lambda_1^h-\lambda_2^h)(\mu_0^h-\mu_u^h) + f\lambda_1^h(\mu_1^h-\mu_u^h) + f\lambda_2^h(\mu_2^h-\mu_u^h), \quad (51)$$

$$e^{-\rho t}C^{-\sigma} = \mu_k\frac{\partial E}{\partial C}, \quad (52)$$

$$\frac{\partial H}{\partial \lambda_1^i} = fU_i(\mu_1^i - \mu_0^i), i \in \{l \ , \ h\} \quad (53)$$

$$\frac{\partial H}{\partial \lambda_2^l} = fU_l[\pi(\mu_2^l - \mu_u^l) - (\mu_0^l - \mu_u^l)], \quad (54)$$

$$\frac{\partial H}{\partial \lambda_2^h} = fU_h(\mu_2^h - \mu_0^h). \quad (55)$$

Decentralized Equilibrium: We then characterize the decentralized economy. The representative household has a continuum of members. Some of them are employed, and they are employed in different industries. Some of them are unemployed. The household can invest their capital $K(t)$ to receive $r(t)K(t)$ capital goods. Let $q(t)$ denote the price of consumption goods denominated by the capital goods.

Households can sell one consumption goods to receive $q(t)$ capital goods. Employed workers receive wage payment $q(t)w_j^i(t)$ in terms of capital goods corresponding to their skill type and the industry they works in. The household's problem is

$$\max \int_{t=0}^{\infty} e^{-\rho t} \frac{C^{1-\sigma}-1}{1-\sigma} dt \tag{56}$$

subject to his budget constraint

$$\dot{K} = rK + q(\sum_{i\in\{l,h\}, j\in\{0,1,2\}} w_j^i E_j^i - C). \tag{57}$$

where E_j^i is again the measure of employed i-skilled workers in industry j. Given the prizes for the equilibrium $\{r(t), q(t), w_j^i(t)\}$, the household decides the consumption $C(t)$ and the fraction of unemployed workers searching in different industries λ_j^i. Evolutions of employed and unemployed workers are the same as (7) – (14).

Firms in the capital goods sector borrow capital at the rental price $r(t)$, and they produce capital goods by using the technology of linear production function $AK(t)$. Firms producing final consumption goods use intermediate goods as input, and the production function is $X = \sum_{j=0}^{3} x_j$. Let $q_j(t)$ denote the price of intermediate goods j. The profit of firms producing final consumption goods is

$$q(t)\sum_{j=0}^{3} x_j - \sum_{j=0}^{3} q_j(t)x_j.$$

Firms in industry j producing intermediate goods uses constant return to scale technology $F_j(k,l)$. They post vacancies at zero cost. When they are matched with workers, they buy capital from the households and pay wage $q_j(t)w_j^i(t)$ to the employed workers. Their profit function $\pi_j(k,l)$ is

$$\pi_j(k,l) = q_j(t)F_j(k,l) - k - q(t)w_j^i(t)l.$$

For the market prizes $q = q_j = \frac{\partial E}{\partial C}, \forall j \in \{0,1,2,3\}; w_j^i = -\frac{\partial E}{\partial E_j^i} / \frac{\partial E}{\partial C}, \forall i \in \{l,h\}, \forall j \in \{0,1,2\}; r = A$, we check that we have an equilibrium. For firms in capital goods sector, $A = r$ indicates zero profit. For firms producing final consumption goods, $q = q_j$ means zero profit. For firms producing intermediate goods j, combining with $E(.)$ and expressions of q and w_j^i, we have the following inequality

$$q \le qw_0^i; q\lambda \le qw_1^i + a; q\lambda^2 \le qw_2^i + a^2; q\lambda^3 \le a^3, i \in \{l,h\}. \tag{58}$$

with equality if the corresponding production of intermediate goods j is positive. It indicates that firms in each industry j earn at most zero profit. By setting the Hamiltonian for the household's problem, we derive the FOCs. And we find the FOCs are the same as those for the social planner's problem after substituting the expressions for r, q and w_j^i. Thus the social planner's solution maximizes the household's utility given our market prizes. Besides, the social planner's solution is also consistent with the optimization problem of firms as they earn zero profit. Market clearing condition is satisfied as production functions are constant return to scale and firms can flexibly adjust their demand with zero profit.

A2. Proof of proposition 2

We first demonstrate asymptotic properties of first order conditions for the social planner in **Lemma 1**.

Lemma 1 *(1)* $\mu_j^i \to \mu_u^i$, $\dot{\mu}_j^i \to \dot{\mu}_u^i$, $i \in \{l, h\}$, $j \in \{0, 1, 2\}$; *(2)* $\dot{\mu}_u^l \to \frac{\mu_k \widehat{f}(\frac{\partial E}{\partial E_0^l}(1-\lambda_1^l-\lambda_2^l)+\frac{\partial E}{\partial E_1^l}\lambda_1^l+\pi\frac{\partial E}{\partial E_2^l}\lambda_2^l)-\pi\widehat{f}\lambda_2^l\xi(\mu_u^h-\mu_u^l)}{\widehat{f}[(1-\lambda_2^l)+\pi\lambda_2^l]+\widehat{\delta}}$, $\dot{\mu}_u^h \to \frac{\mu_k \widehat{f}(\frac{\partial E}{\partial E_0^h}(1-\lambda_1^h-\lambda_2^h)+\frac{\partial E}{\partial E_1^h}\lambda_1^h+\frac{\partial E}{\partial E_2^h}\lambda_2^h)}{\widehat{f}+\widehat{\delta}}$.

Proof. (1)Combining with $\delta = \kappa\widehat{\delta}$, $\kappa \to \infty$, we can transform the first order conditions into

$$\lim_{\kappa\to\infty} -\frac{\dot{\mu}_0^i}{\kappa} = \widehat{\delta}(\mu_u^i - \mu_0^i) - \lim_{\kappa\to\infty} \frac{\mu_k \frac{\partial E}{\partial E_0^i}}{\kappa}, i \in \{l \ , \ h\} \tag{59}$$

$$\lim_{\kappa\to\infty} -\frac{\dot{\mu}_1^i}{\kappa} = \widehat{\delta}(\mu_u^i - \mu_1^i) - \lim_{\kappa\to\infty} \frac{\mu_k \frac{\partial E}{\partial E_1^i}}{\kappa}, i \in \{l \ , \ h\} \tag{60}$$

$$\lim_{\kappa\to\infty} -\frac{\dot{\mu}_2^l}{\kappa} = \widehat{\delta}(\mu_u^l - \mu_2^l) - \lim_{\kappa\to\infty} \frac{\mu_k \frac{\partial E}{\partial E_2^l}}{\kappa} + \lim_{k\to\infty} \frac{\xi(\mu_2^h - \mu_2^l)}{\kappa}, \tag{61}$$

$$\lim_{\kappa\to\infty} -\frac{\dot{\mu}_2^h}{\kappa} = \widehat{\delta}(\mu_u^h - \mu_2^h) - \lim_{\kappa\to\infty} \frac{\mu_k \frac{\partial E}{\partial E_2^h}}{\kappa}. \tag{62}$$

We have $\mu_j^i \to \mu_u^i$. Similarly, by differentiating the first order conditions and taking the limit, we have $\dot{\mu}_j^i \to \dot{\mu}_u^i$. The intuition is straightforward. When the job dropping rate is sufficiently large, the employed workers quickly turn into unemployed workers, so their current value and future dynamics will converge to the same as unemployed ones when δ goes to infinite. Notice that $\mu_j^i \to \mu_u^i$ never implies $\delta(\mu_j^i - \mu_u^i) \to 0$ since $\delta \to \infty$. (2) Transforming the conditions for $\dot{\mu}_u^i$, we have

$$-\dot{\mu}_u^l = \frac{\widehat{f}}{\widehat{\delta}}[(1-\lambda_1^l-\lambda_2^l)\delta(\mu_0^l-\mu_u^l) + \lambda_1^l\delta(\mu_1^l-\mu_u^l) + \pi\lambda_2^l\delta(\mu_2^l-\mu_u^l)], \tag{63}$$

$$-\dot{\mu}_u^h = \frac{\widehat{f}}{\widehat{\delta}}[(1-\lambda_1^h-\lambda_2^h)\delta(\mu_0^h-\mu_u^h) + \lambda_1^h\delta(\mu_1^h-\mu_u^h) + \lambda_2^h\delta(\mu_2^h-\mu_u^h)]. \tag{64}$$

Substituting the conditions for $\dot{\mu}_j^i$ into the above transformed expressions to replace $\delta(\mu_j^i - \mu_u^i)$ and applying the property $\dot{\mu}_j^i \to \dot{\mu}_u^i$, $\mu_j^i \to \mu_u^i$, we demonstrate the second asymptotic property. ■

The next result gives a simple rule for the social planner to optimally allocate unemployed workers.

Lemma 2 *If $\xi < \bar{\xi}$, social planner compares the surplus $-\frac{\widehat{f}}{\widehat{f}+\widehat{\delta}}\frac{\partial E}{\partial E_j^h}$ for high skilled workers in industry $j = 0, 1, 2$, and the surplus $-\frac{\widehat{f}}{\widehat{f}+\widehat{\delta}}\frac{\partial E}{\partial E_j^l}$ if $j \neq 2$ and $-\frac{\pi\widehat{f}}{\pi\widehat{f}+\widehat{\delta}}\frac{\partial E}{\partial E_j^l}$ if $j = 2$ for low skilled workers to allocate workers in industries with maximal values at each instant.*

Proof. It provides a simple equivalence for the necessary conditions of λ_j^i,$i \in \{l, h\}, j \in \{1, 2\}$. The Kuhn-Tucker conditions require that $\lambda_j^i > 0$ only if $\frac{\partial H}{\partial \lambda_j^i} \geq 0$. When $\frac{\partial H}{\partial \lambda_1^i}$ and $\frac{\partial H}{\partial \lambda_2^i}$ are both non-negative for any i skilled workers, the constraint $\lambda_1^i + \lambda_2^i \leq 1$ requires that the weight should be put on the one with maximal value. In other words, we have $\lambda_j^i > 0$ only if $\frac{\partial H}{\partial \lambda_j^i}$ is the maximum of $\{0, \frac{\partial H}{\partial \lambda_1^i}, \frac{\partial H}{\partial \lambda_2^i}\}$. For the instantaneous equilibrium, we have

$$\begin{aligned}\frac{\partial H}{\partial \lambda_1^i} &= fU_i(\mu_1^i - \mu_0^i) = \frac{\widehat{f}}{\widehat{\delta}}U_i[\delta(\mu_1^i - \mu_u^i) - \delta(\mu_0^i - \mu_u^i)] = \frac{\widehat{f}}{\widehat{\delta}}U_i[(\dot{\mu}_1^i - \mu_k\frac{\partial E}{\partial E_1^i}) - (\dot{\mu}_0^i - \mu_k\frac{\partial E}{\partial E_0^i})] \\ &= \frac{\widehat{f}}{\widehat{\delta}}U_i[(\dot{\mu}_1^i - \dot{\mu}_0^i) + (-\mu_k\frac{\partial E}{\partial E_1^i} + \mu_k\frac{\partial E}{\partial E_0^i})] \to \frac{\widehat{f}}{\widehat{\delta}}U_i(-\mu_k\frac{\partial E}{\partial E_1^i} + \mu_k\frac{\partial E}{\partial E_0^i}), i \in \{l, h\}\end{aligned} \tag{65}$$

$$\frac{\partial H}{\partial \lambda_2^h} = fU_h(\mu_2^h - \mu_0^h) \to \frac{\widehat{f}}{\widehat{\delta}}U_h(-\mu_k\frac{\partial E}{\partial E_2^h} + \mu_k\frac{\partial E}{\partial E_0^h}), \tag{66}$$

$$\begin{aligned}\frac{\partial H}{\partial \lambda_2^l} &= fU_l[\pi(\mu_2^l - \mu_u^l) - (\mu_0^l - \mu_u^l)] = \frac{\widehat{f}}{\widehat{\delta}}U_l[\pi(\dot{\mu}_2^l + \xi(\mu_2^h - \mu_2^l) - \mu_k\frac{\partial E}{\partial E_2^l}) - (\dot{\mu}_0^l - \mu_k\frac{\partial E}{\partial E_1^l})] \\ &\to \frac{\widehat{f}}{\widehat{\delta}}U_l[(\pi - 1)\dot{\mu}_u^l + \pi\xi(\mu_u^h - \mu_u^l) - \pi\mu_k\frac{\partial E}{\partial E_2^l} + \mu_k\frac{\partial E}{\partial E_0^l})] \\ &\to \frac{(1-\pi)\lambda_1^l\widehat{f}(\frac{\partial E}{\partial E_0^l} - \frac{\partial E}{\partial E_1^l}) + \frac{(\pi\widehat{f}+\widehat{\delta})(\widehat{f}+\widehat{\delta})}{\widehat{f}}(\frac{\widehat{f}}{\widehat{f}+\widehat{\delta}}\frac{\partial E}{\partial E_0^l} - \frac{\pi\widehat{f}}{\pi\widehat{f}+\widehat{\delta}}\frac{\partial E}{\partial E_2^l})}{\widehat{f}[(1-\lambda_2^l) + \pi\lambda_2^l] + \widehat{\delta}} + \frac{\pi\xi(\widehat{f}+\widehat{\delta})(\mu_2^h - \mu_2^l)}{\widehat{f}[(1-\lambda_2^l) + \pi\lambda_2^l] + \widehat{\delta}},\end{aligned} \tag{67}$$

By symmetry,

$$\frac{\partial H}{\partial \lambda_2^h} - \frac{\partial H}{\partial \lambda_1^h} = fU_h(\mu_2^h - \mu_1^h) \to \frac{\widehat{f}}{\widehat{\delta}}U_h(-\mu_k\frac{\partial E}{\partial E_2^h} + \mu_k\frac{\partial E}{\partial E_1^h}), \tag{68}$$

$$\begin{aligned}&\frac{\partial H}{\partial \lambda_2^l} - \frac{\partial H}{\partial \lambda_1^l} = fU_l[\pi(\mu_2^l - \mu_u^l) - (\mu_1^l - \mu_u^l)] \\ \to &\frac{(1-\pi)(1-\lambda_1^l-\lambda_2^l)\widehat{f}(\frac{\partial E}{\partial E_1^l} - \frac{\partial E}{\partial E_0^l}) + \frac{(\pi\widehat{f}+\widehat{\delta})(\widehat{f}+\widehat{\delta})}{\widehat{f}}(\frac{\widehat{f}}{\widehat{f}+\widehat{\delta}}\frac{\partial E}{\partial E_1^l} - \frac{\pi\widehat{f}}{\pi\widehat{f}+\widehat{\delta}}\frac{\partial E}{\partial E_2^l})}{\widehat{f}[(1-\lambda_2^l) + \pi\lambda_2^l] + \widehat{\delta}} + \frac{\pi\xi(\widehat{f}+\widehat{\delta})(\mu_2^h - \mu_2^l)}{\widehat{f}[(1-\lambda_2^l) + \pi\lambda_2^l] + \widehat{\delta}}.\end{aligned} \tag{69}$$

For the high-skilled workers, from (65)(66)(68) we know that $\lambda_j^h > 0$ only if $-\frac{\widehat{f}}{\widehat{f}+\widehat{\delta}}\frac{\partial E}{\partial E_j^h}$ is the maximum of $\{-\frac{\widehat{f}}{\widehat{f}+\widehat{\delta}}\frac{\partial E}{\partial E_0^h}, -\frac{\widehat{f}}{\widehat{f}+\widehat{\delta}}\frac{\partial E}{\partial E_1^h}, -\frac{\widehat{f}}{\widehat{f}+\widehat{\delta}}\frac{\partial E}{\partial E_2^h}\}$. We use the multiplier $-\frac{\widehat{f}}{\widehat{f}+\widehat{\delta}}$ because we want to first highlight the similarity of conditions for high-skilled workers and for low-skilled workers. Moreover, it has an economic meaning: since $-\frac{\partial E}{\partial E_j^h}$ indicates reduced expenditure of capital flows into consumption goods when one more unit of high-skilled worker is employed in industry j, $-\frac{\widehat{f}}{\widehat{f}+\widehat{\delta}}\frac{\partial E}{\partial E_j^h}$ indicates the reduced expenditure of capital flows if one more unit of high-skilled workers are allocated in industry j in the steady state. For low-skilled workers, equation (65) implies that $\frac{\partial H}{\partial \lambda_1^l} > 0$ if and only if $-\frac{\widehat{f}}{\widehat{f}+\widehat{\delta}}\frac{\partial E}{\partial E_1^l} > -\frac{\widehat{f}}{\widehat{f}+\widehat{\delta}}\frac{\partial E}{\partial E_0^l}$. In other words, social planner has larger incentive to allocate low skilled workers in industry 1 rather than industry 0 only if the reduced expenditure of capital flows is larger when low-skilled workers are allocated industry 1 rather than industry 2 in the steady state. If $\frac{\partial H}{\partial \lambda_1^l} > 0$, social planner compares $\frac{\partial H}{\partial \lambda_1^l}$ and $\frac{\partial H}{\partial \lambda_2^l}$ to determine whether $\lambda_1^l > 0$ or $\lambda_2^l > 0$. We must have $1 - \lambda_1^l - \lambda_2^l = 0$, and (69) implies that it is equivalent to compare $-\frac{\pi\widehat{f}}{\pi\widehat{f}+\widehat{\delta}}\frac{\partial E}{\partial E_2^l}$ and $-\frac{\widehat{f}}{\widehat{f}+\widehat{\delta}}\frac{\partial E}{\partial E_1^l}$ if ξ is not large enough so that the first term of (69) is determinant. The intuition is that social planner only focuses on the instant reduced expenditure, if the future revenue of skill premium is not large enough. When $\xi \to 0$, $\frac{\pi\xi(\widehat{f}+\widehat{\delta})(\mu_2^h-\mu_2^l)}{\widehat{f}[(1-\lambda_2^l)+\pi\lambda_2^l]+\widehat{\delta}}$ uniformly converges to 0. It indicates that there must be a required positive threshold $\bar{\xi_1}$, and when $\xi < \bar{\xi_1}$ the first term is dominant. Consider another case when $\frac{\partial H}{\partial \lambda_1^l} < 0$. We have $\lambda_1^l = 0$ and low-skilled workers are allocated either in industry 0 or industry 2. According to (77), it is equivalent to compare $-\frac{\widehat{f}}{\widehat{f}+\widehat{\delta}}\frac{\partial E}{\partial E_0^l}$ and $-\frac{\pi\widehat{f}}{\pi\widehat{f}+\widehat{\delta}}\frac{\partial E}{\partial E_2^l}$ when ξ is below the threshold $\bar{\xi_2}$. If $-\frac{\pi\widehat{f}}{\pi\widehat{f}+\widehat{\delta}}\frac{\partial E}{\partial E_2^l} > -\frac{\widehat{f}}{\widehat{f}+\widehat{\delta}}\frac{\partial E}{\partial E_0^l}$, we have $\frac{\partial H}{\partial \lambda_2^l} > 0$ and the low-skilled workers are only allocated in industry 2 and vice versa. Q.E.D. ■

The rest of proof is straightforward. Taking the limit $\kappa \to \infty$, the behavior of labor market converges to

$$\widehat{f}\lambda_j^h U_h = \widehat{\delta}E_j^h, j \in \{0,1,2\}; \quad \widehat{f}\lambda_j^l U_h = \widehat{\delta}E_j^l, j \in \{0,1\}; \quad \pi\widehat{f}\lambda_2^l U_h = \widehat{\delta}E_2^l, \tag{70}$$

Combining with the definition of $L_j^i(t)$, we have

$$E_j^h = \frac{\widehat{f}}{\widehat{f}+\widehat{\delta}}L_j^h, j \in \{0,1,2\}; \quad E_j^l = \frac{\widehat{f}}{\widehat{f}+\widehat{\delta}}L_j^l, j \in \{0,1\}; \quad E_2^l = \frac{\pi\widehat{f}}{\pi\widehat{f}+\widehat{\delta}}L_2^l. \tag{71}$$

The labor endowment $\{L_l, L_h\}$ is given at each instant, and the social planner is confronted with the constraints that $\sum_{j=0}^2 L_j^i \leq L_i, i \in \{l,h\}$. Combining with ***Lemma 2***, social planner optimally allocates workers to minimize the expenditure of capital flows into the consumption goods sector. Equivalently, it requires an optimal steady state equilibrium to produce $C(t)$ at each instant. Q.E.D.

A3. Proof of proposition 3

Since all resources are used for the optimal solution, we set the Langrange as follows

$$\begin{aligned} L = (E_0^l + E_0^h) + \lambda(E_1^l + E_1^h) + \lambda^2(E_2^l + E_2^h) + \frac{\lambda^3}{a^3}K_3 + \mu_E(E - a(E_1^l + E_1^h) - a^2(E_2^l + E_2^h) - K_3) \\ + \mu_l(L^l - E_0^l\frac{\widehat{f}+\widehat{\delta}}{\widehat{f}} - E_1^l\frac{\widehat{f}+\widehat{\delta}}{\widehat{\delta}} - E_2^l\frac{\pi\widehat{f}+\widehat{\delta}}{\pi\widehat{f}}) + \mu_h(L - (E_0^h + E_1^h + E_2^h)\frac{\widehat{f}+\widehat{\delta}}{\widehat{f}}), \end{aligned} \tag{72}$$

The control variables are E_j^i,$i \in \{l,h\}$,$j \in \{0,1,2\}$ which are used to produce goods $j = 0,1,2$ and K_3 which is used to produce goods 3. The Kuhn-Tucker conditions are

$$(1 - \mu_i \frac{\widehat{f} + \widehat{\delta}}{\widehat{f}})E_0^i = 0, E_0^i \geq 0, 1 - \mu_i \frac{\widehat{f} + \widehat{\delta}}{\widehat{f}} \leq 0, i \in \{l,h\}, \tag{73}$$

$$(\lambda - \mu_i \frac{\widehat{f} + \widehat{\delta}}{\widehat{f}} - a\mu_E)E_1^i = 0, E_1^i \geq 0, \lambda - \mu_i \frac{\widehat{f} + \widehat{\delta}}{\widehat{f}} - a\mu_E \leq 0, i \in \{l,h\}, \tag{74}$$

$$(\lambda^2 - \mu_l \frac{\pi\widehat{f} + \widehat{\delta}}{\pi\widehat{f}} - a^2\mu_E)E_2^l = 0, E_2^l \geq 0, \lambda^2 - \mu_l \frac{\pi\widehat{f} + \widehat{\delta}}{\pi\widehat{f}} - a^2\mu_E \leq 0, \tag{75}$$

$$(\lambda^2 - \mu_h \frac{\widehat{f} + \widehat{\delta}}{\widehat{f}} - a^2\mu_E)E_2^h = 0, E_2^h \geq 0, \lambda^2 - \mu_h \frac{\widehat{f} + \widehat{\delta}}{\widehat{f}} - a^2\mu_E \leq 0, \tag{76}$$

$$(\frac{\lambda^3}{a^3} - \mu_E)K_3 = 0, K_3 \geq 0, \frac{\lambda^3}{a^3} - \mu_E \leq 0. \tag{77}$$

For each condition, we uniformly shift $\mu_i \frac{\widehat{f}+\widehat{\delta}}{\widehat{f}}$ in the last inequality term to the right-hand side and compare the left-hand side value. Since μ_i is the shadow price of i skilled workers and indicates their marginal contribution to the consumption goods, $\mu_i \frac{\widehat{f}+\widehat{\delta}}{\widehat{f}}$ refers to the marginal contribution of any employed worker who does not suffer from mismatch. For $i = l$, the left-hand side values are 1, $\lambda - a\mu_E$, and $\frac{\pi\widehat{f}+\pi\widehat{\delta}}{\pi\widehat{f}+\widehat{\delta}}(\lambda^2 - a^2\mu_E)$ in the condition of $E_j^l = 0,1,2$ respectively. For $i = h$ the values are 1, $\lambda - a\mu_E$, and $\lambda^2 - a^2\mu_E$ in the condition of $E_j^h = 0,1,2$ respectively. For industry 3, the last term is $\frac{\lambda^3}{a^3} - \mu_E$. To determine μ_i and E_j^i we need to find the maximal left-hand side values. Notice that μ_E is the only endogenous variables in these expressions, which is actually the shadow price of capital flows. Assume $a - 1 > \lambda > \frac{\pi\widehat{f}+\widehat{\delta}}{\pi\widehat{f}+\pi\widehat{\delta}}$, and $\frac{a^2}{a\frac{\pi\widehat{f}+\pi\widehat{\delta}}{\pi\widehat{f}+\widehat{\delta}}-1} > \frac{\lambda^2}{\lambda\frac{\pi\widehat{f}+\pi\widehat{\delta}}{\pi\widehat{f}+\widehat{\delta}}-1}$. When μ_k goes down, workers gradually move from the most labor intensive industry 0) to the industry (2). Mismatch delays the industrial upgrading for the low skilled workers.

A4. Time path for the number of employed low-skilled workers in industry 2

From Table 1, we know that

$$E_2^l = \frac{\pi\widehat{f}}{\pi\widehat{f} + \widehat{\delta}}(\frac{E - a^2\frac{\widehat{f}}{\widehat{f}+\widehat{\delta}}(L - L^l) - a\frac{\widehat{f}}{\widehat{f}+\widehat{\delta}}L^l}{a^2\frac{\pi\widehat{f}}{\pi\widehat{f}+\widehat{\delta}} - a\frac{\widehat{f}}{\widehat{f}+\widehat{\delta}}}), \tag{78}$$

$$E = \frac{a^2\frac{\pi\widehat{f}}{\pi\widehat{f}+\widehat{\delta}} - a\frac{\widehat{f}}{\widehat{f}+\widehat{\delta}}}{\lambda^2\frac{\pi\widehat{f}}{\pi\widehat{f}+\widehat{\delta}} - \lambda\frac{\widehat{f}}{\widehat{f}+\widehat{\delta}}}C - \frac{\widehat{f}}{\widehat{f}+\widehat{\delta}}\frac{a(a-\lambda)}{\lambda\frac{\pi\widehat{f}}{\pi\widehat{f}+\widehat{\delta}} - \frac{\widehat{f}}{\widehat{f}+\widehat{\delta}}}(\frac{\pi\widehat{f}}{\pi\widehat{f}+\widehat{\delta}}L^l + \frac{\widehat{f}}{\widehat{f}+\widehat{\delta}}(L - L^l)), \tag{79}$$

and (??) implies

$$C = \lambda L\frac{\widehat{f}}{\widehat{f}+\widehat{\delta}}e^{g_c(t-t_1)}, t > t_1. \tag{80}$$

From the above three equations, we obtain

$$E_2^l(t) = N_1 e^{g_c(t-t_1)} - (N_1 + N_2 L) + N_2 L^l, \tag{81}$$

where

$$N_1 = \frac{L\frac{\widehat{f}}{\widehat{f}+\widehat{\delta}}\frac{\pi\widehat{f}}{\pi\widehat{f}+\widehat{\delta}}}{\lambda\frac{\pi\widehat{f}}{\pi\widehat{f}+\widehat{\delta}} - \frac{\widehat{f}}{\widehat{f}+\widehat{\delta}}}, N_2 = \frac{(\lambda-1)\frac{\widehat{f}}{\widehat{f}+\widehat{\delta}}\frac{\pi\widehat{f}}{\pi\widehat{f}+\widehat{\delta}}}{\lambda\frac{\pi\widehat{f}}{\pi\widehat{f}+\widehat{\delta}} - \frac{\widehat{f}}{\widehat{f}+\widehat{\delta}}}. \tag{82}$$

As E_2^l cannot exceed the total supply of low-skilled workers, so we have

$$E_2^l(t) = \min\{N_1 e^{g_c(t-t_1)} - (N_1 + N_2 L) + N_2 L^l, \frac{\pi\widehat{f}}{\pi\widehat{f}+\widehat{\delta}} L^l\}. \tag{83}$$

A5. Proof of proposition 5

(1) (Learning efficiency) Consider the first half of dynamics when $t_1 < t \leq t_{2,l}$. Taking the partial derivative of E_2^l with respect to ξ, we get

$$\begin{aligned} \frac{\partial E_2^l}{\partial \xi} &= \frac{N_1 N_2 g_c}{(N_2+g_c)^2}[((N_2+g_c)(t-t_1)+1)e^{-N_2\xi(t-t_1)} - e^{g_c(t-t_1)}] \\ &= \frac{N_1 N_2 g_c e^{-N_2\xi(t-t_1)}}{(N_2+g_c)^2}[((N_2+g_c)(t-t_1)+1) - e^{(N_2\xi+g_c)(t-t_1)}] < 0 \end{aligned} \tag{84}$$

The last term is negative for $e^x > x+1$ if $x \neq 0$. This inequality will be frequently used in our following analysis. We have already proved that $E_2^l = N_1 e^{g_c(t-t_1)} - (N_1 + N_2 L) + N_2 L_l$ for the first half. Therefore we also have $\frac{\partial}{\partial \xi} < 0$ for any $t_1 < t \leq t_{2,l}$. $\frac{\partial}{\partial \xi} < 0$ and $\frac{\partial E_2^l}{\partial \xi} < 0$ infer that aggregate unemployment rate and upper bond unemployment rate shift downward with respect to a larger ξ in the first half as is shown in Figure 3. Then we prove the duration of industrial upgrading from industry 1 to industry 2 is extended. Equivalently, the turning time when aggregate unemployment rate reaches it peak value after low-skilled workers first move into industry 2 is delayed. To be more precise, $t'_{2,l} - t'_1 > t_{2,l} - t_1$ as is shown in Figure 3. The turning time is determined by

$$\frac{\pi\widehat{f}+\widehat{\delta}}{\pi\widehat{f}} E_2^l = L_l, \tag{85}$$

Combining with the expression of E_l^2 in terms of L_l and t, we have

$$\frac{\pi\widehat{f}+\widehat{\delta}}{\pi\widehat{f}}(N_1 e^{g_c(t-t_1)} - (N_1 + N_2 L)) = L_l(1 - \frac{\pi\widehat{f}+\widehat{\delta}}{\pi\widehat{f}} N_2), \tag{86}$$

It is simple to find $\frac{\pi\widehat{f}+\widehat{\delta}}{\pi\widehat{f}} N_2 < 1$ since $\pi < 1$. When ξ increases the right-hand side increases since $\frac{\partial L_l}{\partial \xi} > 0$. The left-hand side increases when $t - t_1$ increases. Therefore when ξ increases we must have $t'_{2,l} - t'_1 > t_{2,l} - t_1$. Finally we prove that when ξ increases, aggregate unemployment rate also decreases for the second half when $t > t_{2,l}$. Aggregate unemployment rate decreases smoothly after $t_{2,l}$ when the constraint $L_2^l \le L_l$ is binding. In Figure 3, the extended dotted line from $\frac{\widehat{\delta}}{\pi \widehat{f}+\widehat{\delta}}$ after $t_{2,l}$ is actually above the solid line. For convenience, we define this hypothetical unemployment rate without considering $L_2^l \le L_l$ after $t_{2,l}$ as lower bond unemployment rate. You can see that upper bond unemployment rate turns into lower bond unemployment rate after $t_{2,l}$. We have already proved lower bond unemployment rate shifts downward when ξ increases. Since it is below aggregate unemployment, we must have $U'|_{t'_{2,l}} < U|_{t_1+(t'_{2,l}-t'_1)}$. And then aggregate unemployment decreases exponentially at a larger rate $-\frac{\pi \widehat{f}}{\pi \widehat{f}+\widehat{\delta}}\xi'$. In this sense, aggregate unemployment actually shifts downward for the whole path. (2) (capital-specific productivity) When A increases the growth rate of consumption goods g_c increases. It induces more rapid industrial upgrading. Taking the partial derivative of E_l^2 and L_l with respect to g_c when $t_1 < t \le t_{2,l}$, we have

$$\frac{\partial E_2^l}{\partial g_c} = \frac{N_1}{N_2\xi + g_c}[\frac{N_2\xi}{N_2\xi + g_c}(e^{g_c(t-t_1)} - e^{-N_2\xi(t-t_1)}) + g_c(t-t_1)e^{g_c(t-t_1)}] > 0 \tag{87}$$

$$\begin{aligned}\frac{\partial L_l}{\partial g_c} &= \frac{-N_1}{N_2(N_2\xi + g_c)}[\xi N_2(t-t_1)e^{g_c(t-t_1)} - \frac{N_2\xi}{N_2\xi + g_c}e^{g_c(t-t_1)} + \frac{N_2\xi}{N_2\xi + g_c}e^{-N_2\xi(t-t_1)}] \\ &= \frac{-\xi N_1 N_2 e^{g_c(t-t_1)}}{N_2(N_2\xi + g_c)^2}[(t-t_1)(N_2\xi + g_c) + (e^{-(N_2\xi+g_c)(t-t_1)} - 1)] < 0 \end{aligned} \tag{88}$$

The last term is negative using the inequality $e^x > x+1$ if $x \ne 0$. Therefore, aggregate unemployment rate shifts upward while upper bond unemployment rate shifts downward when $t_1 < t \le t_{2,l}$. Immediately, we have $t'_{2,l} - t'_1 < t_{2,l} - t_1$ indicating that it takes less time for structural unemployment rate to reach its peak value. Meanwhile structural unemployment is more severe during industrial upgrading. Then we prove it reaches a larger peak value when the industrial upgrading is more intensive. We make use of the properties that L_l is decreasing with t while E_2^l is increasing with t before turning point. Aggregate unemployment rate equals to upper bond unemployment rate when $L_2^l = L_l$ at time $t_{2,l}$. We compare $\frac{\frac{\partial L_l}{\partial g_c}}{\frac{\partial L_l}{\partial t}}|_{t_{2,l}}$ and $\frac{\frac{\partial L_2^l}{\partial g_c}}{\frac{\partial L_2^l}{\partial t}}|_{t_{2,l}}$, while the second term equals to $\frac{\frac{\partial E_2^l}{\partial g_c}}{\frac{\partial E_2^l}{\partial t}}|_{t_{2,l}}$ by using $L_2^l = \frac{\pi \widehat{f}+\widehat{\delta}}{\pi \widehat{f}}E_2^l$. $\frac{\frac{\partial L_l}{\partial g_c}}{\frac{\partial L_l}{\partial t}}|_{t_{2,l}}$ indicates how much time is put forward to reach the same value of L_l when g_c increases one unit. $\frac{\frac{\partial L_2^l}{\partial g_c}}{\frac{\partial L_2^l}{\partial t}}|_{t_{2,l}}$ indicates that how much time is put forward to reach the same value of L_2^l if g_c increases one unit. If $\frac{\frac{\partial L_l}{\partial g_c}}{\frac{\partial L_l}{\partial t}}|_{t_{2,l}} > \frac{\frac{\partial L_2^l}{\partial g_c}}{\frac{\partial L_2^l}{\partial t}}|_{t_{2,l}}$, L_l first comes to the initial value. Since L_l is decreasing while L_2^l is increasing, it induces a smaller peak value of aggregate unemployment rate when $L_l = L_2^l$. On the contrary, if $\frac{\frac{\partial L_2^l}{\partial g_c}}{\frac{\partial L_2^l}{\partial t}}|_{t_{2,l}} > \frac{\frac{\partial L_l}{\partial g_c}}{\frac{\partial L_l}{\partial t}}|_{t_{2,l}}$, L_2^l first increases to the initial value and induces a larger peak value of aggregate unemployment rate. Taking the partial derivative of

E_2^l, we have

$$\begin{aligned}\frac{\frac{\partial E_2^l}{\partial g_c}}{\frac{\partial E_2^l}{\partial t}}|_{t_{2,l}} &= \frac{\frac{N_1}{N_2\xi+g_c}[\frac{N_2\xi}{N_2\xi+g_c}(e^{g_c(t_{2,l}-t_1)} - e^{-N_2\xi(t_{2,l}-t_1)}) + g_c(t_{2,l}-t_1)e^{g_c(t_{2,l}-t_1)}]}{\frac{N_1 g_c}{N_2\xi+g_c}[g_c e^{g_c(t_{2,l}-t_1)} + N_2\xi e^{-N_2\xi(t_{2,l}-t_1)}]} \\ &= \frac{\frac{N_2\xi}{N_2\xi+g_c}(e^{g_c(t_{2,l}-t_1)} - e^{-N_2\xi(t_{2,l}-t_1)}) + g_c(t_{2,l}-t_1)e^{g_c(t_{2,l}-t_1)}}{g_c[g_c e^{g_c(t_{2,l}-t_1)} + N_2\xi e^{-N_2\xi(t_{2,l}-t_1)}]} \\ = \frac{\frac{1}{g_c}(e^{g_c(t_{2,l}-t_1)} - e^{-N_2\xi(t_{2,l}-t_1)}) - \frac{1}{N_2\xi+g_c}(e^{g_c(t_{2,l}-t_1)} - e^{-N_2\xi(t_{2,l}-t_1)}) + (t_{2,l}-t_1)e^{g_c(t_{2,l}-t_1)}}{g_c e^{g_c(t_{2,l}-t_1)} + N_2\xi e^{-N_2\xi(t_{2,l}-t_1)}}&,\end{aligned} \tag{89}$$

Similarly, taking the partial derivative of L_l

$$\begin{aligned}\frac{\frac{\partial L_l}{\partial g_c}}{\frac{\partial L_l}{\partial t}}|_{t_{2,l}} &= \frac{\frac{\xi N_1 N_2}{N_2(N_2\xi+g_c)^2}[((t_{2,l}-t_1)(N_2\xi+g_c)-1)e^{g_c(t_{2,l}-t_1)} + e^{-N_2\xi(t_{2,l}-t_1)}]}{\frac{\xi N_1 N_2 g_c}{N_2(N_2\xi+g_c)}[e^{g_c(t_{2,l}-t_1)} - e^{-N_2\xi(t_{2,l}-t_1)}]} \\ &= \frac{\frac{1}{N_2\xi+g_c}[((t_{2,l}-t_1)(N_2\xi+g_c)-1)e^{g_c(t_{2,l}-t_1)} + e^{-N_2\xi(t_{2,l}-t_1)}]}{g_c[e^{g_c(t_{2,l}-t_1)} - e^{-N_2\xi(t_{2,l}-t_1)}]} \\ &= \frac{-\frac{1}{N_2\xi+g_c}(e^{g_c(t_{2,l}-t_1)} - e^{-N_2\xi(t_{2,l}-t_1)}) + (t_{2,l}-t_1)e^{g_c(t_{2,l}-t_1)}}{g_c[e^{g_c(t_{2,l}-t_1)} - e^{-N_2\xi(t_{2,l}-t_1)}]},\end{aligned} \tag{90}$$

We demonstrate that $\frac{\frac{\partial E_2^l}{\partial g_c}}{\frac{\partial E_2^l}{\partial t}}|_{t_{2,l}} > \frac{\frac{\partial L_l}{\partial g_c}}{\frac{\partial L_l}{\partial t}}|_{t_{2,l}}$ by the following equivalence

$$\begin{aligned}&\frac{\frac{\partial E_2^l}{\partial g_c}}{\frac{\partial E_2^l}{\partial t}}|_{t_{2,l}} > \frac{\frac{\partial L_l}{\partial g_c}}{\frac{\partial L_l}{\partial t}}|_{t_{2,l}} \\ \Leftrightarrow\ &[e^{g_c(t_{2,l}-t_1)} - e^{-N_2\xi(t_{2,l}-t_1)}]^2 > [e^{-N_2\xi(t_{2,l}-t_1)} + ((N_2\xi+g_c)(t_{2,l}-t_1)-1)e^{g_c(t_{2,l}-t_1)}]e^{-N_2\xi(t_{2,l}-t_1)} \\ \Leftrightarrow\ &e^{g_c(t_{2,l}-t_1)} > [(N_2\xi+g_c)(t_{2,l}-t_1)+1]e^{-N_2\xi(t_{2,l}-t_1)} \\ \Leftrightarrow\ &e^{(N_2\xi+g_c)(t_{2,l}-t_1)} > [(N_2\xi+g_c)(t_{2,l}-t_1)+1]\end{aligned} \tag{91}$$

We finally prove that aggregate unemployment rate in the long run is lower with a larger g_c by contradiction. It is equivalent to show that aggregate unemployment rate is below the initial value after the duration $t_{2,l} - t_1$. If not, aggregate unemployment rate must be above the initial value in the transitional time period $(0, t_{2,l})$. It induces that E_2^l is larger and L_l decreases at a larger rate during this period. Since L_l comes to a lower value, aggregate unemployment must be below the initial value after the duration $t_{2,l} - t_1$ which leads to the contradiction. (3) (Mismatch) We prove that when π decreases (i.e. low-skilled workers suffer from greater mismatch), aggregate unemployment rate increases permanently after the industrial upgrading of low-skilled workers from industry 1 to industry 2. In this case, $\frac{\pi\widehat{f}}{\pi\widehat{f}+\widehat{\delta}}$ decreases which induces greater mismatch unemployment rate $\frac{\widehat{\delta}}{\pi\widehat{f}+\widehat{\delta}}$. Firstly, we prove that $\frac{\partial E_2^l}{\partial \frac{\pi f}{\pi \widehat{f}+\widehat{\delta}}}|_t < 0$ when $t_1 < t \leq t_{2,l}$. Since aggregate unemployment can be transformed into

$$U = \frac{\widehat{\delta}}{\widehat{f}+\widehat{\delta}}L + (\frac{\widehat{\delta}}{\pi\widehat{f}+\widehat{\delta}} - \frac{\widehat{\delta}}{\widehat{f}+\widehat{\delta}})L_2^l = \frac{\widehat{\delta}}{\widehat{f}+\widehat{\delta}}L + (\frac{\frac{\widehat{f}}{\widehat{f}+\widehat{\delta}}}{\frac{\pi\widehat{f}}{\pi\widehat{f}+\widehat{\delta}}} - 1)E_2^l, \tag{92}$$

we immediately get $\frac{\partial U}{\partial \frac{\pi \widehat{f}}{\pi \widehat{f}+\widehat{\delta}}}|_t < 0$. It is simple to check that $\frac{N_1}{N_2} = \frac{L}{\lambda-1}$ and $\frac{\partial N_2}{\partial \frac{\pi \widehat{f}}{\pi \widehat{f}+\widehat{\delta}}} < 0$. We have

$$\frac{\partial E_2^l}{\partial \frac{\pi \widehat{f}}{\pi \widehat{f}+\widehat{\delta}}} = \frac{N_1}{N_2} \frac{\partial \frac{N_2 g_c}{N_2\xi+g_c}(e^{g_c(t-t_1)} - e^{-N_2\xi(t-t_1)})}{\partial N_2} \frac{\partial N_2}{\partial \frac{\pi \widehat{f}}{\pi \widehat{f}+\widehat{\delta}}}$$

$$= \frac{N_1}{N_2} \frac{\partial N_2}{\partial \frac{\pi \widehat{f}}{\pi \widehat{f}+\widehat{\delta}}} [(\frac{g_c}{N_2\xi+g_c})^2 (e^{g_c(t-t_1)} - e^{-N_2\xi(t-t_1)}) + \frac{N_2\xi g_c}{N_2\xi+g_c}(t-t_1)e^{-N_2\xi(t-t_1)}] < 0 \tag{93}$$

It indicates that when mismatch is more serious, aggregate unemployment rate shifts upward for the first half. To explain for the second half, we need to find out how the upper bond unemployment rate changes with respect to $\frac{\pi \widehat{f}}{\pi \widehat{f}+\widehat{\delta}}$. We prove that it shifts upward before time $t_{2,l}$. According to its definition, we can tranform the expression as follows

$$U^* \equiv \frac{\widehat{\delta}}{\widehat{f}+\widehat{\delta}} L + (\frac{\widehat{\delta}}{\pi \widehat{f}+\widehat{\delta}} - \frac{\widehat{\delta}}{\widehat{f}+\widehat{\delta}}) L_l = \frac{\widehat{\delta}}{\widehat{f}+\widehat{\delta}} L + (\frac{\widehat{f}}{\widehat{f}+\widehat{\delta}} - \frac{\pi \widehat{f}}{\pi \widehat{f}+\widehat{\delta}}) L_l, \tag{94}$$

Taking the partial derivative, we have

$$\frac{\partial U^*}{\partial \frac{\pi \widehat{f}}{\pi \widehat{f}+\widehat{\delta}}} = (\frac{\widehat{f}}{\widehat{f}+\widehat{\delta}} - \frac{\pi \widehat{f}}{\pi \widehat{f}+\widehat{\delta}}) \frac{\partial L_l}{\partial \frac{\pi \widehat{f}}{\pi \widehat{f}+\widehat{\delta}}} - L_l$$

$$= -\frac{N_1}{N_2} (\frac{\widehat{f}}{\widehat{f}+\widehat{\delta}} - \frac{\pi \widehat{f}}{\pi \widehat{f}+\widehat{\delta}}) \frac{\partial [\frac{\xi N_2}{(N_2\xi+g_c)} e^{g_c(t-t_1)} + \frac{g_c}{N_2\xi+g_c} e^{-N_2\xi(t-t_1)}]}{\partial N_2} \frac{\partial N_2}{\partial \frac{\pi \widehat{f}}{\pi \widehat{f}+\widehat{\delta}}} - L_l, \tag{95}$$

We derive the expression for each part

$$\frac{\partial N_2}{\partial \frac{\pi \widehat{f}}{\pi \widehat{f}+\widehat{\delta}}} = -\frac{(\lambda-1)(\frac{\widehat{f}}{\widehat{f}+\widehat{\delta}})^2}{(\lambda \frac{\pi \widehat{f}}{\pi \widehat{f}+\widehat{\delta}} - \frac{\widehat{f}}{\widehat{f}+\widehat{\delta}})^2} = -(\frac{N_2}{\frac{\pi \widehat{f}}{\pi \widehat{f}+\widehat{\delta}}})^2 \frac{1}{\lambda-1}, \tag{96}$$

$$\frac{\partial [\frac{\xi N_2}{(N_2\xi+g_c)} e^{g_c(t-t_1)} + \frac{g_c}{N_2\xi+g_c} e^{-N_2\xi(t-t_1)}]}{\partial N_2} = \frac{\xi}{N_2\xi+g_c} [\frac{-N_2\xi}{N_2\xi+g_c} e^{g_c(t-t_1)} - \frac{g_c}{N_2\xi+g_c} e^{-N_2\xi(t-t_1)}$$

$$+ e^{g_c(t-t_1)} - g_c(t-t_1)e^{-N_2\xi(t-t_1)}] = \frac{\xi g_c}{(N_2\xi+g_c)^2} [e^{g_c(t-t_1)} - (1+(N_2\xi+g_c)(t-t_1))e^{-N_2\xi(t-t_1)}], \tag{97}$$

Substituting the above expressions and $\frac{N_1}{N_2} = \frac{L}{\lambda-1}$ into $\frac{\partial U^*}{\partial \frac{\pi \widehat{f}}{\pi \widehat{f}+\widehat{\delta}}}$, we can get

$$\frac{\partial U^*}{\partial \frac{\pi \widehat{f}}{\pi \widehat{f}+\widehat{\delta}}} = \frac{L}{(\lambda-1)^2} (\frac{N_2}{\frac{\pi \widehat{f}}{\pi \widehat{f}+\widehat{\delta}}})^2 \frac{\xi g_c}{(N_2\xi+g_c)^2} [e^{g_c(t-t_1)} - (1+(N_2\xi+g_c)(t-t_1))e^{-N_2\xi(t-t_1)}] (\frac{\widehat{f}}{\widehat{f}+\widehat{\delta}} - \frac{\pi \widehat{f}}{\pi \widehat{f}+\widehat{\delta}}) - L_l, \tag{98}$$

Tranform the expression of L_l as follows

$$L_l = -\frac{L}{\lambda-1} [\frac{\xi N_2}{(N_2\xi+g_c)} e^{g_c(t-t_1)} + \frac{g_c}{N_2\xi+g_c} e^{-N_2\xi(t-t_1)}] + \frac{\lambda L}{\lambda-1}, \tag{99}$$

Combining (102) and (103), we have

$$\frac{\partial U^*}{\partial \frac{\pi\widehat{f}}{\pi\widehat{f}+\widehat{\delta}}} < 0 \Leftrightarrow \lambda > (\frac{N_2}{\frac{\pi\widehat{f}}{\pi\widehat{f}+\widehat{\delta}}})^2 \frac{\frac{1}{\lambda-1}\xi g_c}{(N_2\xi+g_c)^2}[e^{g_c(t-t_1)} - (1+(N_2\xi+g_c)(t-t_1))e^{-N_2\xi(t-t_1)}](\frac{\widehat{f}}{\widehat{f}+\widehat{\delta}} - \frac{\pi\widehat{f}}{\pi\widehat{f}+\widehat{\delta}})$$
$$+[\frac{\xi N_2}{(N_2\xi+g_c)}e^{g_c(t-t_1)} + \frac{g_c}{N_2\xi+g_c}e^{-N_2\xi(t-t_1)}], \tag{100}$$

The right-hand side expression is an increasing function of t. When $t = t_1$, the value of the right-hand side expression is 1, and the inequality holds automatically. Obviously when $t \to \infty$, the value converges to infinity. Therefore the inequality holds only for a finite interval. To demonstrate it holds for $(t_1, t_{2,l})$, we simply need to check it holds at time $t_{2,l}$. And $t_{2,l}$ is determined by the following equality

$$L_l|_{t_{2,l}} = \frac{E_2^l}{\frac{\pi\widehat{f}}{\pi\widehat{f}+\widehat{\delta}}}|_{t_{2,l}} \Leftrightarrow$$
$$-\frac{L}{\lambda-1}[\frac{\xi N_2}{(N_2\xi+g_c)}e^{g_c(t_{2,l}-t_1)} + \frac{g_c}{N_2\xi+g_c}e^{-N_2\xi(t_{2,l}-t_1)}] + \frac{\lambda L}{\lambda-1} = \frac{\frac{L}{\lambda-1}}{\frac{\pi\widehat{f}}{\pi\widehat{f}+\widehat{\delta}}}\frac{N_2 g_c}{N_2\xi+g_c}(e^{g_c(t_{2,l}-t_1)} - e^{-N_2\xi(t_{2,l}-t_1)})$$
$$\lambda = \frac{1}{\frac{\pi\widehat{f}}{\pi\widehat{f}+\widehat{\delta}}}\frac{N_2 g_c}{N_2\xi+g_c}(e^{g_c(t_{2,l}-t_1)} - e^{-N_2\xi(t_{2,l}-t_1)}) + [\frac{\xi N_2}{(N_2\xi+g_c)}e^{g_c(t_{2,l}-t_1)} + \frac{g_c}{N_2\xi+g_c}e^{-N_2\xi(t_{2,l}-t_1)}], \tag{101}$$

Combining (104) and (105), we have

$$\frac{\partial U^*}{\partial \frac{\pi\widehat{f}}{\pi\widehat{f}+\widehat{\delta}}}|_{t_{2,l}} < 0 \Leftrightarrow (\frac{N_2}{\frac{\pi\widehat{f}}{\pi\widehat{f}+\widehat{\delta}}})^2 \frac{\frac{1}{\lambda-1}\xi g_c}{(N_2\xi+g_c)^2}[e^{g_c(t_{2,l}-t_1)} - (1+(N_2\xi+g_c)(t_{2,l}-t_1))e^{-N_2\xi(t_{2,l}-t_1)}](\frac{\widehat{f}}{\widehat{f}+\widehat{\delta}} - \frac{\pi\widehat{f}}{\pi\widehat{f}+\widehat{\delta}})$$
$$< \frac{1}{\frac{\pi\widehat{f}}{\pi\widehat{f}+\widehat{\delta}}}\frac{N_2 g_c}{N_2\xi+g_c}(e^{g_c(t_{2,l}-t_1)} - e^{-N_2\xi(t_{2,l}-t_1)}) \Leftarrow (\frac{N_2}{\frac{\pi\widehat{f}}{\pi\widehat{f}+\widehat{\delta}}})^2 \frac{\frac{1}{\lambda-1}\xi g_c}{(N_2\xi+g_c)^2}(\frac{\widehat{f}}{\widehat{f}+\widehat{\delta}} - \frac{\pi\widehat{f}}{\pi\widehat{f}+\widehat{\delta}}) < \frac{1}{\frac{\pi\widehat{f}}{\pi\widehat{f}+\widehat{\delta}}}\frac{N_2 g_c}{N_2\xi+g_c}$$
$$\Leftrightarrow \frac{N_2\xi}{N_2\xi+g_c}(\frac{\widehat{f}}{\widehat{f}+\widehat{\delta}} - \frac{\pi\widehat{f}}{\pi\widehat{f}+\widehat{\delta}}) < (\lambda-1)\frac{\pi\widehat{f}}{\pi\widehat{f}+\widehat{\delta}}$$
$$\Leftarrow \frac{\widehat{f}}{\widehat{f}+\widehat{\delta}} < \lambda\frac{\pi\widehat{f}}{\pi\widehat{f}+\widehat{\delta}} \tag{102}$$

The last inequality holds according to our assumption $\lambda > \frac{\pi\widehat{f}+\widehat{\delta}}{\pi\widehat{f}+\pi\widehat{\delta}}$. It must be satisfied to ensure that mismatch is not serious enough and low-skilled workers eventually move into industry 2. It also ensures that $N_1$and N_2 both are positive. Since both upper bond unemployment rate and aggregate unemployment rate rise up when π decreases, aggregate unemployment rate reaches a larger peak value with greater mismatch. The time to reach its peak value is also put forward. We only need to show this by showing

$\frac{\partial L_l}{\partial \frac{\pi \hat{f}}{\pi \hat{f}+\hat{\delta}}}|_t > 0$ when $t_1 < t < t_{2,l}$. According to our former analysis, we have

$$\frac{\partial L_l}{\partial \frac{\pi \hat{f}}{\pi \hat{f}+\hat{\delta}}} = -\frac{N_1}{N_2} \frac{\partial [\frac{\xi N_2}{(N_2\xi+g_c)} e^{g_c(t-t_1)} + \frac{g_c}{N_2\xi+g_c} e^{-N_2\xi(t-t_1)}]}{\partial N_2} \frac{\partial N_2}{\partial \frac{\pi \hat{f}}{\pi \hat{f}+\hat{\delta}}}$$
$$= -\frac{N_1}{N_2} \frac{\xi g_c}{(N_2\xi+g_c)^2} [e^{g_c(t-t_1)} - (1+(N_2\xi+g_c)(t-t_1))e^{-N_2\xi(t-t_1)}] \frac{\partial N_2}{\partial \frac{\pi \hat{f}}{\pi \hat{f}+\hat{\delta}}} > 0 \tag{103}$$

The last inequality holds because $e^{(N_2\xi+g_c)(t-t_1)} > (N_2\xi+g_c)(t-t_1)+1$ and $\frac{\partial N_2}{\partial \frac{\pi \hat{f}}{\pi \hat{f}+\hat{\delta}}} < 0$. Since $\frac{E_2^l}{\frac{\pi \hat{f}}{\pi \hat{f}+\hat{\delta}}}$ rises up and L_l goes down when π decreases, the waiting time $t_{2,l} - t_1$ when they intersects each other is shortened. The rest of proof is straightforward. Since the upper bond unemployment rate turns into lower bond unemployment rate, aggregate unemployment rate must be larger at the same waiting time $t_{2,l} - t_1$ when π decreases to π'. And then aggregate unemployment rate decreases exponentially at a even slower rate $\frac{\pi' \hat{f}}{\pi' \hat{f}+\hat{\delta}}\xi < \frac{\pi \hat{f}}{\pi \hat{f}+\hat{\delta}}\xi$. Thus for the second half, aggregate unemployment rate is also above the initial level. Figure 7 shows the shift of aggregate unemployment rate.

A6. Proof of proposition 6

With some abuse of notation, we set the Lagrange as follows

$$(E_0^l + E_0^h) + \lambda \min\{E_1^l + E_1^h, \frac{K_1}{a}\} + \lambda^2 \min\{E_1^l + E_2^h, \frac{K_2}{a^2}\} + \lambda^3 \frac{K_3}{a^3} - \sum_{i\in\{l,h\}} \sum_{j=0}^{2} V_j^i c$$
$$+\mu_k(E - a\sum_{i\in\{l,h\}} E^i - a^2 \sum_{i\in\{l,h\}} E_2^i - K_3) + \sum_{i\in\{l,h\}} \mu_i (L_i - \sum_{j=0}^{2}(E_j^i + U_j^i)) + \sum_{i\in\{l,h\}} \sum_{j=0}^{2} (\hat{A}_j^i q(\theta_j^i) U_j^i - \hat{\delta} E_j^i), \tag{104}$$

where $\hat{A}_j^i = \hat{A}$ if $i \neq l$ or $j \neq 2$, and $\hat{A}_2^l = \pi \hat{A}$. The Kuhn-Tucker conditions are

$$1 - \mu_i - \hat{\delta}\mu_0^i \leq 0, E_0^i \geq 0, (1 - \mu_i - \hat{\delta}\mu_0^i)E_0^i = 0, i \in \{l, h\} \tag{105}$$

$$\lambda - \mu_i - \hat{\delta}\mu_1^i - a\mu_k \leq 0, E_1^i \geq 0, (\lambda - \mu_i - \hat{\delta}\mu_1^i - a\mu_k)E_1^i = 0, i \in \{l, h\} \tag{106}$$

$$\lambda^2 - \mu_i - \hat{\delta}\mu_2^i - a^2\mu_k \leq 0, E_2^i \geq 0, (\lambda^2 - \mu_i - \hat{\delta}\mu_2^i - a^2\mu_k)E_2^i = 0, i \in \{l, h\} \tag{107}$$

$$-\mu_i + \mu_j^i \hat{A}_j^i (q(\theta_j^i) - \theta_j^i q\prime(\theta_j^i)) \leq 0, U_j^i \geq 0, [-\mu_i + \mu_j^i \hat{A}_j^i (q(\theta_j^i) - \theta_j^i q\prime(\theta_j^i))]U_j^i = 0, i \in \{l, h\}, j \in \{0, 1, 2\} \tag{108}$$

$$-c + \mu_j^i \hat{A}_j^i q'(\theta_j^i) \leq 0, V_j^i \geq 0, [-c + \mu_j^i \hat{A}_j^i q'(\theta_j^i)]V_j^i = 0, i \in \{l, h\}, j \in \{0, 1, 2\} \tag{109}$$

Consider the first case where industry 0 and industry 1 are in production. We must have $E_0^i > 0$, $U_0^i > 0$, and $V_0^i > 0$ which implies that $1 - \mu_i - \hat{\delta}\mu_0^i = 0$, $-\mu_i + \mu_0^i(q(\theta_0^i) - \theta_0^i q'(\theta_0^i)) = 0$, and $-c + \mu_0^i q'(\theta_0^i) = 0$.

Combining these conditions, we can get

$$\mu_i = 1 - \frac{\widehat{\delta}}{\widehat{A}q'(\theta_0^i)}c = \frac{q(\theta_0^i) - \theta_0^i q'(\theta_0^i)}{q'(\theta_0^i)}c, \tag{110}$$

The concavity induces that $\frac{d(q(\theta_0^i)-\theta_0^i q'(\theta_0^i))}{d\theta_0^i} = -\theta_0^i q''(\theta_0^i) > 0$. Since the second part is decreasing with θ_0^i while the third part is increasing with θ_0^i, θ_0^i is uniquely determined by c, $\widehat{\delta}$ and $\widehat{A}$. Therefore the job finding rate $\widehat{A}q(\theta_0^i)$ in industry 0is uniquely determined and constant. In terms of industry 1, we have

$$\mu_i = \frac{q(\theta_1^i) - \theta_1^i q'(\theta_1^i)}{q'(\theta_1^i)}c, \tag{111}$$

The last part indicates that $\theta_1^i = \theta_0^i$ and the job finding rate in industry $1\widehat{A}q(\theta_1^i)$ is just the same as the job finding rate in industry 0. Consider the case where industry 1 and industry 2 are in production. For high-skilled workers, Kuhn-Tucker conditions imply that

$$\mu_h = \lambda - \frac{\widehat{\delta}}{\widehat{A}q'(\theta_1^h)}c - a\mu_k = \frac{q(\theta_1^h) - \theta_1^h q'(\theta_1^h)}{q'(\theta_1^h)}c, \tag{112}$$

$$\mu_h = \lambda^2 - \frac{\widehat{\delta}}{\widehat{A}q'(\theta_2^h)}c - a^2\mu_k = \frac{q(\theta_2^h) - \theta_2^h q'(\theta_2^h)}{q'(\theta_2^h)}c, \tag{113}$$

Combining above expressions, we must have $\theta_1^h = \theta_2^h$. By eliminating μ_k, we derive the following equation to determine θ_1^h and θ_2^h

$$\lambda\frac{a-\lambda}{a-1} = \frac{q(\theta_1^h) - \theta_1^h q'(\theta_1^h) + \frac{\widehat{\delta}}{\widehat{A}}}{q'(\theta_1^h)}c, \tag{114}$$

The first part is constant while the second part is increasing with θ_1^h. Thus θ_1^h (or θ_2^h) is uniquely determined by c, $\widehat{\delta}$, $\widehat{A}$, a and λ. The job finding rates are the same for high-skilled workers in industry 1 and industry 2. For low-skilled workers, we can similarly prove $\theta_1^l = \theta_2^l$ and derive the following equation

$$\lambda\frac{a-\lambda}{a-1} = \frac{q(\theta_1^l) - \theta_1^l q'(\theta_1^l) + \frac{\widehat{\delta}}{\widehat{A}}\frac{a-\frac{1}{\pi}}{a-1}}{q'(\theta_1^l)}c, \tag{115}$$

The expression induces a unique solution for θ_1^l (or θ_2^l) determined by c, $\widehat{\delta}$, $\widehat{A}$, a and λ and π. The job finding rate of low-skilled workers in industry 2 is π times that of low-skilled workers in industry 1 because of mismatch. Finally, it is straightforward to check for both cases that given the labor endowment L_l and L_h, with larger capital flow E, a larger share of workers searches and works in industry $j+1$ to meet the market clearing conditions.

A7. Supplementary proof of infinite industries model

We apply the forward induction method to solve the dynamic problem. Within each upgrading from industry j to industry $j+1$, the process to derive a tractable solution is very similar to one in the benchmark model. Let $\bar{L}_j^l \equiv L_j^l(t_{j+1}^l)$ and $\bar{L}_j^h \equiv L_j^h(t_{j+1}^l)$ The consumption goods reaches $\lambda^j(\frac{\pi\widehat{f}}{\pi\widehat{f}+\widehat{\delta}}\bar{L}_j^l + \frac{\widehat{f}}{\widehat{f}+\widehat{\delta}}\bar{L}_j^h)$ at the start of transition t_{j+1}^l and grows exponentially at rate g_c. The production of final commodity in the consumption goods sector must satisfy the following condition

$$\lambda^j(\frac{\pi\widehat{f}}{\pi\widehat{f}+\widehat{\delta}}\bar{L}_j^l + \frac{\widehat{f}}{\widehat{f}+\widehat{\delta}}\bar{L}_j^h)e^{g_c(t-t_{j+1}^l)} = \lambda^j(\frac{\widehat{f}}{\widehat{f}+\widehat{\delta}}L_j^h + \frac{\pi\widehat{f}}{\pi\widehat{f}+\widehat{\delta}}L_j^l) + \lambda^{j+1}(\frac{\widehat{f}}{\widehat{f}+\widehat{\delta}}L_{j+1}^h + \frac{\pi\widehat{f}}{\pi\widehat{f}+\widehat{\delta}}L_{j+1}^l), \quad (116)$$

Eliminating the term λ^j on both sides, we can get

$$(\frac{\pi\widehat{f}}{\pi\widehat{f}+\widehat{\delta}}\bar{L}_j^l + \frac{\widehat{f}}{\widehat{f}+\widehat{\delta}}\bar{L}_j^h)e^{g_c(t-t_{j+1}^l)} = (\frac{\widehat{f}}{\widehat{f}+\widehat{\delta}}L_j^h + \frac{\pi\widehat{f}}{\pi\widehat{f}+\widehat{\delta}}L_j^l) + \lambda(\frac{\widehat{f}}{\widehat{f}+\widehat{\delta}}L_{j+1}^h + \frac{\pi\widehat{f}}{\pi\widehat{f}+\widehat{\delta}}L_{j+1}^l), \quad (117)$$

The form of expression is independent of j except for the time t_{j+1}^l and initial value $\bar{L}_j^l$ and $\bar{L}_j^h$ which are the low-skilled workers and high-skilled workers in industry j when transition begins. Since low-skilled workers move first and learning efficiency is the same ξ across industries, overall low-skilled workers in both industry j and industry $j+1$ are decreasing exponentially. That is

$$L_j^l + L_{j+1}^l = e^{-\frac{\pi\widehat{f}}{\pi\widehat{f}+\widehat{\delta}}\xi(t-t_{j+1}^l)}\bar{L}_j^l, \quad (118)$$

Combining the above expressions, we have

$$\begin{aligned}(\frac{\pi\widehat{f}}{\pi\widehat{f}+\widehat{\delta}}\bar{L}_j^l + \frac{\widehat{f}}{\widehat{f}+\widehat{\delta}}\bar{L}_j^h)e^{g_c(t-t_{j+1}^l)} &= \frac{\widehat{f}}{\widehat{f}+\widehat{\delta}}(L - e^{-\frac{\pi\widehat{f}}{\pi\widehat{f}+\widehat{\delta}}\xi(t-t_{j+1}^l)}\bar{L}_j^l - L_{j+1}^h) \\ &+ \frac{\pi\widehat{f}}{\pi\widehat{f}+\widehat{\delta}}(e^{-\frac{\pi\widehat{f}}{\pi\widehat{f}+\widehat{\delta}}\xi(t-t_{j+1}^l)}\bar{L}_j^l - L_{j+1}^l) + \lambda(\frac{\widehat{f}}{\widehat{f}+\widehat{\delta}}L_{j+1}^h + \frac{\pi\widehat{f}}{\pi\widehat{f}+\widehat{\delta}}L_{j+1}^l), \quad (119)\end{aligned}$$

Rearranging it we can get

$$E_{j+1}^l = \frac{\pi\widehat{f}}{\pi\widehat{f}+\widehat{\delta}}L_{j+1}^l = M_j e^{g_c(t-t_{j+1}^l)} + \Phi_j e^{-\xi\frac{\pi\widehat{f}}{\pi\widehat{f}+\widehat{\delta}}(t-t_{j+1}^l)} - \frac{\widehat{f}}{\widehat{f}+\widehat{\delta}}\frac{L}{\lambda-1} - \frac{\widehat{f}}{\widehat{f}+\widehat{\delta}}L_{j+1}^h, \quad (120)$$

where

$$M_j = \frac{\frac{\pi\widehat{f}}{\pi\widehat{f}+\widehat{\delta}}\bar{L}_j^l + \frac{\widehat{f}}{\widehat{f}+\widehat{\delta}}\bar{L}_j^h}{\lambda-1}, \Phi_j = \frac{\frac{\widehat{f}}{\widehat{f}+\widehat{\delta}} - \frac{\pi\widehat{f}}{\pi\widehat{f}+\widehat{\delta}}}{\lambda-1}\bar{L}_j^l, \quad (121)$$

The evolution of high-skilled workers in industry $j+1$ follows

$$\dot{L}_{j+1}^h = \xi E_{j+1}^l \quad (122)$$

Substituting the expression of E_{j+1}^l into the evolution equation, we can get a linear and inhomogeneous differential equation. By solving the differential equation, we can find the expression of E_{j+1}^l and L_{j+1}^h with respect to time t if $t_{j+1}^l \leq t < t_{j+1}^h$

$$L_{j+1}^h = \frac{M_j\xi}{g_c + \xi\frac{\widehat{f}}{\widehat{f}+\widehat{\delta}}}e^{g_c(t-t_{j+1}^l)} + \frac{\bar{L}_j^l}{\lambda-1}e^{-\xi\frac{\pi\widehat{f}}{\pi\widehat{f}+\widehat{\delta}}(t-t_{j+1}^l)} + (\frac{\bar{L}_j^h}{\lambda-1} - \frac{\xi M_j}{g_c + \xi\frac{\widehat{f}}{\widehat{f}+\widehat{\delta}}})e^{-\xi\frac{\widehat{f}}{\widehat{f}+\widehat{\delta}}(t-t_{j+1}^l)} - \frac{L}{\lambda-1}, \quad (123)$$

$$E^l_{j+1} = \frac{M_j g_c}{g_c + \xi\frac{\widehat{f}}{\widehat{f}+\widehat{\delta}}} e^{g_c(t-t^l_{j+1})} - \frac{\frac{\pi\widehat{f}}{\pi\widehat{f}+\widehat{\delta}}\bar{L}^l_j}{\lambda-1} e^{-\xi\frac{\pi\widehat{f}}{\pi\widehat{f}+\widehat{\delta}}(t-t^l_{j+1})} - (\frac{\frac{\widehat{f}}{\widehat{f}+\widehat{\delta}}\bar{L}^h_j}{\lambda-1} - \frac{M_j\xi\frac{\widehat{f}}{\widehat{f}+\widehat{\delta}}}{g_c+\xi\frac{\widehat{f}}{\widehat{f}+\widehat{\delta}}}) e^{-\xi\frac{\widehat{f}}{\widehat{f}+\widehat{\delta}}(t-t^l_{j+1})}. \tag{124}$$

We next prove that L^h_{j+1} and E^l_{j+1} are increasing functions of t. We only need to prove E^l_{j+1} is increasing with time since L^h_{j+1} increases automatically if $E^l_{j+1} > 0$. Firstly, we assume $\frac{dE^l_{j+1}}{dt}|t^l_{j+1} > 0$. Later on we will prove this premise is true. Taking the derivation of E^l_{j+1} with respect to t, we have

$$\begin{aligned}\frac{dE^l_{j+1}}{dt} &= \frac{M_j g_c^2}{g_c + \xi\frac{\widehat{f}}{\widehat{f}+\widehat{\delta}}} e^{g_c(t-t^l_{j+1})} + \xi\frac{(\frac{\pi\widehat{f}}{\pi\widehat{f}+\widehat{\delta}})^2\bar{L}^l_j}{\lambda-1} e^{-\xi\frac{\pi\widehat{f}}{\pi\widehat{f}+\widehat{\delta}}(t-t^l_{j+1})} \\ &+\xi\frac{\widehat{f}}{\widehat{f}+\widehat{\delta}}(\frac{\frac{\widehat{f}}{\widehat{f}+\widehat{\delta}}\bar{L}^h_j}{\lambda-1} - \frac{M_j\xi\frac{\widehat{f}}{\widehat{f}+\widehat{\delta}}}{g_c+\xi\frac{\widehat{f}}{\widehat{f}+\widehat{\delta}}}) e^{-\xi\frac{\widehat{f}}{\widehat{f}+\widehat{\delta}}(t-t^l_{j+1})},\end{aligned} \tag{125}$$

Combining with following inequality condition

$$g_c > -\xi\frac{\pi\widehat{f}}{\pi\widehat{f}+\widehat{\delta}} > -\xi\frac{\widehat{f}}{\widehat{f}+\widehat{\delta}}, \tag{126}$$

It indicates that

$$\frac{dE^l_{j+1}}{dt} \geq e^{-\xi\frac{\widehat{f}}{\widehat{f}+\widehat{\delta}}(t-t^l_{j+1})}\frac{dE^l_{j+1}}{dt}|t^l_{j+1} > 0. \tag{127}$$

Notice that without the initial inequality $\frac{dE^l_{j+1}}{dt}|t^l_{j+1} > 0$, E^l_{j+1} can be a decreasing function and shows a negative value. Consider the case where learning efficiency ξ is large and consumption growth rate g_c is small, and $\bar{L}^l_j$ is much larger than $\bar{L}^h_j$. Quick learning allows low-skilled workers to stay at industry j to produce all the consumption goods, and there is no incentive for them to move into industry $j+1$. However, $\bar{L}^l_j$ and $\bar{L}^h_j$ are not arbitrary. They are history-dependent and related to ξ and g_c. The induced values of $\bar{L}^l_j$ and $\bar{L}^h_j$ can rule out these extreme cases. In terms of a mathematical term, it is equivalent to show $\frac{dE^l_{j+1}}{dt}|t^l_{j+1} > 0$ by using the forward induction method. Our former analysis indicates that $\frac{dE^l_{j+1}}{dt}|t^h_{j+1} > 0$ if $\frac{dE^l_{j+1}}{dt}|t^l_{j+1} > 0$. So we only need to show $\frac{dE^l_{j+1}}{dt}|t^l_{j+1} > 0$ if $\frac{dE^l_j}{dt}|t^h_j > 0$ which is proved in the second half. Consider the industrial upgrading of high-skilled workers from industry j to industry $j+1$ from time t^h_{j+1}. Similarly we have the following equality equation for the consumption goods production

$$\begin{aligned}[\lambda^j\frac{\widehat{f}}{\widehat{f}+\widehat{\delta}}\bar{\bar{\bar{L}}}^h_j + \lambda^{j+1}(\frac{\pi\widehat{f}}{\pi\widehat{f}+\widehat{\delta}}\bar{\bar{L}}^l_{j+1} + \frac{\widehat{f}}{\widehat{f}+\widehat{\delta}}\bar{\bar{L}}^h_{j+1})]e^{g_c(t-t^h_{j+1})} &= \lambda^j\frac{\widehat{f}}{\widehat{f}+\widehat{\delta}}(L - L^h_{j+1} - L^l_{j+1}) \\ &+\lambda^{j+1}(\frac{\widehat{f}}{\widehat{f}+\widehat{\delta}}L^h_{j+1} + \frac{\pi\widehat{f}}{\pi\widehat{f}+\widehat{\delta}}L^l_{j+1}),\end{aligned} \tag{128}$$

Transforming the equation, we have

$$E^l_{j+1} = Q_j e^{g_c(t-t^h_{j+1})} - N_2 L^h_{j+1} - N_1. \tag{129}$$

where N_1 and N_2 follow the former definitions and

$$Q_j = \frac{\frac{\pi\widehat{f}}{\pi\widehat{f}+\widehat{\delta}}(\frac{\widehat{f}}{\widehat{f}+\widehat{\delta}}\bar{\bar{\bar{L}}}_j^h + \lambda\frac{\pi\widehat{f}}{\pi\widehat{f}+\widehat{\delta}}\bar{\bar{L}}_{j+1}^l + \lambda\frac{\widehat{f}}{\widehat{f}+\widehat{\delta}}\bar{\bar{L}}_{j+1}^h)}{\lambda\frac{\pi\widehat{f}}{\pi\widehat{f}+\widehat{\delta}} - \frac{\widehat{f}}{\widehat{f}+\widehat{\delta}}}, \tag{130}$$

Combining with the evolution equation of L_{j+1}^h, we find a first order linear differential equation. By solving this equation, we derive a tractable solution for E_{j+1}^l and L_{j+1}^h from t_{j+1}^h to t_{j+2}^l which is shown as follows

$$L_{j+1}^h = (\bar{\bar{L}}_{j+1}^h + \frac{N_1}{N_2} - \frac{\xi Q_j}{g_c + \xi N_2})e^{-\xi N_2(t-t_{j+1}^h)} + \frac{\xi Q_j}{g_c+\xi N_2}e^{g_c(t-t_{j+1}^h)} - \frac{N_1}{N_2}, \tag{131}$$

$$E_{j+1}^l = \frac{Q_j g_c}{N_2\xi + g_c}e^{g_c(t-t_{j+1}^h)} - (N_1 + N_2\bar{\bar{L}}_{j+1}^h - \frac{Q_j\xi N_2}{g_c+\xi N_2})e^{-N_2\xi(t-t_{j+1}^h)}, \tag{132}$$

Consider the case where $j = 1$. It refers to the industrial upgrading from industry 1 to industry 2 when mismatch firstly takes place. We know the initial values $\bar{\bar{\bar{L}}}_1^h = L$, $\bar{\bar{L}}_2^l = \bar{\bar{L}}_2^h = 0$ since all workers can be treated as high-skilled in industry 1. Thus we must have $Q_1 = N_1$ and expressions of L_2^h and E_2^l are the same as those in the benchmark model. As shown earlier, $\frac{dE_2^l}{dt} > 0$ and $\frac{dL_2^h}{dt} > 0$ must hold at any t from t_2^h to t_3^l. Especially at the time t_3^l it indicates that without industrial upgrading workers can not produce the required amount of consumption goods by improvement of skill in the sunrise industry. It requires low-skilled workers in industry 2 to move into industry 3. In this sense, it provides us an initial inequality condition for this upward induction. Consider any $j \geq 2$. We only need to prove E_{j+1}^l is increasing. Taking the derivation of E_{j+1}^l, we have

$$\frac{dE_{j+1}^l}{dt} = \frac{Q_j g_c^2}{N_2\xi + g_c}e^{g_c(t-t_{j+1}^h)} + N_2\xi(N_1 + N_2\bar{\bar{L}}_{j+1}^h - \frac{Q_j\xi N_2}{g_c+\xi N_2})e^{-N_2\xi(t-t_{j+1}^h)}, \tag{133}$$

Likewise $\frac{dE_{j+1}^l}{dt}$ can be negative for some ξ,g_c, $\bar{\bar{\bar{L}}}_{j+1}^h$, $\bar{\bar{L}}_{j+1}^l$, $\bar{\bar{L}}_{j+1}^h$ since $(N_1 + N_2\bar{\bar{L}}_{j+1}^h - \frac{Q_j\xi N_2}{g_c+\xi N_2})$ is not necessarily positive. If the learning efficiency is large, the consumption growth rate is small and there exist masses of low-skilled workers in the sunrise industry $j+1$, workers in industry j can have no incentive to move into industry $j+1$. But with the initial condition $\frac{dE_{j+1}^l}{dt}|t_{j+1}^h > 0$, we must have $\frac{dE_{j+1}}{dt} > 0$ using the following inequality

$$\frac{dE_{j+1}^l}{dt} \geq e^{-N_2\xi(t-t_{j+1}^h)}\frac{dE_{j+1}^l}{dt}|t_{j+1}^h > 0 \tag{134}$$

At time t_{j+2}^l when workers all stay in industry $j+1$, $\frac{dE_{j+1}^l}{dt}|t_{j+2}^l > 0$ indicates that without further industrial upgrading, workers can not produce enough consumption goods. Thus it also implies $\frac{dE_{j+2}^l}{dt}|t_{j+2}^l > 0$ since there is no worker in industry j for upgrading. Combining the above discussion, the case $j = 1$ gives the initial condition that $\frac{dE_3^l}{dt}|t_3^l > 0$, and then $\frac{dE_3^l}{dt}|t_3^l > 0$, and then $\frac{dE_3^l}{dt}|t_3^h > 0$. Applying this forward induction, we complete the proof of increasing E_{j+1}^l and L_{j+1}^h from t_{j+1}^l to t_{j+2}^l. Simultaneously, we can uniquely determine the span of each upgrading $t^h_{j+1} - t^l_{j+1}$ and $t^l_{j+2} - t^h_{j+1}$ and initial values $\bar{L}^l_{j+1}$, $\bar{\bar{L}}^l_{j+1}$, $\bar{L}^h_{j+1}$, $\bar{\bar{L}}^h_{j+1}$ and $\bar{\bar{\bar{L}}}^h_{j+1}$. From t^l_{j+1} to t^h_{j+1}, since E^l_{j+1} is increasing while $E^l_j + E^l_{j+1}$ is decreasing, it induces a unique $t^h_{j+1} - t^l_{j+1}$ when all low-skilled workers in industry j move into industry $j+1$. From t^h_{j+1} to t^l_{j+2}, the trend that E^l_{j+1} and L^h_{j+1} are increasing with time induces a unique $t^l_{j+2} - t^h_{j+1}$ when all workers move into industry $j+1$. The dynamics of industrial upgrading as well as labor market performance can be easily derived using the above equations and we omit this part. It is obvious that the equilibrium shows a symmetric pattern. In terms of E^l_{j+1} and L^h_{j+1}, the function form is identical except for parameteres related to the initial allocation of workers. The reason to derive a symmetric solution for this infinite-industry model is owing to our symmetric model setting. We assume that mismatch is the same when workers first move in, and learning efficiency is the same across industries. We also assume that capital intensity a^j increases a times and the TFP of industry production λ^j increases λ times for the following sunrise industry. It helps us to derive a very simple form of solution to characterize the whole dynamics of economy with unceasing industrial upgrading.

References

[1] Acemoglu, Daron, and Veronica Guerrieri. 2008. "Capital Deepening and Nonbalanced Economic Growth", *Journal of Political Economy* 116 (June): 467-498.

[2] Alvarez, Fernando and Robert Shimer. 2011. "Search and Rest Unemployment", *Econometrica,* Vol. 79, No. 1 (January), 75-122.

[3] Hosios, A. J. 1994. "Unemployment and Vacancies with Sectoral Shifts", *American Economic Review*, 84(1), 124.

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

[5] Buera, Francisco J. and Joseph P. Kaboski. 2012. "The Rise of the Service Economy", *American Economic Review*, 102(6): 2540-69.

[6] Chenery, Hollis B. 1960. "Patterns of Industrial Growth," *American Economic Review* 50 (September): 624-654.

[7] Comin, Diego & Danial Lashkari & Mart Mestieri, 2021. "Structural Change With Long-Run Income and Price Effects," Econometrica vol. 89(1), pages 311-374

[8] Davis, S. and J. Haltiwanger. 1992, "Gross Job Creation, Gross Job Destruction, and Employment Reallocation", The *Quarterly Journal of Economics*, Vol.107, 819-863.

[9] Haltiwanger et al., 2003, "Job Creation and Job Destruction in Transition Countries", *Economics of Transition*, Vol.11, 205-219.

[10] Herrendorf, Rogerson and Valentinyi. 2011. "Growth and Structural Transformation", manuscript prepared for the *Handbook of Economic Growth.*

[11] Hosios, A. J. 1994. "Unemployment and Vacancies with Sectoral Shifts", *American Economic Review*, 84(1), 124.

[12] Jovanovic, Boyan. 1982. "Selection and the Evolution of Industry," *Econometrica,* 50(3): 649-670.

[13] Ju, Jiandong, Justin Yifu Lin, and Yong Wang. 2011. "Industrial Dynamics, Endowment Structure, and Economic Growth", HKUST working paper

[14] Kuznets, Simon. 1966. *Modern Economic Growth: Rate, Structure, and Spread.* New Heaven: Yale University Press

[15] Lee, Donghoon and Kenneth Wolpin. 2006. "Intersectoral Labor Mobility and the Growth of the Service Sector," *Econometrica*, 74(1): 1-46

[16] Lilien, David M. 1982. "Sectoral Shifts and Cyclical Unemployment." The Journal of Political Economy, 90(4), 777-793.

[17] Lucas, Robert E., Jr., and Prescott, Edward C. 1974. "Equilibrium Search and Unemployment." J. Econ. Theory 7 (February): 188-209.

[18] Pascual Restrepo. 2015. "Skill Mismatch and Structural Unemployment", *Working Paper.*

[19] Pissarides, Christopher. 2000. *Equilibrium Unemployment Theory* (2nd edition). Cambridge: The MIT Press

[20] Pissarides, Christopher. 2007. "Unemployment and Hours of Work: The North Atlantic Divide Revisited." *International Economic Review*, 48 (February):1-36.

[21] Rogerson, Richard. 2008. "Structural Transformation and the Deterioration of European Labor Market Outcomes," *Journal of Political Economy*, vol. 116(2): 235-259.

[22] Romer, Paul M. "Endogenous Technological Change." *Journal of Political Economy* 98, no. 5 (1990): S71-102.

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

[24] Ngai, L. Rachel, and Barbara Petrongolo. 2012, "Structural Transformation, Marketization and Female Employment.", LSE working paper

[25] Şahin, Ayşegül, Joseph Song, Giorgio Topa, and Giovanni L. Violante. 2014. "Mismatch Unemployment." *American Economic Review,* 104 (11): 3529-64.

[26] Shimer, Robert. 2007. "Mismatch," *American Economic Review*, 97(4): 1074-1101.

[27] Shimer, Robert. 2005. "The Cyclical Behavior of Equilibrium Unemployment and Vacancies," *American Economic Review,* 95(1): 25–49.

[28] Wang, Yong. 2012. "Market Structure, Factor Endowment, and Industrial Upgrading", HKUST working paper

[29] Zagler M. 2009. "Economic growth, structural change, and search unemployment". *Journal of Economics*, 96(1):63-78.