Labor Market Mismatch, Structural Unemployment and Industry Dynamics

Siyu Chen Yong Wang Lijun Zhu*

April 16, 2023

Abstract

We develop a multi-industry growth model with labor market mismatch to explore the interaction between life-cycle industry dynamics, aggregate unemployment rate and economic growth. We show that, without any exogenous aggregate shocks, the aggregate unemployment rate exhibits recurrent cycles along with the perpetual structural change, which is in turn driven by capital accumulation. Experienced workers in an old industry lose their industry-specific expertise when they are relocated to a more capital-intensive industry and suffer mismatch, so the aggregate unemployment rate rises. However, these workers gradually become experienced through on-the-job learning, so mismatch attenuates and the aggregate unemployment rate declines, till the sunrise industry itself becomes a sunset one and workers have to move to an even more capital-intensive industry, so the unemployment rate rises again, ad infinitum. We analytically characterize the properties of dynamic labor market performance, life-cycle dynamics of each of the underlying infinite industries, and the aggregate consumption growth rates. We show that the model predictions are consistent with data.

JEL classification: E24, E37, O14, O41

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

*Chen: Department of Economics, Washington University in St. Louis, E-mail: siyu.c@wustl.edu. Wang (corresponding author): Institute of New Structural Economics, Peking University, E-mail: yongwang@nsd.pku.edu.cn. Zhu: Institute of New Structural Economics, Peking University, E-mail: lijunzhu@nsd.pku.edu.cn. For useful comments, we thank Rachel Ngai, Robert Shimer, Yongs Shin, Ping Wang and other participants of numerous seminars and conferences. Wang acknowledges financial support from the grant NSSF no:20AJL017.The usual disclaimer applies.

1 Introduction

The aggregate growth of an economy is accompanied with changes in its underlying composition of industries. It has been well documented that, as GDP per capita increases, the employment (or value added) share of the agriculture sector declines, the share of the manufacturing sector follows a hump shape, and the share of the service sector increases(Kuznets, 1966). The causes and consequences of such phenomena, also known as the Kuznets facts, have been intensively studied in the literature (Kongsamut et al., 2001; Ngai and Pissarides, 2007; Herrendorf et al., 2014). Meanwhile, structural change also occurs at more disaggregated levels. In particular, within the manufacturing sector, it is empirically established that labor-intensive industries are gradually replaced by more capital-intensive ones as capital accumulates along the growth path. Moreover, each industry exhibits a hump-shaped life-cycle dynamic pattern (Chenery, 1960; Ju et al., 2015). Such a dynamics process is also referred to as industrial upgrading in this paper.

Jobs in different industries usually require different expertise. A Worker newly relocated from a sunset industry is typically unfamiliar with a sunrise industry including the corresponding labor market. She presumably has no clues about implicit job requirement, no tips for job interviews, and no personal connections, so she is less likely to find a job that well matches her skills in the sunrise industry than in the mature but sunset one. Consequently, she faces a lower job finding rate when searching for jobs in a new industry. We refer to such kind of difficulties encountered by relocated workers as labor market mismatch, or mismatch for short. Employment reallocation across sectors might thus generate fluctuations in the aggregate unemployment rate, an observation that has been documented since at least Lilien (1982). Figure 1.1 plots the sector reallocation rate, measured by the standard deviation of employment growth rate across sectors, and the aggregate unemployment rate in the US from 1955 to 2020.[1] As we can see, high unemployment rates are typically associated with significant employment reallocation across sectors. We show in Appendix A that this positive correlation is empirically very robust.

[1] For detailed discussions about the precise definitions and measurement of sector reallocation rate and the related data sources, see Appendix A.

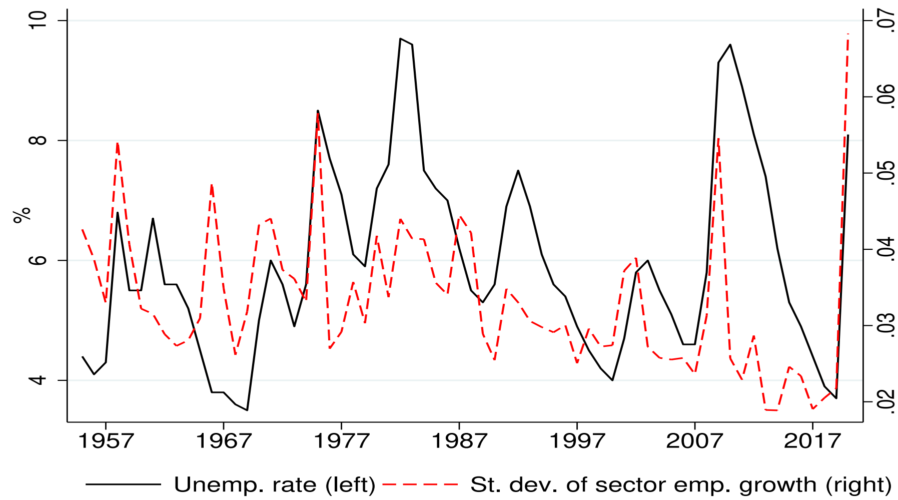


Figure 1.1: US Unemployment Rate and Sector Reallocation Rate: 1955-2020

Note: Sector reallocation rate is measured as the standard deviation of annual employment growth rates across sectors.

However, it remains unclear whether and how the industrial upgrading process, especially the life-cycle dynamics of the sunrise and sunset industries, is logically linked to the fluctuation of aggregate unemployment rates. The primary goal of this paper is, therefore, to theoretically explore this question by investigating how mismatch endogenously emerges and varies in the process of industrial upgrading and how it may impact structural unemployment, industry dynamics and aggregate growth. To this end, we develop a tractable dynamic general equilibrium model with multiple industries, which are heterogeneous in capital intensities. It is the capital accumulation, rather than unbalanced productivity growth or nonhomotheticity of preferences, that endogenously drives life-cycle industry dynamics and structural change, a distinctly important feature shared by Ju et al. (2015). However, different from their assumptions of frictionless labor markets and homogeneous labor, our model explicitly introduces labor market frictions, heterogeneity in experiences across workers and heterogeneity in job requirement across industries.

We model mismatch in a way similar to Restrepo (2015), but our paper differs in two important dimensions: whereas his model assumes that industries are ex ante symmetric and that structural changes are driven by unbalanced sectoral productivity growth, our model assumes that industries are asymmetric and heterogeneous in capital intensities and that the driving force of structural change is capital accumulation. In our model, workers relocated from an old (less capital-intensive) industry are all initially inexperienced in the new (more capital-intensive) industry, and hence suffer a low job finding rate due to mismatch between workers' expertise for the old industry and job requirement of the new industry. These workers, once employed in the new industry, could become experienced at a Poisson arrival rate through on-the-job learning, and thereafter alleviated from mismatch and enjoy a higher job finding rate upon exogenous job separation in this new industry. As capital accumulates, the current new industry also eventually becomes a sunset industry, gradually replaced by an even more capital-intensive industry. Workers then have to reallocate again, so experienced workers become inexperienced workers again after they move into the new industry, suffering mismatch and a low job finding rate, so on and so forth.

We show that in equilibrium the aggregate unemployment rate exhibits a cyclical pattern along with the perpetual changes in underlying industrial compositions. More specifically, in our model, when workers from an old industry have all just been relocated into the new industry, the aggregate unemployment rate reaches the peak because of the mismatch explained above. Then as the sunrise industry grows maturer, the unemployment rate decreases as inexperienced workers gradually turn into experienced ones through on-the-job learning. The decline continues till the current industry itself becomes a sunset industry and gradually replaced by an even more capital-intensive industry, as capital becomes sufficiently cheaper when the economy grows. Inexperienced workers in this mature industry start to move to the new industry, but experienced workers would only move later as their opportunity cost of switching industries is higher. During this phase, mismatch continues to attenuate due to the on-the-job learning in both sunrise and sunset industries, so the average job finding rate in both industries increases and thus the aggregate unemployment rate declines. The unemployment rate reaches the bottom when all the inexperienced workers in the sunset industry have moved to the sunrise industry. It takes some time before experienced workers in the sunset industries start to move out, as it would happen only when capital further accumulates to a certain threshold. During this period, all remaining workers in the sunset industries are experienced ones and free of mismatch, but the aggregate unemployment rate continues to decline because inexperienced workers are still turning into experienced ones in the sunrise industry. After that, experienced workers in the sunset industries start to move into the sunrise industry and lose their industry-specific expertise. The resulting mismatch drives up the aggregate unemployment rate. This process continues till all experienced workers in the sunset industry have moved to the sunrise industry, at which point the old industry exits the market and all workers, employed or unemployed, are in the labor market for the new industry. The equilibrium unemployment rate reaches the maximum value, which completes one cycle. As capital further accumulates for a while, an even more capital-intensive industry emerges as a new sunrise industry and it starts to expand, attracting inexperienced workers in the old sunset industry into it, so the equilibrium unemployment rate declines again, ad infinitum.

Despite the analytical challenges that result from the high dimentionality and nonlinearity of the dynamic structural problem, we are able to obtain closed-form solutions to fully characterize the whole dynamics. Moreover, it also enables us to conduct clean and unambiguous comparative static analyses. It turns out that three parameters play key roles: the capital-goods production efficiency (denoted by A), the Poisson rate of experience accumulation due to on-the-job learning(ξ), and the degree of mismatch (π). We show that, a larger A translates into faster capital accumulation, a higher speed of sunrise industries replacing sunset ones, a shorter life span of each industry, a universally higher level of aggregate unemployment rate, and a faster growth of aggregate consumption. In particular, the aggregate unemployment rate becomes higher because workers must relocate to new industries more frequently and hence mismatch occurs more frequently. Moreover, a higher ξ, i.e. faster experience accumulation, implies a universally lower aggregate unemployment rate and a shorter life span of each industry, because on-the-job learning dampens the negative impact of mismatch and facilitates relocation of experienced workers into new industries. Last, a lower π, or severer mismatch, implies a universally higher unemployment rate and a longer industry life span because inexperienced workers suffer lower job finding rates and so experienced workers become even more reluctant to move to new industries.

Our paper is most 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 (Kongsamut et al., 2001; Buera and Kaboski, 2012; Boppart, 2014; Comin et al., 2021), or substitution effect due to unbalanced productivity growth across sectors (Ngai and

Pissarides, 2007; Acemoglu and Guerrieri, 2008), the driving force of structural transformation (and related industry dynamics) in our model is capital accumulation, which changes relative factor prices, resulting in labor-intensive industries gradually replaced by capital-intensive ones. This mechanism is first articulated in Ju et al. (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 degenerates to Ju et al. (2015).

Whereas full employment is typically assumed in the existing literature of structural transformation (e.g. Ngai and Pissarides (2008)), our paper is, to the best of our knowledge, the first to study the interactions of industrial upgrading, labor market mismatch and the cyclical movement of aggregate unemployment rate. Pissarides (2007) develops a three-sector model with labor search and match, but he assumes that manufacturing and market service require the same skills and the labor markets for the two sectors are fully integrated, so structural transformation itself does not generate structural unemployment. Another distinct difference of his model is that the driving force for structural transformation is exogenous unbalanced productivity growth across sectors rather than capital accumulation. 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, with the driving force for sectoral reallocation being endogenous innovation instead of capital accumulation as we highlight.

Secondly, our paper is related to the macro labor literature on mismatch and fluctuation of unemployment rates, which typically focuses on mismatch between job-seekers and vacancies at the business cycle frequency (Shimer, 2007; Alvarez and Shimer, 2011).[2] To resolve the Shimer puzzle (Shimer, 2005), Shimer (2007) constructs a model in which mismatch results from exogenous and random assignment of identical workers to ex ante symmetric segregated local labor markets, so labor supply exceeds demand in some markets while the opposite is true for others. Instead of such quantity mismatch, we model mismatch between the knowledge possessed by experienced workers from

[2] Sahin et al. (2014) find 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. This finding is further confirmed by Herz and van Rens (2019), who propose an accounting framework to estimate the overall magnitude of mismatch-induced unemployment, and they conclude that mismatch is responsible for 84% of unemployment rate.

old industries and job requirement in new industries. Moreover, workers in our model endogenously choose which sectors to seek jobs instead of being exogenously and randomly assigned to different markets. More importantly, our model does not attempt to explain the business-cycle frequency Shimer puzzle but rather to explain the recurring unemployment rate cycles associated with structural transformation over the long-run economic growth. Restrepo (2015) shows that the impact of great recessions on labor markets can be amplified with mismatch, and generate higher volatility than the canonical Mortensen-Pissarides model. He assumes that both the obsolescence of old jobs and the structural shift are exogenous, whereas in our paper both are endogenous.

The existing literature on cyclical unemployment with a multi-industry setting typically assumes exogenous sectoral productivity shocks to symmetric industries (e.g., Lucas and Prescott (1974), Lilien (1982)). In contrast, our model assumes that industries are heterogeneous in capital intensity and hence endogenously asymmetric in their dynamics. In addition, the endogenous unemployment rate cycles in our model are not caused by any exogenous aggregate or sectoral shocks, but rather by endogenous capital accumulation, because it drives the repeated process of industrial upgrading, which in turn leads to cycles of old experience obsolesces and new experience accumulation via on-the-job learning. In this sense, our model speaks more closely to medium-run cycles than to standard short-run business cycles. In a recent research, Boldrin et al. (2022) build a model in which innovation drives constant upgrading from less to more capital-intensive technologies. That model generates endogenous cycles in factor income shares, but it does not address structural change or unemployment as it assumes both homogeneous labor and perfect labor markets.

The rest of the paper is structured as follows: Section 2 presents a simplified model with finite industries to illustrate the key forces behind the unemployment rate cycle. Section 3 provides a fully-fledged model with infinite industries to show that the unemployment rate cycle obtained in Section 2 is endogenously recurrent and to explain how the life-cycle dynamics of each of the industries and fluctuation of the aggregate unemployment rate are both driven by capital accumulation in the process of structural transformation with frictional labor markets. Section 4 concludes. Technical proofs are mostly delegated to the Appendix.

2 A Simple Finite-Industry Model

This section presents a simple finite-industry dynamic model with imperfect labor markets. The purpose is to analytically characterize the endogenous aggregate unemployment rate and industrial dynamics along the growth path. Industries are heterogeneous in capital intensities; structural transformation, i.e. sectoral reallocation across industries, is mainly driven by capital accumulation, which is essentially the same as Ju et al. (2015). The model has two new features: (1) experience is industry-specific in the sense that all workers relocated from other industries are initially inexperienced workers for the new industry, who may become experienced workers at a Poisson rate; (2) labor markets are frictional in the sense that the job finding rate is finite and it is higher for experienced workers than inexperienced workers.

Consider a continuous-time economy with unit mass of identical households. Each household is initially endowed with physical capital K, inexperienced labor L^l and experienced labor L^h. There are two sectors. 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 $K(t)$ denote 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 a positive coefficient capturing the rate of the investment-specific technological progress. A larger A implies a more efficient capital goods production. The newly produced capital flow is split between two different usages:

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

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

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

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 non-negative intermediate input produced by industry n for $n =$

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 exhibit constant returns to scale:

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

with $a > \lambda > 1$. Therefore an industry with a higher index is not only more capital intensive ($a > 1$), but also has a higher labor productivity ($\lambda > 1$) for $n = 0, 1, 2$. The assumption $a > \lambda$ rules out the trivial case that only the most capital-intensive industry would produce if capital is used to produce consumption goods.

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

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

where $\sigma > 0$ is the reciprocal of the intertemporal elasticity of substitution and $\rho > 0$ is time discount rate. 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 function well, so no mismatch occurs and the job finding follows a Poisson process with constant rate f for all workers in those two industries.[3] However, in industry 2 inexperienced workers suffer from mismatch so their job finding rate is πf, lower than that for experienced workers f, where $\pi \in (0, 1)$. A lower π means severer mismatch. This is the only difference mismatch makes. All employed workers in the same industry have identical productivity, independent of their experience levels. Jobs separate at an exogenous rate δ for all workers in all industries. Inexperienced workers in industry 2, when hired, would benefit from on-the-job learning and become experienced permanently at an exogenous Poisson rate ξ in that industry. A higher ξ implies higher efficiency of on-the-job learning. Experience is industry-specific, so when a worker moves to a new industry, she automatically becomes an inexperienced worker in that industry no matter whether she was experienced or inexperienced in the

[3] The job finding rate is assumed constant not just for simplicity. We show in our appendix B.7.2 that it is consistent with the socially efficient allocation in the model of Mortensen and Pissarides (1994) with a matching function.

old industry.

Experience in our model captures workers' knowledge about a industry. For example, workers from a sunset industry might not have enough knowledge about which subset of jobs matches her the best among the various opportunities offered in a sunrise industry. These types of mismatch thus results in an initially low job finding rate. Overtime, as the worker gradually gains experience and expands her network in the new industry, the mismatch problem become less severe. So that she can locate a suitable job at a greater rate.

Figure 2.1 summarizes the model environment for the consumption good sector.

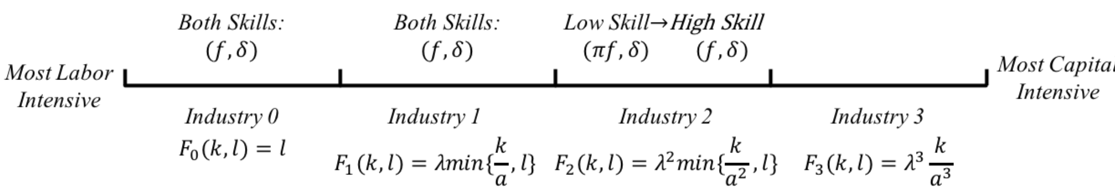


Figure 2.1: Environment in the consumption good sector with multiple industries

We use the subscripts h and l to denote experienced and inexperienced works. Let $U_h(t)$ and $U_l(t)$ denote, respectively, the number of unemployed experienced and inexperienced workers in the whole economy at time t. Let $E^i_j(t), U^i_j(t)$ and $L^i_j(t)$ denote, respectively, the numbers of employed workers, unemployed workers, and total labor force, with experience type i in the labor market for industry j at time t , for $i \in \{l, h\}$, $j \in \{0, 1, 2\}$. Obviously, $L^i_j(t) = U^i_j(t) + E^i_j(t)$.

Let $\tilde{\Omega}(C, E^l_0 + E^h_0, E^l_1 + E^h_1, E^l_2 + E^h_2)$ denote the minimum amount of capital required to produce the final consumption C when current total employment in the first three industries are given by $E^l_0 + E^h_0$, $E^l_1 + E^h_1$ and $E^l_2 + E^h_2$, respectively. It is straightforward to show that

$$\tilde{\Omega}(\cdot) = \begin{cases} \frac{a}{\lambda}C - \frac{a}{\lambda}(E^l_0 + E^h_0), & \text{if} \quad C_0 < C \leq C_1 \\ \frac{a^2}{\lambda^2}C - \frac{a^2}{\lambda^2}(E^h_0 + E^l_0) - \frac{a(a-\lambda)}{\lambda}(E^h_1 + E^l_1), & \text{if} \quad C_1 < C \leq C_2 \\ \frac{a^3}{\lambda^3}C - \frac{a^3}{\lambda^3}(E^h_0 + E^l_0) - \frac{a(a^2-\lambda^2)}{\lambda^2}(E^h_1 + E^l_1) - \frac{a^2(a-\lambda)}{\lambda}(E^h_2 + E^l_2), & \text{if} \quad C_2 < C \end{cases}$$

where

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

Substituting $\tilde{\Omega}(\cdot)$ into (2), we obtain the following capital accumulation equation:

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

Further denote λ_j^i the fraction of unemployed workers with experience type $i \in \{l, h\}$ who seek jobs in industry $j \in \{0, 1, 2\}$.[4] Then the law of motion for inexperienced employment in industry 0 is

$$\dot{E}_0^l = fU_l(1 - \lambda_1^l - \lambda_2^l) - \delta E_0^l, \tag{6}$$

which states that the change in the total number of employed inexperienced workers in industry 0, $\dot{E}_0^l$, is equal to the new employment flow net of exogenous job separation δE_0^l at each time point. There are $U_l \lambda_0^l$ inexperienced unemployed workers in industry 0, where $\lambda_0^l = 1 - \lambda_1^l - \lambda_2^l$, so the flow of new employment in that industry is $fU_l(1 - \lambda_1^l - \lambda_2^l)$ because of no mismatch. Likewise, the law of motion for employed experienced workers in industry 0 is given by

$$\dot{E}_0^h = fU_h(1 - \lambda_1^h - \lambda_2^h) - \delta E_0^h. \tag{7}$$

Similarly, the total number of employment of inexperienced and experienced workers in industry 1 and 2 evolves as follows, respectively:

$$\dot{E}_1^l = fU_l \lambda_1^l - \delta E_1^l, \tag{8}$$

$$\dot{E}_1^h = fU_h \lambda_1^h - \delta E_1^h, \tag{9}$$

$$\dot{E}_2^l = \pi f U_l \lambda_2^l - \delta E_2^l - \xi E_2^l, \tag{10}$$

$$\dot{E}_2^h = fU_l \lambda_2^h - \delta E_2^h + \xi E_2^l. \tag{11}$$

In particular, inexperienced workers in industry 2 suffer from mismatch, so their job finding rate is lower, given by πf, as stated in (10).

[4] By definition, $0 \leq \lambda_1^i, 0 \leq \lambda_2^i, \lambda_1^i + \lambda_2^i \leq 1, i \in \{l, h\}$.

Last, the total number of unemployment for inexperienced and experienced workers, respectively, evolves as follows

$$\dot{U}_h = \delta \sum_{j=0}^{2} E_j^h - fU_h, \tag{12}$$

$$\dot{U}_l = \delta \sum_{j=0}^{2} E_j^l - fU_l(1-\lambda_2^l) - \pi f U_l \lambda_2^l, \tag{13}$$

where (12) states that the change in aggregate experienced unemployment is equal to the new unemployment flow due to exogenous separation in the three industries net of new experienced employment flow and (13 states that inexperienced unemployment is equal to the new unemployment flow caused by 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, $\pi f U_l \lambda_2^l$.

The artificial benevolent social planner's problem is to choose the optimal flow of consumption $C(t)$,investment $I(t)$, capital for consumption production $\Omega(t)$, and how to allocate unemployed workers $U_h(t)$ and $U_l(t)$ into three industries to maximize the utility of a representative household as shown in (5), subject to the constraints specified in equations (2)-(13), given $K(0), U_i(0), E_j^i(0), i \in \{l, h\}, j \in \{0, 1, 2\}$.

We show that the decentralized competitive equilibrium can be characterized by resorting to the above social planner's problem, which is formally stated in the following proposition

Proposition 1. *There exists a decentralized market equilibrium allocation identical to the artificial social planner's allocation.*

Proof. Please see Appendix B.1. ■

The intuition is as follows. Recall that in the standard Diamond-Mortensen-Pissarides model, the 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. When this condition holds, 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 planner. Our model is an extreme and degenerate case where the bargaining power parameter for workers and the matching elasticity with respect to unemployment are both equal to one. Therefore, 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 characterize the transitional dynamics for the planner's problem. It is widely recognized that analytically characterizing the whole transitional dynamic path is unwieldy even in a one-sector growth environment, let alone in a multi-sector dynamic setting such as in our current model. To sharpen the analysis, we assume that gross flows between employment and unemployment are sufficiently large such that labor markets adjust fast enough to fully catch up with changes in the capital stock. See Restrepo (2015) for a similar treatment. In other words, labor allocations resemble the steady state equilibrium for any given capital allocation between the capital goods sector and the consumption good sector. More formally, let $f = \kappa \widehat{f}$, $\delta = \kappa \widehat{\delta}$, and suppose $\kappa \to \infty$. This simplification enables us to derive a tractable solution for the dynamic path of all industries and the aggregate economy. Moreover, such simplification is innocuous for our purpose because our model intends to focus on the medium-run cycles and long-run growth rather than short-run dynamics at the business-cycle frequencies.

Now we formally define the optimal steady state:

Definition 1. *The optimal steady state is the equilibrium of the following economy: technologies are given by (3) and (4); the factor endowment vector is given by (Ω, L^l, L^h), where Ω denotes the capital flow devoted to the production of consumption goods, L^l and L^h denote experienced and inexperienced workers respectively; fraction $\frac{\widehat{f}}{\widehat{f}+\widehat{\delta}}$ of all experienced workers are employed in industries 0,1,2; of all inexperienced workers, fraction $\frac{\widehat{f}}{\widehat{f}+\widehat{\delta}}$ are employed in industries 0 and 1 and fraction $\frac{\pi\widehat{f}}{\pi\widehat{f}+\widehat{\delta}}$ are employed in industry 2; and the output of final consumption commodity is maximized given the factor endowment vector and technologies.*

The following lemma states that the dynamic optimization problem degenerates to static optimization at each instant.

Lemma 1. *For an instantaneous equilibrium (where $f = \kappa \widehat{f}$, $\delta = \kappa \widehat{\delta}$, and $\kappa \to \infty$), dynamic optimization requires that the economy is at an optimal steady state at each instant t if $\xi < \bar{\xi}$*

with $\bar{\bar{\xi}} \equiv \min_{t\in[0,T]}\{\frac{e^{-\rho t}C(t)^{-\sigma}}{\mu_e(t)}(\frac{\pi\widehat{f}+\widehat{\delta}}{\pi\widehat{f}+\pi\widehat{\delta}} + a(\lambda-1) - \lambda^2)\}$, *where* $\mu_e(t)$ *denotes the value to become an experienced worker and* T *is the last time point when the output of final consumption good is equal to* $\lambda^2 \frac{\widehat{f}}{\widehat{f}+\widehat{\delta}}L$.

Proof. Please see Appendix B.2. ∎

When the job finding rate f and separation rate δ are sufficiently large, both employment and unemployment reach the steady state almost immediately. Dynamic optimization requires that labor markets adjust efficiently to reduce the amount of capital flow used for producing consumption goods. It is equivalent to finding an optimal steady state equilibrium to produce the required amount of consumption goods with minimum amount of capital flow. The condition $\xi < \bar{\bar{\xi}}$ is to ensure that the current benefit of capital flow reduction outweighs the future benefit of experience accumulation for inexperienced workers in industry 2. Mismatch must be sufficiently mild (i.e., π is large) so that inexperienced workers have sufficient incentives to enter industry 2.

For the convenience of future references, we collect all the assumptions on the model coefficients and label them as Assumption 1.

Assumption 1: *The following conditions are satisfied:*

$$
\begin{aligned}
&(i)\ 0 < A - \rho < \sigma A;\\
&(ii)\ f = \kappa\widehat{f},\ \delta = \kappa\widehat{\delta},\ \kappa \to \infty;\\
&(iii)\ \xi < \bar{\bar{\xi}};\\
&(iv)\ \frac{\pi\widehat{f}}{\pi\widehat{f}+\widehat{\delta}} > (\frac{1}{\lambda}+\frac{1}{a})\frac{\widehat{f}}{\widehat{f}+\widehat{\delta}}
\end{aligned}
$$

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

$$C = \max_{E^i_j, K_j}\{(E^l_0 + E^h_0) + \lambda\min\{E^l_1 + E^h_1, \frac{K_1}{a}\} + \lambda^2\min\{E^l_1 + E^l_2, \frac{K_2}{a^2}\} + \lambda^3\frac{K_3}{a^3}\} \tag{14}$$

subject to

$$K_1 + K_2 + K_3 \leq \Omega, \tag{15}$$

$$E_0^l \frac{\widehat{f}+\widehat{\delta}}{\widehat{f}} + E_1^l \frac{\widehat{f}+\widehat{\delta}}{\widehat{f}} + E_2^l \frac{\pi\widehat{f}+\widehat{\delta}}{\pi\widehat{f}} \leq L^l, \tag{16}$$

and

$$\left(E_0^h + E_1^h + E_2^h\right) \frac{\widehat{f}+\widehat{\delta}}{\widehat{f}} \leq L^h. \tag{17}$$

Note that (15)-(17) are the resource constraints for capital, inexperienced workers and experienced workers, respectively. In particular, $E_0^l \frac{\widehat{f}+\widehat{\delta}}{\widehat{f}}$ inexperienced workers must be allocated to industry 0 in order to have E_0^l inexperienced workers employed in that industry in the instantaneous equilibrium. The second term on the left-hand side in (16) is the amount of inexperienced workers allocated to industry 1. The third term is the amount of inexperienced workers allocated to industry 2, where mismatch occurs. Observe that ξ does not show up in the above optimization problem because it takes time for inexperienced workers to become experienced, and Assumption 1 ensures that labor markets adjustment (i.e., sectoral employment and unemployment of both inexperienced and experienced workers) reach the instantaneous equilibrium steady state before the factor endowment vector (Ω, L^l, L^h) changes. The following proposition fully characterizes the optimal steady state:

Proposition 2. *Under Assumption 1, in the optimal steady state equilibrium with given capital endowment for consumption production* Ω*, endowment of inexperienced labor* L^l *and experienced labor* L^h*, the labor allocation is summarized in Table 2.1.* [5]

[5]For future references, we also derive the required amount of capital inflow Ω as a function of C, L^h and L^l, see table B.1 in Appendix B.3

Table 2.1: optimal steady state equilibrium

(1). $0 \le \Omega < a\frac{\widehat{f}}{\widehat{f}+\widehat{\delta}}(L^h+L^l)$	(2). $a\frac{\widehat{f}}{\widehat{f}+\widehat{\delta}}(L^h+L^l) \le \Omega < \frac{\widehat{f}}{\widehat{f}+\widehat{\delta}}(aL^l+a^2L^h)$
$C = \frac{\lambda-1}{a}\Omega + \frac{\widehat{f}}{\widehat{f}+\widehat{\delta}}(L^h+L^l)$ $E_0^l + E_0^h = \frac{\widehat{f}}{\widehat{f}+\widehat{\delta}}(L^h+L^l) - \frac{\Omega}{a}$ $E_1^l + E_1^h = \frac{\Omega}{a}$ $E_2^l = 0,\ E_2^h = 0$	$C = \frac{\lambda^2-\lambda}{a^2-a}\Omega + \frac{\lambda(a-\lambda)}{a-1}\frac{\widehat{f}}{\widehat{f}+\widehat{\delta}}(L^h+L^l)$ $E_0^l = 0,\ E_0^h = 0$ $E_1^l = \frac{\widehat{f}}{\widehat{f}+\widehat{\delta}}L^l,\ E_1^h = \frac{a^2\frac{\widehat{f}}{\widehat{f}+\widehat{\delta}}(L^h+L^l)-\Omega}{a^2-a} - \frac{\widehat{f}}{\widehat{f}+\widehat{\delta}}L^l$ $E_2^l = 0,\ E_2^h = \frac{\Omega - a\frac{\widehat{f}}{\widehat{f}+\widehat{\delta}}(L^h+L^l)}{a^2-a}$
(3). $\frac{\widehat{f}}{\widehat{f}+\widehat{\delta}}(aL^l+a^2L^h) \le \Omega < a^2(\frac{\pi\widehat{f}}{\pi\widehat{f}+\widehat{\delta}}L^l + \frac{\widehat{f}}{\widehat{f}+\widehat{\delta}}L^h)$	(4). $\Omega \ge a^2(\frac{\pi\widehat{f}}{\pi\widehat{f}+\widehat{\delta}}L^l + \frac{\widehat{f}}{\widehat{f}+\widehat{\delta}}L^h)$
$C = \frac{\lambda^2\frac{\pi\widehat{f}}{\pi\widehat{f}+\widehat{\delta}} - \lambda\frac{\widehat{f}}{\widehat{f}+\widehat{\delta}}}{a^2\frac{\pi\widehat{f}}{\pi\widehat{f}+\widehat{\delta}} - a\frac{\widehat{f}}{\widehat{f}+\widehat{\delta}}}\Omega + \frac{\widehat{f}}{\widehat{f}+\widehat{\delta}}\frac{\lambda(a-\lambda)}{a\frac{\pi\widehat{f}}{\pi\widehat{f}+\widehat{\delta}} - \frac{\widehat{f}}{\widehat{f}+\widehat{\delta}}}(\frac{\pi\widehat{f}}{\pi\widehat{f}+\widehat{\delta}}L^l + \frac{\widehat{f}}{\widehat{f}+\widehat{\delta}}L^h)$ $E_0^l = 0,\ E_0^h = 0$ $E_1^l = \frac{\widehat{f}}{\widehat{f}+\widehat{\delta}}(\frac{a^2(\frac{\pi\widehat{f}}{\pi\widehat{f}+\widehat{\delta}}L^l + \frac{\widehat{f}}{\widehat{f}+\widehat{\delta}}L^h) - \Omega}{a^2\frac{\pi\widehat{f}}{\pi\widehat{f}+\widehat{\delta}} - a\frac{\widehat{f}}{\widehat{f}+\widehat{\delta}}}),\ E_1^h = 0$ $E_2^l = \frac{\pi\widehat{f}}{\pi\widehat{f}+\widehat{\delta}}(\frac{\Omega - a^2\frac{\widehat{f}}{\widehat{f}+\widehat{\delta}}L^h - a\frac{\widehat{f}}{\widehat{f}+\widehat{\delta}}L^l}{a^2\frac{\pi\widehat{f}}{\pi\widehat{f}+\widehat{\delta}} - a\frac{\widehat{f}}{\widehat{f}+\widehat{\delta}}}),\ E_2^h = \frac{\widehat{f}}{\widehat{f}+\widehat{\delta}}L^h$	$C = \frac{\lambda^3}{a^3}\Omega - \frac{\lambda^2(a-\lambda)}{a}(\frac{\pi\widehat{f}}{\pi\widehat{f}+\widehat{\delta}}L^l + \frac{\widehat{f}}{\widehat{f}+\widehat{\delta}}L^h)$ $E_0^l = 0,\ E_0^h = 0$ $E_1^l = 0,\ E_1^h = 0$ $E_2^l = \frac{\pi\widehat{f}}{\pi\widehat{f}+\widehat{\delta}}L^l,\ E_2^h = \frac{\widehat{f}}{\widehat{f}+\widehat{\delta}}L^h$

Proof. Please see Appendix B.3. ■

As implied by Table 2.1, when $C \in [\frac{\widehat{f}}{\widehat{f}+\widehat{\delta}}(L^h+L^l), \lambda\frac{\widehat{f}}{\widehat{f}+\widehat{\delta}}(L^h+L^l))$, only industry 0 and industry 1 coexist and no mismatch occurs, so the experience structure of the labor force is irrelevant. When $C \in [\lambda\frac{\widehat{f}}{\widehat{f}+\widehat{\delta}}(L^h+L^l), \frac{\widehat{f}}{\widehat{f}+\widehat{\delta}}(\lambda L^l+\lambda^2L^h))$, only industries 1 and 2 coexist, and all workers in industry 2 are experienced. When $C \in [\frac{\widehat{f}}{\widehat{f}+\widehat{\delta}}(\lambda L^l + \lambda^2L^h), \lambda^2(\frac{\pi\widehat{f}}{\pi\widehat{f}+\widehat{\delta}}L^l + \frac{\widehat{f}}{\widehat{f}+\widehat{\delta}}L^h))$, only industry 1 and industry 2 coexist, and all workers in industry 1 are inexperienced. When $C \in [\lambda^2(\frac{\pi\widehat{f}}{\pi\widehat{f}+\widehat{\delta}}L^l + \frac{\widehat{f}}{\widehat{f}+\widehat{\delta}}L^h), \infty)$, only industries 2 and 3 coexist, and all workers are in industry 2.

Observe that total consumption C as a function of Ω, L^l and L^h has different functional forms in the above four different scenarios because of the endogenous differences in the underlying industrial structures, which are in turn determined by the factor endowment structure Ω, L^l and L^h. In particular, when $\widehat{f} \to \infty$ while keeping $\widehat{\delta} < \infty$, there will be no mismatch and experience types will be irrelevant, which is exactly the case in Ju et al. (2015), so Table 2.1 degenerates to the static equilibrium in Ju et al. (2015).

The optimal steady state equilibrium indicates that the required capital flows into the consumption goods sector only depends on $C(t)$, $L^l(t)$ and $L^h(t)$. The artificial social planner's dynamic problem can be rewritten as

$$\max_{C} \int_{t=0}^{\infty} e^{-\rho t} \frac{C^{1-\sigma}-1}{1-\sigma} dt$$
$$\text{s.t. } \dot{K} = AK - \Omega(C, L^l, L - L^l),$$
$$\dot{L}^l = -G(C, L^l, L - L^l),$$

where $\Omega(C, L^l, L - L^l)$ can be derived from Table 2.1 and is directly provided in table B.1 in Appendix B.3, and $G(C, L^l, L - L^l)$, which governs the evolution of inexperienced unemployment, is given by:

$$G(C, L^l, L - L^l) = \begin{cases} 0, & C \in [0, \underline{C}) \\ \xi \frac{\pi \widehat{f}}{\pi \widehat{f} + \widehat{\delta}} \left(\frac{C - \lambda^2 \frac{\widehat{f}}{\widehat{f}+\widehat{\delta}}(L-L^l) - \lambda \frac{\widehat{f}}{\widehat{f}+\widehat{\delta}} L^l}{\lambda^2 \frac{\pi \widehat{f}}{\pi \widehat{f}+\widehat{\delta}} - \lambda \frac{\widehat{f}}{\widehat{f}+\widehat{\delta}}} \right), & C \in [\underline{C}, \bar{C}] \\ \xi \frac{\pi \widehat{f}}{\pi \widehat{f} + \widehat{\delta}} L^l, & C \in (\bar{C}, \infty) \end{cases} \tag{18}$$

with

$$\bar{C} = \frac{\lambda^2 \widehat{f}}{\widehat{f} + \widehat{\delta}}(L - L^l) + \frac{\lambda^2 \pi \widehat{f}}{\pi \widehat{f} + \widehat{\delta}} L^l; \quad \underline{C} = \frac{\lambda^2 \widehat{f}}{\widehat{f} + \widehat{\delta}}(L - L^l) + \frac{\lambda \widehat{f}}{\widehat{f} + \widehat{\delta}} L^l.$$

The current-value Hamiltonian becomes

$$H = \frac{C^{1-\sigma}-1}{1-\sigma} + \mu[AK - \Omega(C, L^l, L - L^l)] - \Lambda G(C, L^l, L - L^l),$$

where Λ is the value of inexperienced workers minus the value of experienced workers, and it must be bounded and negative. The necessary optimality conditions are

$$\dot{\mu} = (\rho - A)\mu, \tag{19}$$

$$C \in arg\max\{\frac{C^{1-\sigma}-1}{1-\sigma} + \mu(AK - \Omega(C, L^l, L - L^l)) - \Lambda G(C, L^l, L - L^l)\}.$$

Thus when ξ is sufficiently small[6] [7],

$$C \approx \hat{C} \in arg\max\{\frac{C^{1-\sigma}-1}{1-\sigma} + \mu(AK - \Omega(C, L^l, L - L^l))\} \tag{20}$$

Assume all workers are inexperienced at time 0, that is, $L^l(0) = L$. Since on-the-job learning is the only way to become experienced workers, $L^l(t)$ is always equal to L before industry 2 emerges. We focus on the transition from industry 1 to 2 as mismatch matters only after inexperienced workers begin to search in industry 2. Since $\Omega(C, L^l, L - L^l)$ is a piece-wise linear function of C and μ decreases at a constant rate, the economy goes through three stages: In stage I, all workers stay in industry 1. Consumption remains constant, capital accumulates over time, and there is no mismatch. In stage II, inexperienced workers in industry 1 move into industry 2. Some of them are employed and luckily become experienced workers via on-the-job learning. All experienced workers stay in industry 2, and inexperienced workers in industry 1 gradually move into this industry. Consumption increases at a constant rate. In stage III, all workers stay in industry 2, consumption increases at a lower speed as more inexperienced workers gradually become experienced, but industry 3 is still inactive at this stage.

Before stage II starts, there are no inexperienced workers in industry 2 so there is no mismatch, the aggregate unemployment is the constant and equal to $\frac{\widehat{\delta}}{\widehat{f}+\widehat{\delta}}L$. After stage II starts, inexperienced workers move from industry 1 to industry 2 and suffer mismatch. The aggregate unemployment is given by

$$U = \frac{\widehat{\delta}}{\widehat{f}+\widehat{\delta}}L + (\frac{\widehat{\delta}}{\pi\widehat{f}+\widehat{\delta}} - \frac{\widehat{\delta}}{\widehat{f}+\widehat{\delta}})L_2^l \tag{21}$$

[6]In appendix B.8, it is shown that even for a large enough ξ, the numerical solution of the equilibrium path is almost the same as our analytical solution.

[7]When t is large, $C(t)$ is larger than $\bar{C}$. In this case, (20) is true without approximation as $G(C, L^l, L - L^l) = \xi\frac{\pi\widehat{f}}{\pi\widehat{f}+\widehat{\delta}}L^l$ which is unrelated to C.

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 when mismatch is absent, i.e., when $L_2^l = 0$. It captures the frictional unemployment resulting from labor market frictions in job search and job destruction. The second term is the extra amount of unemployment due to mismatch of inexperienced workers in industry 2, so it captures structural unemployment resulting from mismatch when workers move from industry 1 to industry 2 during the structural change. Note that $\frac{\widehat{\delta}}{\widehat{f}+\widehat{\delta}}$ 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.

More precisely, after stage II starts, the labor market dynamics is characterized by the following two equations:

$$E_2^l(t) = \begin{cases} \frac{N_1 g_c}{N_2\xi + g_c}\left(e^{g_c(t-t_1)} - e^{-N_2\xi(t-t_1)}\right) & \text{if} \quad t_1 < t \le t_{2,l} \\ e^{-\frac{\pi\widehat{f}}{\pi\widehat{f}+\widehat{\delta}}\xi(t-t_{2,l})}\frac{N_1 g_c}{N_2\xi + g_c}\left(e^{g_c(t_{2,l}-t_1)} - e^{-N_2\xi(t_{2,l}-t_1)}\right) & \text{if} \quad t > t_{2,l} \end{cases} \tag{22}$$

and

$$L^l(t) = \begin{cases} \frac{-N_1}{N_2(N_2\xi+g_c)}\left(\xi N_2 e^{g_c(t-t_1)} + g_c e^{-N_2\xi(t-t_1)}\right) + \frac{N_1}{N_2} + L & \text{if} \quad t_1 < t \le t_{2,l} \\ e^{-\frac{\pi\widehat{f}}{\pi\widehat{f}+\widehat{\delta}}\xi(t-t_{2,l})}\left[\frac{-N_1}{N_2(N_2\xi+g_c)}\left(\xi N_2 e^{g_c(t_{2,l}-t_1)} + g_c e^{-N_2\xi(t_{2,l}-t_1)}\right) + \frac{N_1}{N_2} + L\right] & \text{if} \quad t > t_{2,l} \end{cases} \tag{23}$$

where t_1 denotes the time point when stage II starts, $t_{2,l}$ is the first time when all workers search in industry 2, and ...[please explicitly explain every new notation in the above two equations] Observe that, when $t_1 < t \le t_{2,l}$, $\frac{dL^l(t)}{dt} < 0$ and $\frac{dE_2^l(t)}{dt} > 0$; after $t_{2,l}$, all workers are in industry 2, both $E_2^l(t)$ and L^l decrease exponentially because of on-the-job learning.

Combining equations (21) to (22), we obtain the following proposition about how the aggregate unemployment rate changes over time:

Proposition 3. *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 mismatch and then gradually declines because inexperienced workers gradually become experienced via on-the-job learning.*

Proof. Please see Appendix B.4. ■

To understand the non-monotonic change in the aggregate unemployment rate more intuitively, we plot its time path in Figure 2.2 (the solid curve) after the economy enters stage II at time t_1. The unemployment rate fluctuates endogenously mainly because the number of inexperienced workers in industry 2 endogenously changes. Recall that inexperienced workers in industry 2 suffer mismatch and hence a lower job finding rate than experienced workers, but inexperienced workers could become experienced through on-the-job learning. The number of inexperienced workers in industry 2, and hence unemployment rate, first increases because workers move from industry 1 to 2 as capital increases. It reaches the peak at time $t_{2,l}$, when industry 1 just exits the market and all workers in the economy are in the labor market for industry 2. Afterwards, the number of inexperienced workers in industry 2 and the unemployment rate both decline because of on-the-job learning. The unemployment rate eventually converges to $\frac{\widehat{\delta}}{\widehat{f}+\widehat{\delta}}$ as all workers become experienced and mismatch disappears. Note that industry 3 does not hire labor, so all workers will stay in the labor market for industry 2 and never leave. In other words, structural unemployment will eventually disappear and there will be only frictional unemployment as labor relocation across industries stops.

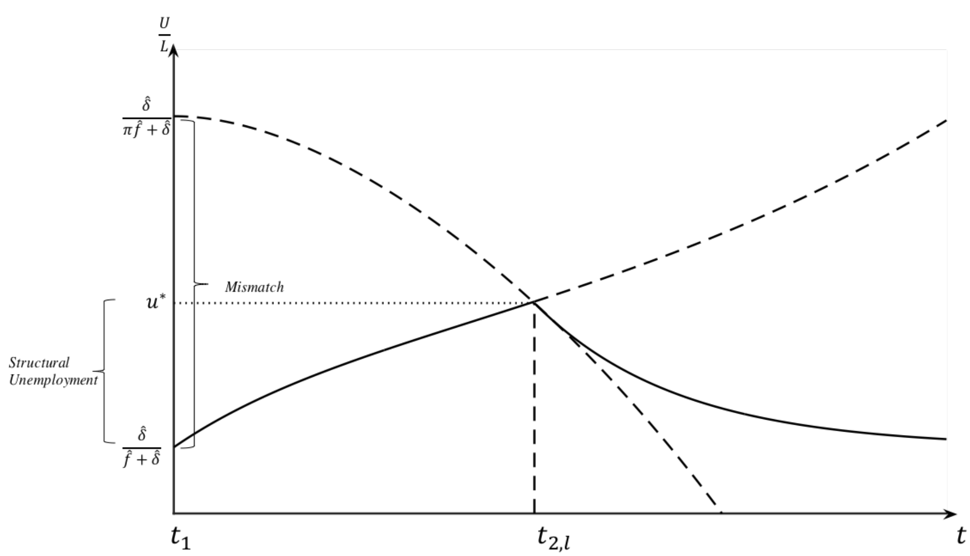


Figure 2.2: Aggregate unemployment rate

The following proposition summarizes some comparative static properties of the dynamic equilibrium.

Proposition 4. *The following three comparative static properties hold after time t_1 in equilibrium: (1) when on-the-job learning efficiency ξ becomes larger, the aggregate unemployment rate u becomes permanently lower at every time point and it reaches the peak later; (2) when capital-goods production efficiency A becomes larger, the unemployment rate u permanently becomes higher permanently before reaching its peak, and the peak value is reached earlier; (3) when the mismatch-afflicted job finding rate becomes larger (i.e., π becomes larger with f given), the unemployment rate u becomes lower permanently and it reaches the peak later.*

Proof. Please see Appendix B.5. ■

The intuition for the above proposition is as follows:

(1) *Learning efficiency* When the learning efficiency ξ increases, inexperienced workers turn into experienced more quickly, and therefore more workers have a higher job finding rate, which results in a lower unemployment rate. In addition, as the number of experienced workers increases, total employment in industry 2 also gradually increases, so industry 2 can sustain a longer period of expansion, that is, $t_{2,l}$ is delayed. Figure B.1 plots how the time path of the aggregate unemployment rate changes when on-the-job learning efficiency ξ increases.

(2) *Capital-goods production efficiency* When capital-goods production efficiency A increases, capital accumulates faster, driving a more rapid structural change from industry 1 to more capital-intensive industry 2. As inexperienced workers rush into industry 2 more quickly, mismatch results in a higher unemployment rate. The peak value of the unemployment rate also becomes higher, because workers on average have less time for on-the-job learning before they have to move into industry 2. As inexperienced workers move into industry 2 earlier and never leave afterwards, they also become experienced earlier via on-the-job learning. The unemployment rate reaches the peak earlier because the most capital-intensive industry 3 emerges earlier, which means that employment in industry 2 reaches the peak at an earlier time. The unemployment rate reaches the peak when the total employment in industry 2 reaches the peak, as shown in Figure B.2.

(3) *Mismatch* When the mismatch-afflicted job finding rate is lower, i.e., π becomes smaller, the aggregate unemployment rate becomes larger permanently and the peak

value also becomes higher. As inexperienced workers are less likely to get employed, fewer of them have the opportunity to become experienced ones via on-the-job learning. Thus the aggregate unemployment rate becomes higher even when all workers stay in the labor market for industry 2. Since consumption growth rate remains unchanged, industry 3 emerges earlier to support consumption growth as the maximum amount of employment in industry 2 becomes lower. It implies that the unemployment rate starts to decline at an earlier time. See Figure B.3.

3 The Infinite-Industry Model

While the simple four-industry model in Section 2 illustrates why the unemployment rate may first rise and then decline during the industrial upgrading from industry 1 to 2, its limitations are also obvious. First, it cannot generate repeated cycles as we observe in the real data because mismatch is only relevant for industry 2 by model construction. Second, the life span of industry 2 is infinite because industry 3 needs no labor in the production. Third, the model cannot produce endlessly repeated cycles of aggregate unemployment rates together with sustained aggregate economic growth because the number of industries within the consumption-goods sector is finite. In order to eliminate these important limitations, we now consider an alternative model, in which everything remains identical as before except that now there are countably infinite industries within the consumption good sector and mismatch could occur in infinite industries. We show that, despite of the seemingly unwieldy complexities of the dynamic problem, we are still able to obtain the closed-form solution to analytically characterize the whole dynamics of the infinite-industry model, including the recurrent cycles of aggregate unemployment rates, finite life spans for all industries, and sustainable long-run economic growth.

The final consumption good is now produced with the following technology:

$$X = \sum_{n=0}^{\infty} x_n, \tag{24}$$

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

$$F_n(k,l) = \begin{cases} l & \text{if} \quad n = 0 \\ \lambda^n \min\{\frac{k}{a^n}, l\} & \text{if} \quad n \geq 1 \end{cases} \tag{25}$$

Following Ju et al. (2015), we impose $a - 1 > \lambda > 1$ to rule out the trivial case that only the most capital-intensive good is produced. Like in the finite-industry model, workers seeking jobs in industry 0 or 1 do not suffer mismatch, and their job finding rate is f. Mismatch takes place only in industries $j \geq 2$. Experience is industry-specific. Workers become inexperienced and face a low job finding rate, πf with $0 < \pi < 1$, when they enter a new industry. After she is employed in the new industry, she might become experienced through on-the-job learning and then face a higher job finding rate f afterwards in this industry if she loses her job. Job separation rates are δ in all industries, independent of experience types of workers. A representative household is initially endowed with capital $K(0)$ and labor L, and all labor is inexperienced initially for any industry j, $j \geq 2$. The household's preference is still given by (5). Figure 3.1 summarizes the model environment for the consumption good sector with infinite industries.

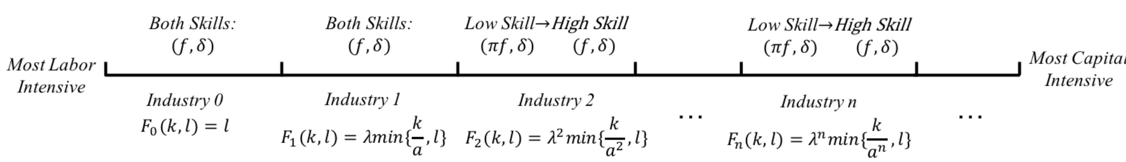


Figure 3.1: Environment in consumption good sector with infinite industries

Assumption 2: *The following conditions are satisfied:*

(i) $f = \kappa\widehat{f}, \delta = \kappa\widehat{\delta},\ \kappa \to \infty$;

(ii) $0 < A - \rho < \sigma A$;

(iii) $\xi < \bar{\xi}^{inf}$;

(iv) $\dfrac{\pi\widehat{f}}{\pi\widehat{f} + \widehat{\delta}} > \dfrac{a+\lambda}{a\lambda + 1}\dfrac{\widehat{f}}{\widehat{f} + \widehat{\delta}}$.

Assumption 2 serves four purposes: Condition (i) ensures that labor reallocation is sufficiently fast so we could sharpen the analysis by focusing on the instantaneous equilibrium for the otherwise unwieldy dynamic problem; condition (ii) ensures that the

long-run growth rate is strictly positive but not explosive; condition (iii)ensures that the on-the-job learning efficiency is small enough so that the benevolent social planner wants to minimize the flow of capital used for producing the consumption good at each instant; and condition (iv) ensures that mismatch is sufficiently mild, or equivalently, π is sufficiently large, so that workers have no incentives to skip any industry, say, directly move from industry j to industry $j+2$.
Notice that conditions (i) and (ii) are identical to those in Assumption 1 in the finite-industry model, but conditions (iii) and (iv) are different from their counterpart in Assumption 1 because....[to be completed]

We characterize the industrial dynamics by forward induction. The step-by-step industrial upgrading process from industry 0 first to industry 1, and then to industry 2, till all workers are in industry 2, is identical to that in the finite-sector model. But afterwards, the economy repeatedly goes through four stages for each upgrading process from industry n to industry $n+1$. We verbally summarize this pattern in the following proposition:

Proposition 5. *In equilibrium, the industrial upgrading process from industry n to industry $n+1$ ($n \geq 2$) experiences four stages:*

- *at stage I, all workers stay in the labor market for industry n, aggregate consumption grows at a strictly positive speed lower than g_c;*
- *at stage II, inexperienced workers in industry n gradually move into industry $n+1$, consumption grows at speed g_c;*
- *at stage III, all the experienced workers in industry n stay in industry n, and consumption grows at a rate which is strictly positive but smaller than g_c;*
- *at stage IV, experienced workers in industry n gradually move into industry $n+1$, the consumption grows at speed g_c.*

Proof. Please see Appendix B.6. ■

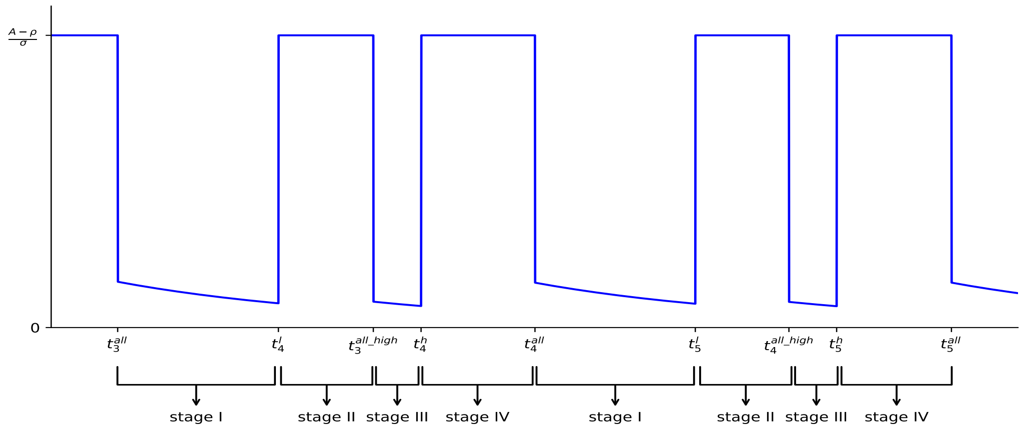


Figure 3.2: Consumption growth rate at four stages

At stage I, all employed workers stay in industry n and all unemployed workers seek jobs in industry n. There is no industrial upgrading or structural change at this stage, but capital accumulates and its shadow price declines over time. Some inexperienced workers in industry n gradually become experienced via on-the-job learning, so more and more workers are employed and we have a strictly positive consumption growth rate due to the employment increase. But the aggregate consumption growth rate is lower than g_c. At stage II, the industrial upgrading takes place. Inexperienced workers in industry n move into $n+1$, while all experienced workers still stay in the labor market for industry n, employed or not. Aggregate consumption grows at speed g_c. The number of employed workers increases in both industry n and industry $n+1$ because more and more inexperienced workers become experienced through on-the-job learning in both industries. This process continues till no inexperienced workers are employed in industry n and no inexperienced workers search for jobs in industry n. At stage III, all inexperienced workers have left industry n, and those who are either employed industry $n+1$ or looking for jobs in industry $n+1$ are all experienced workers, and no workers move across industries ,that is, industrial upgrading pauses. Inexperienced workers in industry $n+1$ gradually become experienced and the aggregate consumption grows at a strictly positive rate but still lower than g_c.[8] At stage IV,

[8]Stages I and III are the phases during which no workers in the labor market for industry n, inexpe-

experienced workers in industry n move into $n+1$, and this stage ends when all workers have entered the labor market for industry $n+1$, employed or unemployed. This four-stage process continues repetitively.[9] The explicit analytical solutions to this dynamic optimization problem are shown in details in Appendix B.6.

Let $t_n^{all}, t_{n+1}^l, t_n^{all\_high}, t_{n+1}^h$ denote the first time when all workers stay at industry n, when inexperienced workers in industry n begin to move into $n+1$, when all workers in industry n are experienced, and when experienced workers in industry n begin to move into $n+1$. So stage I is the period $t \in [t_n^{all}, t_{n+1}^l)$, stage II is $t \in [t_{n+1}^l, t_n^{all\_high})$, stage III is $t \in [t_n^{all\_high}, t_{n+1}^h)$, and stage IV is $t \in [t_{n+1}^h, t_{n+1}^{all})$.

Proposition 6. *The numbers of inexperienced and experienced workers in the economy change in the following fashion when the economy undergoes structural change from industry n to industry $n+1$ ($n \geq 2$):*

- *at stage I, the number of inexperienced workers in industry n decreases, and the number of experienced workers in industry n increases because of on-the-job learning;*
- *at stage II, the number of inexperienced workers in industry n decreases, while experienced workers in industry n increases; the number of both experienced and inexperienced workers in industry $n+1$ increases because of on-the-job learning and labor reallocation across industries;*
- *at stage III, there are no inexperienced workers working or looking for jobs in industry n, and the number of experienced workers stays constant in industry n; the number of inexperienced workers in industry $n+1$ decreases and that of experienced workers in industry $n+1$ increases because of on-the-job learning;*
- *at stage IV, the number of experienced workers in industry n decreases, but the number of*

rienced or experienced, employed or unemployed, leave for the labor market of industry $n+1$. Both of these two phases last for some period so that the relative rental price of capital changes continuously, otherwise the relative rental price of capital would exhibit discontinuity as the marginal cost of consumption in terms of capital in the consumption-good sector exhibits discontinuity when the underlying industrial composition shifts between single industry and two neighboring industries. See Appendix B.6 for details.

[9] The model result that inexperienced workers leave a sunset industry earlier than experienced ones during industrial upgrading is consistent with data. Using PSID data, we find that workers who change sectors have an average working experience of 2.74 years in the old sector, significantly smaller than 4.10, the average experience of workers who stay. See Appendix A for details.

experienced and that of inexperienced workers in industry $n+1$ both increase because of on-the-job learning and labor reallocation across industries.

Proof. *Please see Appendix B.6.* ■

Figure 3.3 shows the cyclical pattern in the numbers of inexperienced and experienced labor force (employed plus unemployed), in the labor market for each industry. [10] We analytically prove that both the size of the experienced labor force in the labor market for industry $n+1$ (denoted by L^h_{n+1}) and the size of the inexperienced counterpart (denoted by L^l_{n+1}) are increasing over time when $t \in [t^l_{n+1}, t^{all\_high}_n)$, because inexperienced workers in the labor market for industry n move into that for industry $n+1$ and some of them become experienced. When $t \in [t^{all\_high}_n, t^h_{n+1})$, no workers move across labor markets for different industries and L^h_{n+1} increases, but L^l_{n+1} decreases as some inexperienced workers become experienced with the total number of labor force (employed plus unemployed) for industry $n+1$ remaining unchanged. When $t \in [t^h_{n+1}, t^{all}_{n+1})$, both L^h_{n+1} and L^l_{n+1} increase over time, because experienced workers in industry n move into industry $n+1$ and immediately become inexperienced, but some of them get employed and become experienced in industry $n+1$. When $t \in [t^l_{n+2}, t^{all\_high}_{n+1})$, industry n has exited and inexperienced workers in industry $n+1$ either move into industry $n+2$ or stay and become experienced in $n+1$. So L^l_{n+1} decreases over time and L^h_{n+1} increases over time at this stage. When $t \in [t^h_{n+2}, t^{all}_{n+2})$, experienced workers in industry $n+1$ move into industry $n+2$, and industry $n+1$ gradually declines and eventually disappears at time t^{all}_{n+2}. As we see from figure 3.3 and figures below, the industrial upgrading from industry 1 to industry 2 is slightly different from those from industry n to industry $n+1$ ($n \geq 2$). It is because we assume all workers face the same job finding rate f in industry 1 so experience types are irrelevant but the job finding rates are different for workers of different experience types in any industry n for $n \geq 2$.

[10] We set the coefficient of relative risk aversion $1/\sigma = 1$, discount rate $\rho = 0.9$, the capital intensity coefficient $a = 3.5$, the productivity coefficient $\lambda = 2.0$, the capital-goods production efficiency $A = 1$, two unemployment rates $\frac{\widehat{f}}{\widehat{f}+\widehat{\delta}} = \frac{1}{3}$ and $\frac{\pi\widehat{f}}{\pi\widehat{f}+\widehat{\delta}} = \frac{2}{5}$, and the learning rate $\xi = 0.2$. It is mainly for illustration purposes instead of attempting to quantitatively match the real-life data.

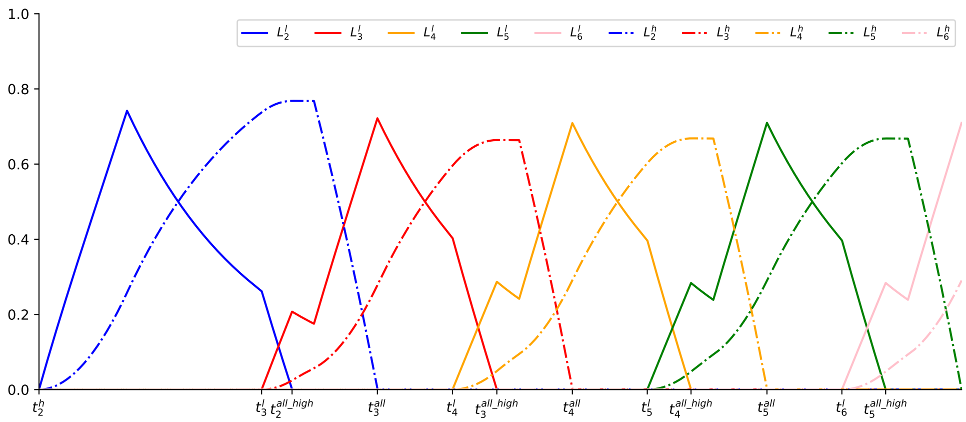


Figure 3.3: Inexperienced and experienced workers in each industry

Figures 3.4 shows the dynamics of employment and unemployment within each industry. The employment of an industry reaches its peak when all labor force is in the labor market for that industry. Figure 3.5 shows how the unemployment rate in the labor market for industry n changes over time during the whole life span of this industry. Recall t_n^l is the time point when industry n just emerges, and the unemployment rate in industry n is equal to $\frac{\hat{\delta}}{\pi\hat{f}+\hat{\delta}}$. From t_n^l to $t_{n-1}^{allhigh}$ (i.e., stage II), inexperienced workers in industry $n-1$ gradually move into industry n, and the industrial unemployment rate in industry n declines as the share of inexperienced workers gradually decreases via on-the-job learning. At stage III, when there is no industrial upgrading, the unemployment rate in industry n keeps declining due to on-the-job learning. At stage IV, when experienced workers in industry $n-1$ move into n, the unemployment rate can either decline or increase, depending on the duration of stage III. If it is short (e.g., when π is close to 1), then the share of inexperienced workers in industry n remains large, and the industrial unemployment rate declines over time. If stage III lasts long (e.g., when π is small), then the share of inexperienced workers is small, so the reallocation process in stage IV drives up the share of inexperienced workers in industry n, and the unemployment rate for industry n increases. At time point t_n^{all}, all labor force of the economy is the in the labor market for industry n. From that point to t_{n+1}^l the industrial unemployment rate keeps declining, as inexperienced workers in industry n either become experienced or move into $n+1$.

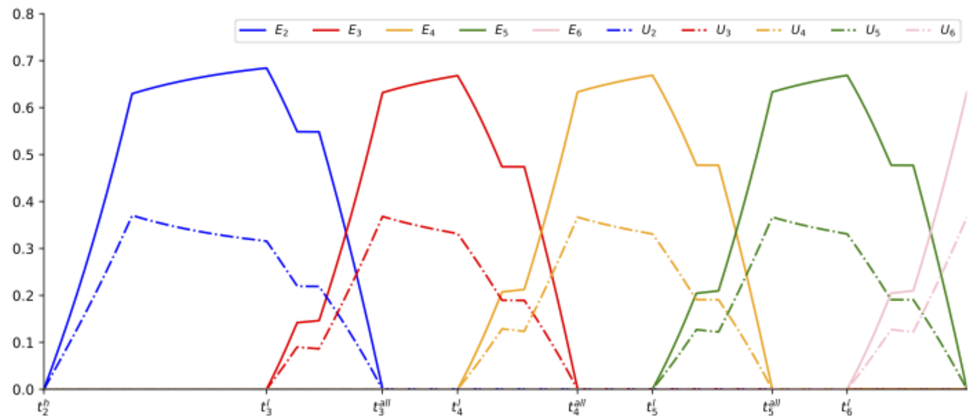


Figure 3.4: Employed and unemployed workers in each industry

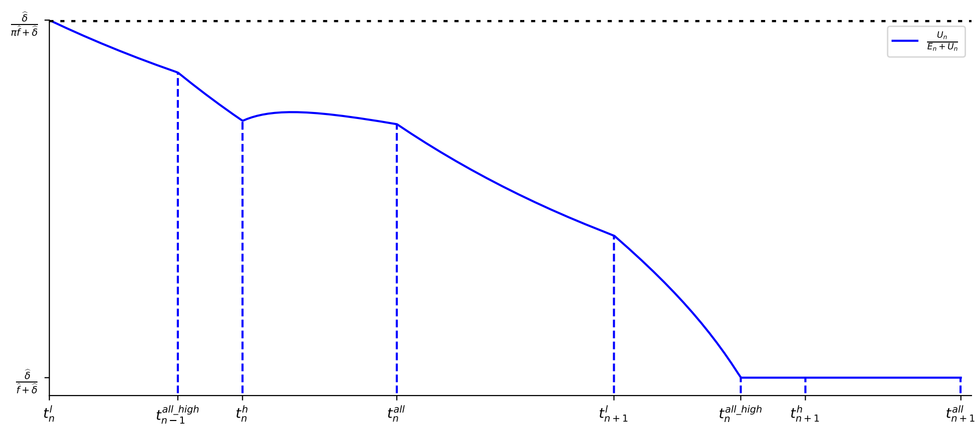


Figure 3.5: Industrial unemployment rate for labor market of industry n

Before moving to present other results of the model, we briefly discuss the empirical relevance of the analysis above regarding sectoral employment and unemployment during industrial upgrading. The theory implies that when a new sector emerges and expands, new and inexperienced workers enter into this sector. The job finding rate is low and unemployment rate is high. The opposite holds for declining sectors. Put it differently, the model implies a positive correlation between sectoral employment change and sectoral unemployment rate, and a negative correlation between sectoral employment change and sectoral job finding rate. We collect data from Current Population Survey and Bureau of Labor Statistics and show in Appendix A that both correlations are consistent with that in data.

We now analyze the model implications on industry life span. As seen in Figure 3.4 the life span of industry $n, n \geq 2$, converges. This is due to the symmetric structure of industrial upgrading. The life span of industry n equals to $t^{all}_{n+2} - t^{l}_{n+1}$, the length between the time point when the first inexperienced worker appears in that industry, t^{l}_{n+1}, and the time point when the last experienced worker leaves that industry, t^{all}_{n+2}. When ξ is small, we have the following result:

Proposition 7. *If ξ converges to 0, the life span of an industry in the long run is*

$$\lim_{\substack{\xi \to 0 \\ n \to \infty}} (t^{all}_{n+2} - t^{l}_{n+1}) = \frac{2\log(\lambda)}{g_c} - \frac{\log\left(\frac{a(a-1)}{\lambda(\lambda-1)} \frac{\lambda^{\frac{\pi\hat{f}}{\pi\hat{f}+\hat{\delta}} - \frac{\hat{f}}{\hat{f}+\hat{\delta}}}}{a^{\frac{\pi\hat{f}}{\pi\hat{f}+\hat{\delta}} - \frac{\hat{f}}{\hat{f}+\hat{\delta}}}}\right)}{A - \rho}, \tag{26}$$

which implies that more serious mismatch (smaller π) results in a longer industry life span.

Proof. *Please see Appendix B.6.* ∎

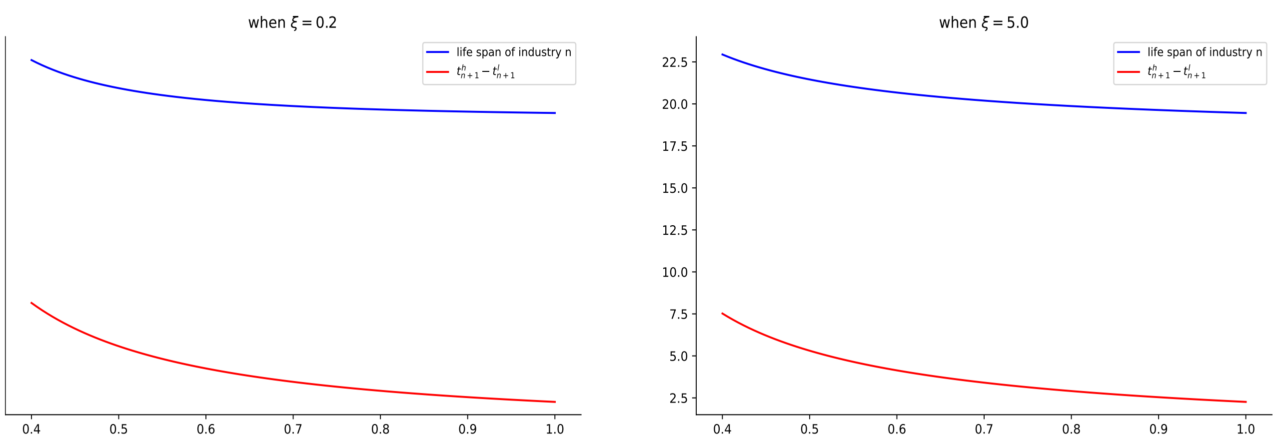


Figure 3.6: How does the life span of industries change with the degree of mismatch

Fixing other parameters except for the mismatch π as our benchmark calibration, Figure 3.6 shows how the life span of industry n changes with the degree of mismatch. We also add the graph for ξ equals to 5.0 where workers have a much faster on-the-job learning rate. Our numerical exercise suggests that even for a large learning rate, the industries will exist for a longer time when the mismatch becomes more serious. The reason why the industry has a larger life span is that mismatch makes experienced workers in the old industries reluctant to move into the new industries as they have a lower job finding rate in these industries. They are stuck in the old industries. The red line in Figure 3.6 computes the difference of the upgrading time for two types of workers $t^h_{n+1} - t^l_{n+1}$, which is also the difference of the time when the industry $n+1$ emerges and when the first experienced workers in industry n move into industry $n+1$. Larger $t^h_{n+1} - t^l_{n+1}$ indicates that experienced workers in the old industries wait longer before they move into the new industries compared with the inexperienced ones, and it consequently extends the life cycle of industries.

We next investigates the dynamics of employment and unemployment. The time path of aggregate unemployment satisfies:

(1) if $t^{all}_n \leq t < t^h_{n+1}$,

$$U(t) = (\frac{\widehat{\delta}}{\pi\widehat{f}+\widehat{\delta}} - \frac{\widehat{\delta}}{\widehat{f}+\widehat{\delta}})e^{-\frac{\pi\widehat{f}}{\pi\widehat{f}+\widehat{\delta}}\xi(t-t^{all}_n)}L^l_n(t^{all}_n) + \frac{\widehat{\delta}}{\widehat{f}+\widehat{\delta}}L \quad (27)$$

(2) if $t^h_{n+1} \leq t < t^{all}_{n+1}$

$$U(t) = \frac{1-\pi}{\pi}\frac{\widehat{\delta}}{\widehat{f}+\widehat{\delta}}[\frac{Q_n g_c}{N_2\xi+g_c}e^{g_c(t-t^h_{n+1})} - (\frac{Q_n g_c}{N_2\xi+g_c} - \frac{\pi\widehat{f}}{\pi\widehat{f}+\widehat{\delta}}L^l_{n+1}(t^h_{n+1}))e^{-N_2\xi(t-t^h_{n+1})}] + \frac{\widehat{\delta}}{\widehat{f}+\widehat{\delta}}L \quad (28)$$

As shown in Appendix B.6, aggregate unemployment decreases during the period $t \in [t^{all}_n, t^h_{n+1})$. For $t \in [t^{all}_n, t^l_{n+1}) \cup [t^{all\_high}_n, t < t^h_{n+1})$, there is no industrial upgrading, and the aggregate unemployment rate declines as inexperienced workers in industry n or in both industries n and $n+1$ gradually become experienced ones via on-the-job learning. For $t \in [t^l_{n+1}, t^{all\_high}_n)$, inexperienced workers in industry n move into $n+1$, and since they are inexperienced in both industries, this reallocation of workers does not increase the mismatch. As a result, the aggregate unemployment keeps declining as inexperienced workers in industries n and $n+1$ continue to become experienced at this stage. The aggregate unemployment rate rises for $t \in [t^h_{n+1}, t^{all}_{n+1})$ as experienced workers in industry n move into $n+1$ and they become inexperienced which causes more mismatch. It increases and reaches its peak at t^{all}_{n+1}, the first time when all workers stay all industry $n+1$. After that, aggregate unemployment declines again in the period $t \in [t^{all}_{n+1}, t^h_{n+2})$, so on and so forth. As a result, the aggregate unemployment rate exhibits a cyclical pattern along with industrial upgrading in the aggregate growth path. Note the unemployment rate cycle is not caused by any exogenous aggregate shocks, but rather by the cyclicality in the number of inexperienced workers resulting from two processes: the experience-to-inexperience transition associated with the labor reallocation process, and the inexperience-to-experience transition through on-the-job learning. Formally, we have the following proposition regarding the cyclical pattern of unemployment rate:

Proposition 8. *The aggregate unemployment rate in the infinite-industry model exhibits a recurrent cyclical pattern: it increases when experienced workers in an old industry move into the new industry and decreases at all the other stages.*

Proof. Please see Appendix B.6. ■

Figure 3.7 illustrates this cyclical pattern of aggregate unemployment rate.

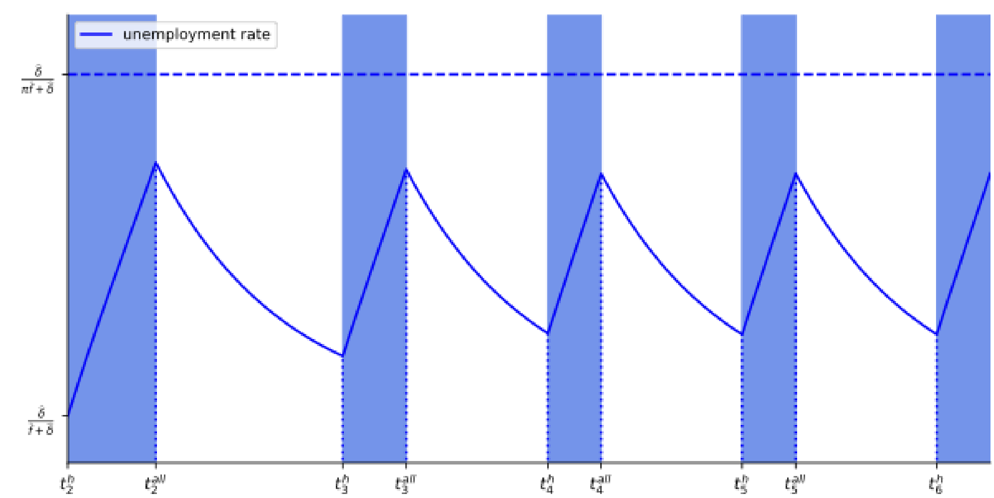


Figure 3.7: Cyclical hump-shaped pattern of aggregate unemployment rate

Next we investigate the impact on aggregate unemployment of changes in parameters. Particularly, we are interested in the following three parameters: the degree of mismatch π, the on-the-job learning rate ξ, and the capital-goods production efficiency A. The next proposition summarizes the comparative static results.

Proposition 9. *When the labor market mismatch becomes severer, or the on-the-job learning rate decreases, or the capital-goods production efficiency increases, both the peak and bottom values of the aggregate unemployment rate become larger.*

Proof. *Please see Appendix B.6.* ∎

This proposition can be more intuitively seen in Figure 3.8. When mismatch becomes severer, the aggregate unemployment rate increases faster and reaches larger peak and bottom values within each industrial upgrading as inexperienced workers has lower job finding rate, and have less chance to be employed to become experienced ones. When the learning rate increases, employed inexperienced workers quickly become experienced and their job finding rate increases. So the aggregate unemployment rate shifts downwards. When the capital-goods production efficiency increases, industrial upgrading takes place more rapidly which results in larger reallocation and leads to greater aggregate unemployment rate. Figure 3.8 illustrates how changes in mismatch π, learning efficiency ξ and the capital-goods production efficiency A affect the aggregate unemployment rate.

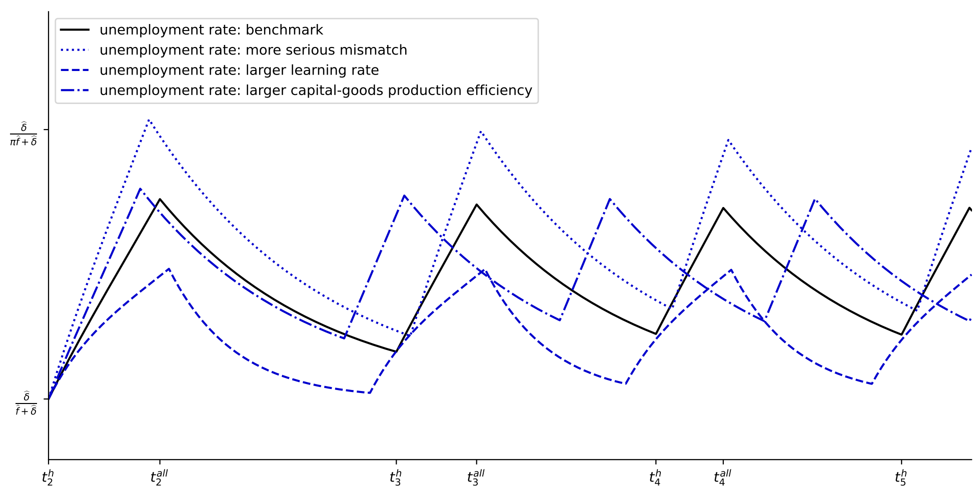


Figure 3.8: The comparative statics of the aggregate unemployment rate and consumption

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 the 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 undertakes structural changes driven by endogenous capital accumulation. The aggregate unemployment rate exhibits a hump-shaped pattern: it rises at the beginning when experienced workers in a sunset industry move into a sunrise industry and suffers from mismatch in the new industry. The unemployment rate declines later on when inexperienced workers become experienced through on-the-job learning in the new industry. The unemployment rate goes up again as the current new industry gradually declines and is eventually replaced by an even more capital-intensive industry, ad infinitum.

We find that there exist three critical forces: capital-goods production efficiency, on-the-job learning efficiency and labor market mismatch. When the capital-goods production efficiency is larger, the consumption growth rate is higher, the speed of old industries being replaced by new ones faster, industry life span shorter, and the aggregate unemployment rate universally higher. When learning efficiency increases, the aggregate unemployment rate shifts downward and the life span of an industry becomes shorter. In addition, severer mismatch drives up the aggregate unemployment rate and extends the life span of an industry. We made necessary simplifications to analytically characterize the impact of various economic forces on industrial dynamics and aggregate unemployment rate, view a quantitative evaluation of these forces with an extended model as a natural extension of current work, and leave it to future research.

References

Acemoglu, Daron and Veronica Guerrieri, "Capital Deepening and Nonbalanced Economic Growth," *Journal of Political Economy*, 2008, *116*, 467–4.

Alvarez, Fernando and Robert Shimer, "Search and Rest Unemployment," *Econometrica*, 2011, *79* (1), 75–122.

Boldrin, Michele, David K. Levine, Yong Wang, and Lijun Zhu, "A Theory of the Dynamics of Factor Shares," *Working Paper*, 2022.

Boppart, Timo, "Structural Change and the Kaldor Facts in a Growth Model with Relative Price Effects and Non-Gorman Preferences," *Econometrica*, 2014, *82* (6), 2167–2196.

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

Chenery, Hollis B., "Patterns of Industrial Growth," *American Economic Review*, 1960, *50* (4), 624–654.

Comin, Diego, Danial Lashkari, and Mart Mestieri, "Structural Change With Long-Run Income and Price Effects," *Econometrica*, 2021, *89* (1), 311–374.

Herrendorf, Berthold, Richard Rogerson, and Akos Valentinyi, "Growth and Structural Transformation," *Handbook of Economic Growth*, 2014, *2*, 855–941.

Herz, Benedikt and Thijs van Rens, "Accounting for Mismatch Unemployment," *Journal of the European Economic Association*, 2019, *18* (4), 1619–1654.

Ju, Jiandong, Justin Yifu Lin, and Yong Wang, "Endowment structures, industrial dynamics, and economic growth," *Journal of Monetary Economics*, 2015, *76*, 244–263.

Kongsamut, Piyabha, Sergio Rebelo, and Danyang Xie, "Beyond Balanced Growth," *Review of Economic Studies*, 2001, *68* (4), 869–882.

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

Lilien, David M., "Sectoral Shifts and Cyclical Unemployment," *The Journal of Political Economy*, 1982, *90* (4), 777–793.

Lucas, Robert E. and Edward C. Prescott, "Sectoral Shifts and Cyclical Unemployment," *Journal of Economic Theory*, 1974, *7*, 188–209.

Mortensen, Dale T. and Christopher A. Pissarides, "Job Creation and Job Destruction in the Theory of Unemployment," *Review Economic Studies*, 1994, *61* (3), 397–415.

Ngai, L. Rachel and Christopher Pissarides, "Structural Change in a Multi-Sector Model of Growth," *American Economic Review*, 2007, *97*, 429–443.

_ **and** _ , "Trends in Hours and Economic Growth," *Review of Economic Dynamics*, 2008, *11*, 239–256.

Pissarides, Christopher, "Unemployment and Hours of Work: The North Atlantic Divide Revisited," *International Economic Review*, 2007, *48*, 1–36.

Restrepo, Pascual, "Skill Mismatch and Structural Unemployment," *Working paper*, 2015.

Sahin, Aysegul, Joseph Song, Giorgio Topa, and Giovanni L. Violante, "Mismatch Unemployment," *American Economic Review*, 2014, *104* (11), 3529–64.

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

_ , "Mismatch," *American Economic Review*, 2007, *97* (4), 1074–1101.

Zagler, Martin, "Economic Growth, Structural Change, and Search Unemployment," *Journal of Economics*, 2009, *96* (1), 63–78.

Online Appendix of "Labor Market Mismatch, Structural Unemployment and Industry Dynamics"

Appendix A Data and Facts

Cross time correlation between unemployment rate and sector reallocation rate Following Lilien (1982), we measure the cross-sector employment reallocation rate the standard deviation of sector employment growth rate. In particular, denote e_{it} employment of sector i in year t, and $g_{it} \equiv \frac{e_{it}-e_{it-1}}{e_{it-1}}$ the growth rate of sector $i's$ employment from year $t-1$ to t. The sector reallocation rate is defined as

$$Reallocation_t \equiv \{\sum_i \omega_i (g_{it} - \sum_i \omega_i g_{it})^2\}^{1/2}$$

where ω_i is the weight for sector i. In the baseline, we choose sector $i's$ employment share in period t as the weight, i.e. $\omega_i = \frac{e_{it}}{\sum_i e_{it}}$. As a robustness check, we further use the employment share in the previous period as weight, $\omega_i = \frac{e_{it-1}}{\sum_i e_{it-1}}$, and calculate the unweighted standard deviation, equivalently $\omega_i = 1$.[11] Figure 1.1 plots the such measured sector reallocation rate.[12]

To formally test the correlation between sector reallocation rate, $Reallocation_t$, and unemployment rate, $UNRATE_t$, we use the time series data from 1955-2020 and run the following regressions

$$UNRATE_t = \beta_0 + \beta_1 * Reallocation_t + Ctrls + \epsilon_t$$

Two controls are added in the regression: a recession dummy and a dummy for the post-2001 period to account for the fact that the sector code in the original data switches from SIC to NAICS after 2001. Table A.1 presents the regression results under differ-

[11] The sector employment data is from the National Income and Product Accounts provided by U.S. Bureau of Economic Analysis. The tables used are "Table 6.8 Persons Engaged in Production by Industry", which contain employment for 63 3-digit SIC sectors from 1948-2000 and for 70 3-digit NAICS sectors from 1998-2020.

[12] The Bureau of labor statistics provides sectoral employment and unemployment from 2002-2021, as detailed below. We obtain sectoral labor force by adding these two items, and calculate the reallocation rate using sectoral labor force instead of employment. The correlation coefficient between labor force and employment based reallocation rate is as high as 0.9984.

ent settings. The coefficient for *reallocation* is positive and significant under the baseline measurement. An increase of 1 percentage point in sector reallocation rate leads to 0.55 percentage point increase in the aggregate unemployment rate. The sign and significance hold if the recession dummy (and the post-2001 dummy) is added, suggesting that the positive correlation between sector reallocation rate and unemployment rate are not driven by recession points only. Further, the baseline result is robust to alternative measures of sector reallocation rate, as seen in regression results (2) and (3) in Table A.1.

Table A.1: Correlation between Unemployment Rate and Cross-Sector Employment Reallocation Rate

	(1) wgt.=emp_t		(2) wgt.=emp_{t-1}		(3) unweighted	
Reallocation	0.55***	0.59**	0.53***	0.57***	0.30**	0.26*
	(0.19)	(0.22)	(0.19)	(0.21)	(0.13)	(0.13)
Recession D	N	Y	N	Y	N	Y
Post2001 D	N	Y	N	Y	N	Y
R^2	0.11	0.14	0.11	0.14	0.08	0.10
Obs.	66	66	66	66	66	66

Note: This table show the results of regressing unemployment rate on cross-sector employment reallocation rate from 1955-2020.

Sector Reallocation and Worker Experience Sector reallocation in the aggregate is composed of workers' sector movement at the individual level. An average worker typically stays in a sector for finite years. For example, in 2013, the average years a worker has worked in their 2013 3-digit sector is 10 years.[13] Experience is an important factor in determining if a worker switches sectors. Particularly, less experienced workers are more likely to move to a new sector, as shown in Table A.2. From 1999-2019, workers who changed sectors from sector A to sector B in year $(t, t+2]$ have, in average, worked in sector A for 2.74 years until year t, while those who did not have

[13]For all workers in a given sector in PSID 2013, we calculate the average years these workers stay in their 2013 3-digit sector, and then average them into the sector level. Figure A.1 plots the distribution of this average working time across all 3-digit sectors.

an average working experience of 4.10 years, which is about 50% longer than the first group. This is consistent with the model result that inexperienced workers move out of an old industry earlier than experienced ones during industrial upgrading.

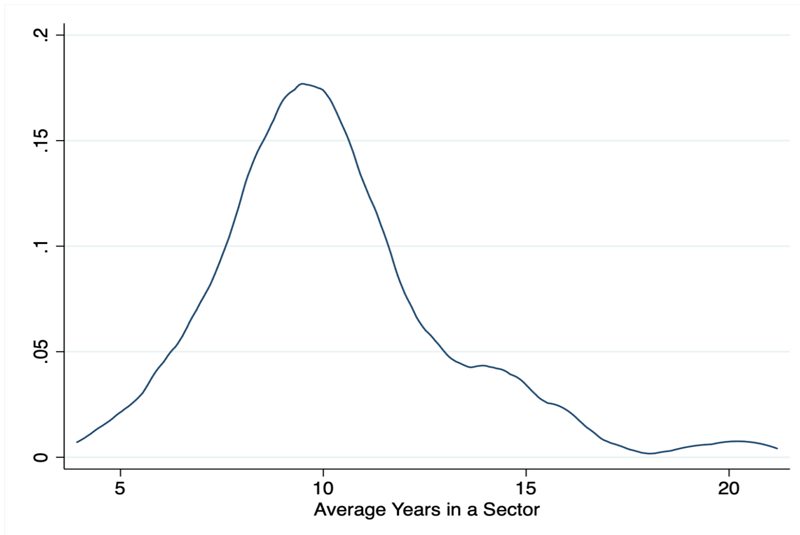


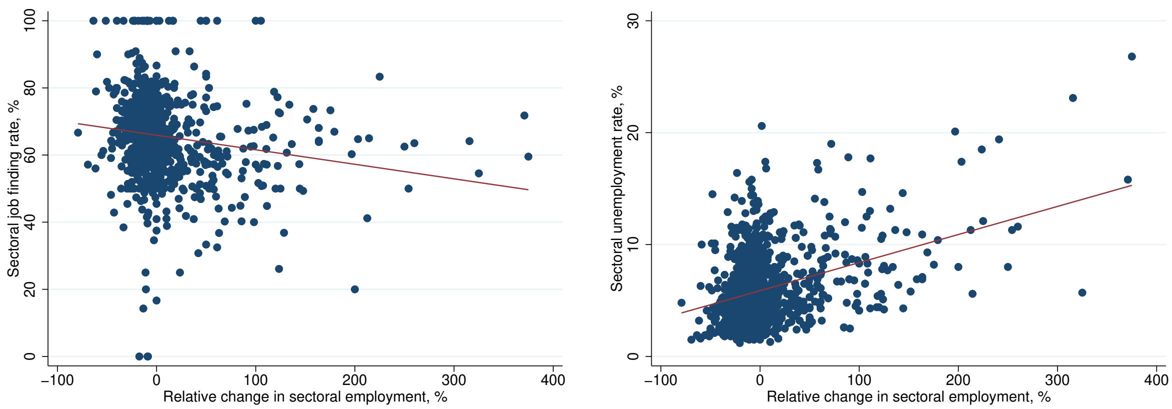
Figure A.1: Distribution of Average Years a Worker Stays in a Sector

Table A.2: Average Working Experience

	Working experience until year t
Workers who change sectors in years (t,t+2]	2.74
Workers who do not change sectors in year (t,t+2]	4.10

Data Source: PSID 1999-2019. This table shows the average working years until year t in a sector between those who changed to a different sector in year $(t, t+2]$ and those who did not.

Sectoral Employment Change, Job Finding Rate and Unemployment Rate The model implies a positive correlation between sectoral employment change rate (EC) and sectoral unemployment rate (UN), and a negative correlation between sectoral employment change and sectoral job finding rate (JF). We now examine if these results are supported by data. The U.S. Bureau of labor statistics provides unemployment rate for 55 3-digit sectors from 2003-2021, with unemployment in a sector i and year t referring to those who are unemployed in year t and whose last job was in sector i. We define employment change rate (EC) of sector i in year t as the relative change of that sector's employment from year $t-1$ to t. Job finding rate (JF) in sector i and year t is defined as the ratio of *"number of persons who is unemployed in sector i and year t, and employed in year* $t+1$" to *"number of persons who is unemployed in sector i and year t"*.[14] Figure A.2 plots the such measured correlations across the 55 sectors from 2003-2021. Consistent with the model implication, at the sector level, job finding rate and employment change rate show a clear negative correlation, while unemployment rate and employment change rate show a positive correlation.



(a) Job Finding Rate

(b) Unemployment Rate

Figure A.2: Correlation between Sectoral Employment Change and Sectoral Job Finding Rate & Unemployment Rate

The figure plots the unemployment rate (vertical) *v.s.* the relative change in employment across 55 sectors annually from 2003-2021. Unemployment in a sector refers to the unemployed whose last job was in the given sector.

More formally, we run the following regressions

$$Y_{i,t} = \beta_0 + \beta_1 * EC_{i,t} + industry + year + \epsilon_{i,t}, \quad Y \in \{JF, UN\}$$

Table A.3 shows the regression results and confirms that the correlations in Figure A.2 are robust to controlling for sector and time fixed effects.

[14]The data source for calculating sectoral job finding rate is Current Population Survey (CPS). We manually match the CPS 4-digit industry code to BLS sectors based on sector names.

Table A.3: Correlation between Job Finding Rate, Unemployment Rate and Sector Employment Change

	Dep. var: JF_{it}		Dep. var: UN_{it}	
	(1)	(2)	(1)	(2)
EC_{it}	-0.043^{***}	-0.027^{**}	0.025^{***}	0.026^{***}
	(0.009)	(0.012)	(0.002)	(0.001)
Sector FE	N	Y	N	Y
Year FE	N	Y	N	Y
R^2	0.03	0.32	0.14	0.84
Obs.	863	863	1086	1086

Note: See the text above for definitions of dependent and independent variables. The sample covers 55 3-digit sectors annually from 2003 to 2021.

Appendix B Derivation and Proofs

B.1 Proof of proposition 1

Planner's Problem: We start by setting the Hamiltonian for the planner's problem

$$\begin{aligned} H = e^{-\rho t}\frac{C^{1-\sigma}-1}{1-\sigma} + \mu_k[AK - \Omega(C, E_0^l + E_0^h, E_1^l + E_1^h, E_2^l + E_2^h)] + \sum_{i=l,h} \mu_0^i[fU_i(1-\lambda_1^i - \lambda_2^i) - \delta E_0^i] \\ + \sum_{i=l,h} \mu_1^i(fU_i\lambda_1^i - \delta E_1^i) + \mu_2^l(\pi f U_l \lambda_2^l - \delta E_2^l - \xi E_2^l) + \mu_2^h(fU_h\lambda_2^h - \delta E_2^h + \xi E_2^l) \\ + \mu_u^l[-fU_l(1-\lambda_2^l) - \pi f U_l \lambda_2^l + \delta(E_0^l + E_1^l + E_2^l)] + \mu_u^h[-fU_h + \delta(E_0^h + E_1^h + E_2^h)] \quad (29) \end{aligned}$$

where μ_k, μ_j^i, and μ_u^i are the multipliers associated with capital constraint, associated with $i-$experience-typed employees in industry j, and associated with unemployment of $i$experience-typed workers, respectively. The first-order conditions are

$$\dot{\mu}_k = -A\mu_k \tag{30}$$

$$-\dot{\mu}_0^i = -\delta\mu_0^i + \delta\mu_u^i - \mu_k \frac{\partial\Omega}{\partial E_0^i}, i \in \{l \;\;,\;\; h\} \tag{31}$$

$$-\dot{\mu}_1^i = -\delta\mu_1^i + \delta\mu_u^i - \mu_k \frac{\partial\Omega}{\partial E_1^i}, i \in \{l \;\;,\;\; h\} \tag{32}$$

$$-\dot{\mu}_2^l = -(\delta + \xi)\mu_2^l + \delta\mu_u^l + \xi\mu_2^h - \mu_k \frac{\partial\Omega}{\partial E_2^l}, \tag{33}$$

$$-\dot{\mu}_2^h = -\delta\mu_2^h + \delta\mu_u^h - \mu_k \frac{\partial\Omega}{\partial E_2^h}, \tag{34}$$

$$-\dot{\mu}_u^l = f(1 - \lambda_1^l - \lambda_2^l)(\mu_0^l - \mu_u^l) + f\lambda_1^l(\mu_1^l - \mu_u^l) + \pi f\lambda_2^l(\mu_2^l - \mu_u^l), \tag{35}$$

$$-\dot{\mu}_u^h = f(1 - \lambda_1^h - \lambda_2^h)(\mu_0^h - \mu_u^h) + f\lambda_1^h(\mu_1^h - \mu_u^h) + f\lambda_2^h(\mu_2^h - \mu_u^h), \tag{36}$$

$$e^{-\rho t}C^{-\sigma} = \mu_k \frac{\partial\Omega}{\partial C}, \tag{37}$$

$$\frac{\partial H}{\partial\lambda_1^i} = fU_i(\mu_1^i - \mu_0^i), i \in \{l \;\;,\;\; h\} \tag{38}$$

$$\frac{\partial H}{\partial\lambda_2^l} = fU_l[\pi(\mu_2^l - \mu_u^l) - (\mu_0^l - \mu_u^l)], \tag{39}$$

$$\frac{\partial H}{\partial\lambda_2^h} = fU_h(\mu_2^h - \mu_0^h). \tag{40}$$

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 household members 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 good in units of the capital good. The household can sell one consumption good to receive $q(t)$ units of capital good. Employed workers of experience type $i, i = \{h, l\}$, in industry j, receive wage payment $q(t)w_j^i(t)$.

The representative household's problem is

$$\max \int_{t=0}^{\infty} e^{-\rho t} \frac{C^{1-\sigma} - 1}{1-\sigma} dt \tag{41}$$

subject to the budget constraint

$$\dot{K} = rK + q\Big(\sum_{i \in \{l,h\}, j \in \{0,1,2\}} w_j^i E_j^i - C\Big). \tag{42}$$

where E_j^i is again the measure of employed i-experience-typed workers in industry j. Given equilibrium prices $\{r(t), q(t), w_j^i(t)\}$, the household chooses consumption $C(t)$ and the fraction of unemployed workers searching in different industries λ_j^i. The evolution of employed and unemployed workers is the same as Equations (6)-(13).

Firms in the capital goods sector borrow capital at the rental price $r(t)$, and access to a linear production technology, $AK(t)$. Firms producing the final consumption good use intermediate goods as inputs, 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 the final consumption good is

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

Firms in industry j produce intermediate goods with a constant return to scale technology, $F_j(k,l)$. They post vacancies at zero cost. Firm matched with workers purchase capital from households and make wage payment, $q_j(t) w_j^i(t)$, to employed workers. Their profit function, $\pi_j(k,l)$, is

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

We check that when the market prices satisfy $q = q_j = \frac{\partial \Omega}{\partial C}, \forall j \in \{0,1,2,3\}$, $w_j^i = -\frac{\partial \Omega}{\partial E_j^i} / \frac{\partial \Omega}{\partial C}, \forall i \in \{l,h\}, \forall j \in \{0,1,2\}$ and $r = A$, we then have an equilibrium. For firms in the capital good sector, $A = r$ indicates zero profit. For firms producing final the consumption good, $q = q_j$ follows from the zero profit condition as well. For firms producing intermediate goods j, combining with $E(.)$ and expressions of q and w_j^i, we have the following inequality

$$q \leq qw_0^i;\ q\lambda \leq qw_1^i + a;\ q\lambda^2 \leq qw_2^i + a^2;\ q\lambda^3 \leq a^3, i \in \{l,h\}. \tag{43}$$

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

B.2 Proof of proposition 2

For an instantaneous equilibrium, the employed and unemployed $i-$type workers in each industry are only a fraction of total $i-$type workers in that industry, with these fractions depending on the relative job finding and separation rate. Labor market clearing implies

$$E_0^l \frac{\widehat{f} + \widehat{\delta}}{\widehat{f}} + E_1^l \frac{\widehat{f} + \widehat{\delta}}{\widehat{f}} + E_2^l \frac{\pi\widehat{f} + \widehat{\delta}}{\pi\widehat{f}} \leq L^l, \quad (E_0^h + E_1^h + E_2^h)\frac{\widehat{f} + \widehat{\delta}}{\widehat{f}} \leq L^h. \tag{44}$$

For planner's problem, there should be no waste of employed workers in order to maximize the final good consumption. In other words, any employed worker has the required amount of capital flow to produce intermediate goods in that industry. If not, those without enough capital should move into industry 0, which increases the final output. Thus the required capital inflow at each instant t is

$$\bar{\Omega}(E_0^h + E_0^l, E_1^h + E_1^l, E_2^h + E_2^l, K_3) = \sum_{i=1}^{2} a^i (E_i^h + E_i^l) + K_3$$

And the consumption of final good is

$$\bar{C}(E_0^h + E_0^l, E_1^h + E_1^l, E_2^h + E_2^l, K_3) = \sum_{i=0}^{2} \lambda^i (E_i^h + E_i^l) + \frac{\lambda^3}{a^3} K_3$$

The Hamiltonian of planner's problem can then be rewritten as

$$H = e^{-\rho t}\frac{C^{1-\sigma}-1}{1-\sigma} + \mu_k[AK - \bar{\Omega}(E_0^h + E_0^l, E_1^h + E_1^l, E_2^h + E_2^l, K_3)]$$
$$+\mu_c[\bar{C}(E_0^h + E_0^l, E_1^h + E_1^l, E_2^h + E_2^l, K_3) - C] + \mu_l(L^l - (E_0^l\frac{\widehat{f}+\widehat{\delta}}{\widehat{f}} + E_1^l\frac{\widehat{f}+\widehat{\delta}}{\widehat{f}} + E_2^l\frac{\pi\widehat{f}+\widehat{\delta}}{\pi\widehat{f}}))$$
$$+\mu_h(L^h - (E_0^h + E_1^h + E_2^h)\frac{\widehat{f}+\widehat{\delta}}{\widehat{f}}) + \mu_e(\xi E_2^l) \quad (45)$$

We have three state variables $\{K, L^h, L^l\}$ and eight control variables $\{C, E_j^i, K_3\}$, $i \in \{h,l\}$, $j \in \{0,1,2\}$. The last term of Hamiltonian is related to the equation of experience accumulation $\dot{L^h} = \xi E_2^l$. All multipliers are positive. Let $C^*(t)$, $E_j^{i*}(t)$ and $K_3^*(t)$ denote the optimal consumption level, the optimal level of employment of $i-$experience-typed workers in industry j , and the optimal level of capital inflow in industry 3, at time t. If $E_2^{l*}(t) = 0$, the last term ξE_2^l is zero, and the maximization of H requires the planner to minimize the capital inflow $\bar{\Omega}(E_0^h + E_0^l, E_1^h + E_1^l, E_2^h + E_2^l, K_3)$ for a given path $C^*(t)$ with the market clearing conditions (44) hold.

If $E_2^{l*}(t) > 0$, the first order conditions of $E_i^l(t), i \in \{0,1,2\}$ become

$$\mu_c - \frac{\widehat{f}+\widehat{\delta}}{\widehat{f}}\mu_l \leq 0, \lambda\mu_c - a\mu_k - \frac{\widehat{f}+\widehat{\delta}}{\widehat{f}}\mu_l \leq 0, \lambda^2\mu_c - a^2\mu_k - \frac{\pi\widehat{f}+\widehat{\delta}}{\pi\widehat{f}}\mu_l + \xi\mu_e = 0 \quad (46)$$

Notice that μ_h and μ_l indicate the value of experienced and inexperienced workers. Then we must have $\mu_h > \mu_l$ [15], and it implies

$$\frac{\partial H}{\partial E_0^h} = \mu_c - \frac{\widehat{f}+\widehat{\delta}}{\widehat{f}}\mu_h < 0, \frac{\partial H}{\partial E_1^h} = \lambda\mu_c - a\mu_k - \frac{\widehat{f}+\widehat{\delta}}{\widehat{f}}\mu_h < 0$$

It follows that $E_0^{h*}(t) = E_1^{h*}(t) = 0$ and $E_2^{h*}(t) = \frac{\widehat{f}}{\widehat{f}+\widehat{\delta}}L^h$. Since inexperienced workers have a lower fraction of employment in industry 2, all experienced workers are first

[15] The f.o.c of C in (45) implies that $\mu_c = e^{-\rho t}C^{-\sigma} > 0$. The first term in (46) implies $\mu_l \geq \frac{\widehat{f}}{\widehat{f}+\widehat{\delta}}\mu_c > 0$. When ξ is small, from the last term of (46) and $\mu_h \leq \mu_l$, we will have $\frac{\partial H}{\partial E_2^h} = \lambda^2\mu_c - a^2\mu_k - \frac{\widehat{f}+\widehat{\delta}}{\widehat{f}}\mu_h > 0$. Then it is contradictory as the f.o.c of E_2^h requires $\frac{\partial H}{\partial E_2^h} \leq 0$.

allocated into industry 2 before any inexperienced ones [16].

(i) For the case $K_3^*(t) > 0$, the first order conditions with respect to $K_3(t)$ implies that $\frac{\lambda^3}{a^3} = \frac{\mu_k}{\mu_c}$. We have $\frac{\partial H}{\partial E_0^l} < 0$ and $\frac{\partial H}{\partial E_1^l} < 0$ if assumption 1 holds. Then $E_0^{l*}(t) = 0$, $E_1^{l*}(t) = 0$, and $E_2^{l*}(t) = \frac{\pi\widehat{f}}{\pi\widehat{f}+\widehat{\delta}}L^l(t)$. All workers stay in industry 2, and it is straightforward to check that such labor allocation minimizes the capital inflow to produce $C^*(t)$.

(ii) For the case $K_3^*(t) = 0$, the labor allocation minimizes capital inflow if and only if $E_0^{l*}(C) = 0$ [17]. So we need $\frac{\partial H}{\partial E_0^l} < 0$. If $\frac{\partial H}{\partial E_0^l} = 0$, we have

$$\mu_c - \frac{\widehat{f}+\widehat{\delta}}{\widehat{f}}\mu_l = 0, \lambda\mu_c - a\mu_k - \frac{\widehat{f}+\widehat{\delta}}{\widehat{f}}\mu_l \leq 0, \lambda^2\mu_c - a^2\mu_k - \frac{\pi\widehat{f}+\widehat{\delta}}{\pi\widehat{f}}\mu_l + \xi\mu_e = 0 \quad (47)$$

which implies

$$\frac{\mu_k}{\mu_c} \geq \frac{\lambda-1}{a};$$
$$\xi \geq \frac{\mu_c}{\mu_e}(\frac{\pi\widehat{f}+\widehat{\delta}}{\pi\widehat{f}+\pi\widehat{\delta}} + a(\lambda-1) - \lambda^2)$$

The first order condition with respect to consumption reads, $\mu_c(t) = e^{-\rho t}C(t)^{-\sigma}$. So to ensure that $\frac{\partial H}{\partial E_0^l} < 0$, we need

$$\xi < \frac{e^{-\rho t}C(t)^{-\sigma}}{\mu_e(t)}(\frac{\pi\widehat{f}+\widehat{\delta}}{\pi\widehat{f}+\pi\widehat{\delta}} + a(\lambda-1) - \lambda^2)$$

The above inequality needs to hold at time t when $K_3^*(t) = 0$. When the capital good is produced using an AK technology, consumption grows unboundedly to infinity. Let

[16] For our benchmark model, experienced (inexperienced) workers means particularly they are experienced (inexperienced) in industry 2. For the infinity-industry model, we use the term experienced (inexperienced) in industry n to avoid ambiguity. Assuming that all workers are initially inexperienced in industry $n+1$, we find the inexperienced workers in industry n are first allocated in industry $n+1$ before any experienced workers in industry n. So these two implications are not conflicting.

[17] For the if part, given that our assumption 1 holds, the economy minimizes the capital inflow when inexperienced workers only stay in industry 1 and industry 2, and experienced workers only stay in industry 2; for the only if part, when inexperienced workers stay in both industry 0 and industry 2, we can reallocate some of them in industry 1 to reduce the capital inflow without decreasing the output $C(t)$.

T be the last time when $C(T) = \lambda^2 \frac{\widehat{f}}{\widehat{f}+\widehat{\delta}} L$. For $t \geq T$, we must have $K_3(t) > 0$. The sufficient condition to have the labor allocation minimizing capital inflow at any time t is

$$\xi < \bar{\xi} \equiv \min_{t \in [0,T]} \{ \frac{e^{-\rho t} C(t)^{-\sigma}}{\mu_e(t)} (\frac{\pi \widehat{f} + \widehat{\delta}}{\pi \widehat{f} + \pi \widehat{\delta}} + a(\lambda - 1) - \lambda^2) \} \tag{48}$$

The right hand side is the minimum of positive values in a compact set. So $\bar{\xi}$ is also positive. The maximization of $C(t)$ given capital inflow $\Omega(t)$ is a dual problem of the minimization of $\Omega(t)$ given $C(t)$. When $\xi < \bar{\xi}$, the economy is at an optimal steady state equilibrium at each instant.

B.3 Proof of proposition 3

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

$$L = (E_0^l + E_0^h) + \lambda(E_1^l + E_1^h) + \lambda^2(E_2^l + E_2^h) + \frac{\lambda^3}{a^3} K_3 + \mu_\Omega(\Omega - a(E_1^l + E_1^h) - a^2(E_2^l + E_2^h) - K_3)$$
$$+ \mu_l(L^l - E_0^l \frac{\widehat{f} + \widehat{\delta}}{\widehat{f}} - E_1^l \frac{\widehat{f} + \widehat{\delta}}{\widehat{\delta}} - E_2^l \frac{\pi \widehat{f} + \widehat{\delta}}{\pi \widehat{f}}) + \mu_h(L^h - (E_0^h + E_1^h + E_2^h) \frac{\widehat{f} + \widehat{\delta}}{\widehat{f}}), \tag{49}$$

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 good 3. The Kuhn-Tucker conditions are

$$(1 - \mu_i \frac{\widehat{f} + \widehat{\delta}}{\widehat{f}}) E_0^i = 0, E_0^i \geq 0, 1 - \mu_i \frac{\widehat{f} + \widehat{\delta}}{\widehat{f}} \leq 0, i \in \{l, h\}, \tag{50}$$

$$(\lambda - \mu_i \frac{\widehat{f} + \widehat{\delta}}{\widehat{f}} - a\mu_\Omega) E_1^i = 0, E_1^i \geq 0, \lambda - \mu_i \frac{\widehat{f} + \widehat{\delta}}{\widehat{f}} - a\mu_\Omega \leq 0, i \in \{l, h\}, \tag{51}$$

$$(\lambda^2 - \mu_l \frac{\pi \widehat{f} + \widehat{\delta}}{\pi \widehat{f}} - a^2 \mu_\Omega) E_2^l = 0, E_2^l \geq 0, \lambda^2 - \mu_l \frac{\pi \widehat{f} + \widehat{\delta}}{\pi \widehat{f}} - a^2 \mu_\Omega \leq 0, \tag{52}$$

$$(\lambda^2 - \mu_h \frac{\widehat{f} + \widehat{\delta}}{\widehat{f}} - a^2 \mu_\Omega) E_2^h = 0, E_2^h \geq 0, \lambda^2 - \mu_h \frac{\widehat{f} + \widehat{\delta}}{\widehat{f}} - a^2 \mu_\Omega \leq 0, \tag{53}$$

$$(\frac{\lambda^3}{a^3} - \mu_\Omega) K_3 = 0, K_3 \geq 0, \frac{\lambda^3}{a^3} - \mu_\Omega \leq 0. \tag{54}$$

Given Assumption 1, there exist the following cases:

(1) if $\mu_\Omega > \frac{\lambda-1}{a}$: $\mu_l = \mu_h = \frac{\widehat{f}}{\widehat{f}+\widehat{\delta}}$, and only industry 0 exists: $E_0^l = \frac{\widehat{f}}{\widehat{f}+\widehat{\delta}} L^l, E_0^h = \frac{\widehat{f}}{\widehat{f}+\widehat{\delta}} L^h$.

(2) if $\mu_\Omega = \frac{\lambda-1}{a}$: $\mu_l = \mu_h = \frac{\widehat{f}}{\widehat{f}+\widehat{\delta}}$, and industry 0 and industry 1 coexist: $\sum_{j=0}^{1} E_j^l = \frac{\widehat{f}}{\widehat{f}+\widehat{\delta}} L^l, \sum_{j=0}^{1} E_j^h = \frac{\widehat{f}}{\widehat{f}+\widehat{\delta}} L^h$.

(3) if $\frac{\lambda^2-\lambda}{a^2-a} < \mu_\Omega < \frac{\lambda-1}{a}$: $\mu_l = \mu_h = \frac{\widehat{f}}{\widehat{f}+\widehat{\delta}}(\lambda - a\mu_E)$, and only industry 1 exists: $E_1^l = \frac{\widehat{f}}{\widehat{f}+\widehat{\delta}} L^l, E_1^h = \frac{\widehat{f}}{\widehat{f}+\widehat{\delta}} L^h$.

(4) if $\mu_\Omega = \frac{\lambda^2-\lambda}{a^2-a}$: $\mu_h = \mu_l = \frac{\widehat{f}}{\widehat{f}+\widehat{\delta}} \frac{(a-\lambda)\lambda}{a-1}$. Experienced workers stay in industry 1 and industry 2 : $\sum_{j=1}^{2} E_j^h = \frac{\widehat{f}}{\widehat{f}+\widehat{\delta}} L^h$. Inexperienced workers only stay in industry 1: $E_1^l = \frac{\widehat{f}}{\widehat{f}+\widehat{\delta}} L^l$.

(5) if $\frac{\lambda^2 \frac{\pi\widehat{f}}{\pi\widehat{f}+\widehat{\delta}} - \lambda \frac{\widehat{f}}{\widehat{f}+\widehat{\delta}}}{a^2 \frac{\pi\widehat{f}}{\pi\widehat{f}+\widehat{\delta}} - a \frac{\widehat{f}}{\widehat{f}+\widehat{\delta}}} < \mu_\Omega < \frac{\lambda^2-\lambda}{a^2-a}$: $\mu_l = \frac{\widehat{f}}{\widehat{f}+\widehat{\delta}}(\lambda - a\mu_\Omega)$ and $\mu_h = \frac{\widehat{f}}{\widehat{f}+\widehat{\delta}}(\lambda^2 - a^2\mu_\Omega)$. Experienced workers only stay in industry 2: $E_2^h = \frac{\widehat{f}}{\widehat{f}+\widehat{\delta}} L^h$ while inexperienced workers only stay in industry 1: $E_1^l = \frac{\widehat{f}}{\widehat{f}+\widehat{\delta}} L^l$.

(6) if $\mu_\Omega = \frac{\lambda^2 \frac{\pi\widehat{f}}{\pi\widehat{f}+\widehat{\delta}} - \lambda \frac{\widehat{f}}{\widehat{f}+\widehat{\delta}}}{a^2 \frac{\pi\widehat{f}}{\pi\widehat{f}+\widehat{\delta}} - a \frac{\widehat{f}}{\widehat{f}+\widehat{\delta}}}$: $\mu_l = \frac{\widehat{f}}{\widehat{f}+\widehat{\delta}} \frac{\pi\widehat{f}}{\pi\widehat{f}+\widehat{\delta}} \frac{(a-\lambda)\lambda}{a\frac{\pi\widehat{f}}{\pi\widehat{f}+\widehat{\delta}} - \frac{\widehat{f}}{\widehat{f}+\widehat{\delta}}}$ and $\mu_h = \frac{\widehat{f}}{\widehat{f}+\widehat{\delta}} \frac{\widehat{f}}{\widehat{f}+\widehat{\delta}} \frac{(a-\lambda)\lambda}{a\frac{\pi\widehat{f}}{\pi\widehat{f}+\widehat{\delta}} - \frac{\widehat{f}}{\widehat{f}+\widehat{\delta}}}$. Experienced workers stay in industry 2: $E_2^h = \frac{\widehat{f}}{\widehat{f}+\widehat{\delta}} L^h$ while inexperienced workers stay in industry 1 and industry 2: $\frac{\widehat{f}+\widehat{\delta}}{\widehat{f}} E_1^l + \frac{\pi\widehat{f}+\widehat{\delta}}{\pi\widehat{f}} E_2^l = L^l$.

(7) if $\frac{\lambda^3}{a^3} < \mu_\Omega < \frac{\lambda^2 \frac{\pi\widehat{f}}{\pi\widehat{f}+\widehat{\delta}} - \lambda \frac{\widehat{f}}{\widehat{f}+\widehat{\delta}}}{a^2 \frac{\pi\widehat{f}}{\pi\widehat{f}+\widehat{\delta}} - a \frac{\widehat{f}}{\widehat{f}+\widehat{\delta}}}$: $\mu_l = \frac{\pi\widehat{f}}{\pi\widehat{f}+\widehat{\delta}}(\lambda^2 - a^2\mu_E)$ and $\mu_h = \frac{\widehat{f}}{\widehat{f}+\widehat{\delta}}(\lambda^2 - a^2\mu_E)$. Workers of both types stay in industry 2: $E_2^l = \frac{\pi\widehat{f}}{\pi\widehat{f}+\widehat{\delta}} L^l$ and $E_2^h = \frac{\widehat{f}}{\widehat{f}+\widehat{\delta}} L^h$ and industry 3, which only uses capital, is not in production.

(8) if $\mu_\Omega = \frac{\lambda^3}{a^3}$: Both experienced workers only stay in industry 2 and industry 3 (which only uses capital) is in production.

In equilibrium, $\mu_\Omega \geq \frac{\lambda^3}{a^3}$ and the market clearing condition of capital flow Ω implies the results presented in Table B.1:

Table B.1: the required capital flow in the optimal steady state equilibrium

(1). $\frac{\widehat{f}}{\widehat{f}+\widehat{\delta}}(L^h+L^l) \leq C < \lambda\frac{\widehat{f}}{\widehat{f}+\widehat{\delta}}(L^h+L^l)$	(2). $\lambda\frac{\widehat{f}}{\widehat{f}+\widehat{\delta}}(L^h+L^l) \leq C < \frac{\widehat{f}}{\widehat{f}+\widehat{\delta}}(\lambda L^l+\lambda^2 L^h)$
$\Omega = \frac{a}{\lambda-1}C - \frac{a}{\lambda-1}\frac{\widehat{f}}{\widehat{f}+\widehat{\delta}}(L^h+L^l)$ $E_0^l + E_0^h = \frac{\lambda\frac{\widehat{f}}{\widehat{f}+\widehat{\delta}}(L^h+L^l)-C}{\lambda-1}$ $E_1^l + E_1^h = \frac{C-\frac{\widehat{f}}{\widehat{f}+\widehat{\delta}}(L^h+L^l)}{\lambda-1}$ $E_2^l = 0,\ E_2^h = 0$	$\Omega = \frac{a^2-a}{\lambda^2-\lambda}C - \frac{a(a-\lambda)}{\lambda-1}\frac{\widehat{f}}{\widehat{f}+\widehat{\delta}}(L^h+L^l)$ $E_0^l = 0,\ E_0^h = 0$ $E_1^l = \frac{\widehat{f}}{\widehat{f}+\widehat{\delta}}L^l,\ E_1^h = \frac{\lambda^2\frac{\widehat{f}}{\widehat{f}+\widehat{\delta}}(L^h+L^l)-C}{\lambda^2-\lambda} - \frac{\widehat{f}}{\widehat{f}+\widehat{\delta}}L^l$ $E_2^l = 0,\ E_2^h = \frac{C-\lambda\frac{\widehat{f}}{\widehat{f}+\widehat{\delta}}(L^h+L^l)}{\lambda^2-\lambda}$
(3). $\frac{\widehat{f}}{\widehat{f}+\widehat{\delta}}(\lambda L^l+\lambda^2 L^h) \leq C < \lambda^2(\frac{\pi\widehat{f}}{\pi\widehat{f}+\widehat{\delta}}L^l + \frac{\widehat{f}}{\widehat{f}+\widehat{\delta}}L^h)$	(4). $C \geq \lambda^2(\frac{\pi\widehat{f}}{\pi\widehat{f}+\widehat{\delta}}L^l + \frac{\widehat{f}}{\widehat{f}+\widehat{\delta}}L^h)$
$\Omega = \frac{a^2\frac{\pi\widehat{f}}{\pi\widehat{f}+\widehat{\delta}}-a\frac{\widehat{f}}{\widehat{f}+\widehat{\delta}}}{\lambda^2\frac{\pi\widehat{f}}{\pi\widehat{f}+\widehat{\delta}}-\lambda\frac{\widehat{f}}{\widehat{f}+\widehat{\delta}}}C - \frac{\widehat{f}}{\widehat{f}+\widehat{\delta}}\frac{a(a-\lambda)}{\lambda\frac{\pi\widehat{f}}{\pi\widehat{f}+\widehat{\delta}}-\frac{\widehat{f}}{\widehat{f}+\widehat{\delta}}}(\frac{\pi\widehat{f}}{\pi\widehat{f}+\widehat{\delta}}L^l + \frac{\widehat{f}}{\widehat{f}+\widehat{\delta}}L^h)$ $E_0^l = 0,\ E_0^h = 0$ $E_1^l = \frac{\widehat{f}}{\widehat{f}+\widehat{\delta}}(\frac{\lambda^2(\frac{\pi\widehat{f}}{\pi\widehat{f}+\widehat{\delta}}L^l+\frac{\widehat{f}}{\widehat{f}+\widehat{\delta}}L^h)-C}{\lambda^2\frac{\pi\widehat{f}}{\pi\widehat{f}+\widehat{\delta}}-\lambda\frac{\widehat{f}}{\widehat{f}+\widehat{\delta}}}),\ E_1^h = 0$ $E_2^l = \frac{\pi\widehat{f}}{\pi\widehat{f}+\widehat{\delta}}(\frac{C-\lambda^2\frac{\widehat{f}}{\widehat{f}+\widehat{\delta}}L^h-\lambda\frac{\widehat{f}}{\widehat{f}+\widehat{\delta}}L^l}{\lambda^2\frac{\pi\widehat{f}}{\pi\widehat{f}+\widehat{\delta}}-\lambda\frac{\widehat{f}}{\widehat{f}+\widehat{\delta}}}),\ E_2^h = \frac{\widehat{f}}{\widehat{f}+\widehat{\delta}}L^h$	$\Omega = \frac{a^3}{\lambda^3}C - \frac{a^2(a-\lambda)}{\lambda}(\frac{\pi\widehat{f}}{\pi\widehat{f}+\widehat{\delta}}L^l + \frac{\widehat{f}}{\widehat{f}+\widehat{\delta}}L^h)$ $E_0^l = 0,\ E_0^h = 0$ $E_1^l = 0,\ E_1^h = 0$ $E_2^l = \frac{\pi\widehat{f}}{\pi\widehat{f}+\widehat{\delta}}L^l,\ E_2^h = \frac{\widehat{f}}{\widehat{f}+\widehat{\delta}}L^h$

.

B.4 Solving the dynamic path for the benchmark model

The necessary conditions are

$$\dot{\mu} = (\rho - A)\mu, \tag{55}$$

$$\hat{C} \in \text{argmax}\{\frac{C^{1-\sigma}-1}{1-\sigma} + \mu(AK - \Omega(C, L^l, L-L^l))\} \tag{56}$$

Under Assumption 1, the value of capital, μ, decreases exponentially at a constant rate. We know that $\Omega(C, L^l, L-L^l)$ is a piece-wise linear function. When $\hat{C}$ is not at the endpoints of linear parts, and we have $\hat{C}^{-\sigma} = \mu\frac{\partial E(\hat{C},L^l,L-L^l)}{\partial C}$, then it solves equation (56). As $\frac{\partial E(\hat{C},L^l,L-L^l)}{\partial C}$ remains constant around the local area, locally we have the Euler equation:

$$\frac{\dot{\hat{C}}}{\hat{C}} = \frac{A-\rho}{\sigma} \tag{57}$$

When $\hat{C}$ reaches the endpoints, the following conditions hold:

$$\hat{C}^{-\sigma} - \mu\frac{\partial\Omega(\hat{C}, L^l, L-L^l)}{\partial C_-} \geq 0,\ \hat{C}^{-\sigma} - \mu\frac{\partial\Omega(\hat{C}, L^l, L-L^l)}{\partial C_+} \leq 0 \tag{58}$$

where $\frac{\partial\cdot}{\partial x_-}$ and $\frac{\partial\cdot}{\partial x_+}$ denote the left and right derivatives. To maximize the value in (56), $\hat{C}$ should stay at the endpoints until $\hat{C}^{-\sigma} - \mu\frac{\partial\Omega(\hat{C},L^l,L-L^l)}{\partial C_+} > 0$. Notice that in our model, inexperienced workers become experienced, so the endpoints ($\bar{C}$, $\underline{C}$) as well as the solution $\hat{C}$ may be moving with time. It then follows that there are three stages for the industrial upgrading from industry 1 to industry 2:

Stage I: All workers search in industry 1:

$$C(t) = \lambda\frac{\widehat{f}}{\widehat{f}+\widehat{\delta}}L, \quad \text{if} \quad \mu \in ((\frac{\lambda\widehat{f}}{\widehat{f}+\widehat{\delta}}L)^{-\sigma}M_1, (\frac{\lambda\widehat{f}}{\widehat{f}+\widehat{\delta}}L)^{-\sigma}\frac{(\lambda-1)}{a}] \tag{59}$$

where $M_1 = \frac{\lambda^2\frac{\pi\widehat{f}}{\pi\widehat{f}+\widehat{\delta}} - \lambda\frac{\widehat{f}}{\widehat{f}+\widehat{\delta}}}{a^2\frac{\pi\widehat{f}}{\pi\widehat{f}+\widehat{\delta}} - a\frac{\widehat{f}}{\widehat{f}+\widehat{\delta}}}$. The life span of this first stage is $\frac{log(a/((\lambda-1)M_1))}{A-\rho}$. In stage I, consumption remains the same while capital accumulates with time.

Stage II: Inexperienced workers are employed in both industries while experienced in industry 2 only:

$$C(t) = \lambda\frac{\widehat{f}}{\widehat{f}+\widehat{\delta}}Le^{g_c(t-t_1)}, \quad \text{if} \quad \mu \in [\bar{C}^{-\sigma}M_1, (\frac{\lambda\widehat{f}}{\widehat{f}+\widehat{\delta}}L)^{-\sigma}M_1] \tag{60}$$

where $g_c \equiv \frac{A-\rho}{\sigma}$, and t_1 denotes the time when $\mu(t_1) = (\frac{\lambda\widehat{f}}{\widehat{f}+\widehat{\delta}}L)^{-\sigma}M_1$ is satisfied. That is, t_1 is the time when inexperienced workers begin to search in industry 2. We will derive the closed form solution of $L^l(t)$ later. For the condition (60) to hold, we implicitly assume that experienced workers in industry 2 are not enough to satisfy the increase of consumption demand, i.e. $\frac{\partial\Omega(C,L-L^l,L^l)}{\partial C} = \frac{1}{M_1}$ instead of $\frac{\lambda(\lambda-1)}{a(a-1)}$. So that there will always be inexperienced workers in industry 2 when the industrial upgrading from industry 1 to 2 takes place. We verify this later by using the expression of $L^l(t)$.

Stage III: All workers search in industry 2 and industry 3 is not in production:

$$C(t) = \bar{C}(t), \quad \text{if} \quad \mu \in [\bar{C}(t)^{-\sigma}\frac{\lambda^3}{a^3}, \bar{C}(t)^{-\sigma}M_1] \tag{61}$$

In this stage, consumption increases as inexperienced workers learn to become experienced, and the intermediate goods of industry 2 increases. Capital also accumulates over time but industry 3, which uses capital only, is not in production, because of the jump in marginal cost.

From Table 2.1, we know that after t_1:

$$E_2^l = \frac{\pi\widehat{f}}{\pi\widehat{f}+\widehat{\delta}}\left(\frac{C - \lambda^2\frac{\widehat{f}}{\widehat{f}+\widehat{\delta}}(L-L^l) - \lambda\frac{\widehat{f}}{\widehat{f}+\widehat{\delta}}L^l}{\lambda^2\frac{\pi\widehat{f}}{\pi\widehat{f}+\widehat{\delta}} - \lambda\frac{\widehat{f}}{\widehat{f}+\widehat{\delta}}}\right), \tag{62}$$

Combining with (60), we obtain

$$E_2^l(t) = N_1 e^{g_c(t-t_1)} - (N_1 + N_2 L) + N_2 L^l, \tag{63}$$

where

$$N_1 = \frac{L\frac{\widehat{f}}{\widehat{f}+\widehat{\delta}}\frac{\pi\widehat{f}}{\pi\widehat{f}+\widehat{\delta}}}{\lambda\frac{\pi\widehat{f}}{\pi\widehat{f}+\widehat{\delta}} - \frac{\widehat{f}}{\widehat{f}+\widehat{\delta}}}, N_2 = \frac{(\lambda-1)\frac{\widehat{f}}{\widehat{f}+\widehat{\delta}}\frac{\pi\widehat{f}}{\pi\widehat{f}+\widehat{\delta}}}{\lambda\frac{\pi\widehat{f}}{\pi\widehat{f}+\widehat{\delta}} - \frac{\widehat{f}}{\widehat{f}+\widehat{\delta}}}. \tag{64}$$

As E_2^l cannot exceed the total supply of inexperienced workers, we have

$$E_2^l(t) = \min\{N_1 e^{g_c(t-t_1)} - (N_1 + N_2 L) + N_2 L^l, \frac{\pi\widehat{f}}{\pi\widehat{f}+\widehat{\delta}}L^l\}. \tag{65}$$

As inexperienced workers in industry 2 gradually become experienced through on-the-job learning, then

$$\dot{L}^l = -\xi E_2^l. \tag{66}$$

Substituting (65) into the above equation yields

$$\dot{L}_l = -\xi \min\{N_1 e^{g_c(t-t_1)} - (N_1 + N_2 L) + N_2 L^l, \frac{\pi\widehat{f}}{\pi\widehat{f}+\widehat{\delta}}L^l\}. \tag{67}$$

Solving the above differential equation with the initial condition that $L^l(t_1) = L$, we obtain Equations (22) and (23) in the main text.

From equation (22), we know that $E_2^l(t)$ is positive during $t_1 < t \leq t_{2,l}$, which verifies that the consumption increased brought by the experience accumulation is not enough to support the increasing consumption demand. More inexperienced workers are allocated to industry 2. After $t_{2,l}$, all workers stay in industry 2, and $C(t) = \bar{C}(t)$. We can also compute $\frac{\dot{\bar{C}}(t)}{\bar{C}(t)}$ as

$$\frac{\dot{\bar{C}}(t)}{\bar{C}(t)} = \frac{\frac{\pi\widehat{f}}{\pi\widehat{f}+\widehat{\delta}}\xi\left(\frac{\widehat{f}}{\widehat{f}+\widehat{\delta}} - \frac{\pi\widehat{f}}{\pi\widehat{f}+\widehat{\delta}}\right)L^l(t)}{\frac{\widehat{f}}{\widehat{f}+\widehat{\delta}}(L - L^l(t)) + \frac{\pi\widehat{f}}{\pi\widehat{f}+\widehat{\delta}}L^l(t)}, \tag{68}$$

which is a decreasing function of time as $L^l(t)$ is a decreasing function of time. The value $\frac{\dot{\bar{C}}(t)}{\bar{C}(t)}$ at $t_{2,l}$ is smaller than g_c, so it is smaller than g_c after $t_{2,l}$. This verifies that the experience accumulation is not enough to support the increasing consumption demand, and inexperienced workers will not be allocated back to industry 1 after $t_{2,l}$.

B.5 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

$$\begin{aligned}\frac{\partial E_2^l}{\partial \xi} &= \frac{N_1 N_2 g_c}{(N_2\xi + g_c)^2}[((N_2\xi + g_c)(t - t_1) + 1)e^{-N_2\xi(t-t_1)} - e^{g_c(t-t_1)}] \\ &= \frac{N_1 N_2 g_c e^{-N_2\xi(t-t_1)}}{(N_2\xi + g_c)^2}[((N_2\xi + g_c)(t - t_1) + 1) - e^{(N_2\xi+g_c)(t-t_1)}] < 0\end{aligned} \tag{69}$$

The last term is negative for $e^x > x + 1$ if $x \neq 0$. $\frac{\partial E_2^l}{\partial \xi} < 0$ implies that aggregate unemployment rate shifts downward in the first half. Similarly, we can prove that $\frac{\partial L^l}{\partial \xi} < 0$. Then we show $t'_{2,l} - t'_1 > t_{2,l} - t_1$: the expansion stage of industry 2 gets longer. Notice that $t_{2,l}$ is the first time when

$$E_2^l = \frac{\pi\widehat{f}}{\pi\widehat{f} + \widehat{\delta}}L^l, \tag{70}$$

Substituting (22) and (23) into (70) gives

$$LHS(t) \equiv \frac{N_1}{N_2\xi+g_c}\left(\frac{\pi\widehat{f}}{\pi\widehat{f}+\widehat{\delta}}\xi e^{g_c(t-t_1)} + \frac{\pi\widehat{f}}{\pi\widehat{f}+\widehat{\delta}}\frac{g_c}{N2}e^{-N_2\xi(t-t_1)} + g_c e^{g_c(t-t_1)} - g_c e^{-N_2\xi(t-t_1)}\right)$$
$$= \frac{\pi\widehat{f}}{\pi\widehat{f}+\widehat{\delta}}\left(L + \frac{N_1}{N_2}\right) \quad (71)$$

It is straightforward to check that $LHS(t)$ is an increasing function of t, and that

$$\frac{\partial LHS(t)}{\partial \xi} = \frac{N_1 g_c e^{-N_2\xi(t-t_1)}\left(N_2 - \frac{\pi\widehat{f}}{\pi\widehat{f}+\widehat{\delta}}\right)}{(N_2\xi+g_c)^2}\left[\left((N_2\xi+g_c)(t-t_1)+1\right) - e^{(N_2\xi+g_c)(t-t_1)}\right] \quad (72)$$

From (64), we know $N_2 > \frac{\pi\widehat{f}}{\pi\widehat{f}+\widehat{\delta}}$. Thus $\frac{\partial LHS(t)}{\partial \xi} < 0$. Therefore when ξ increases we must have $t'_{2,l} - t'_1 > t_{2,l} - t_1$ to ensure that (71) holds. As $L^l(t)$ is a decreasing function of ξ and t, when $\xi' > \xi$ and $t'_{2,l} - t'_1 > t_{2,l} - t_1$, we have $L^{l'}(t'_{2,l}) < L^l(t_{2,l})$. Combining with (70) and (21), it implies that $U'(t'_{2,l}) < U(t_{2,l})$. Aggregate unemployment rate reaches a lower peak value when ξ increases.
Finally we prove that aggregate unemployment rate decreases for the second half for $t > t_{2,l}$ if ξ increases. Define

$$L^l_{hypo}(t) \equiv \frac{-N_1}{N_2(N_2\xi+g_c)}\left(\xi N_2 e^{g_c(t-t_1)} + g_c e^{-N_2\xi(t-t_1)}\right) + \frac{N_1}{N_2} + L, \forall t > t_1 \quad (73)$$

It follows from the definitions that $L^l_{hypo}(t) = L^l(t), \forall t_1 < t \le t_{2,l}$. Besides,

$$\dot{L}^l_{hypo}(t) < \dot{L}^l_{hypo}(t_{2,l}) = \dot{L}^l(t_{2,l}) < \dot{L}^l(t), \forall t > t_{2,l} \quad (74)$$

Then $L^l_{hypo}(t) < L^l(t), \forall t > t_{2,l}$. We want to show $U'(t'_{2,l}) < U(t'_{2,l} - t'_1 + t_1)$. If this is true, we know that the aggregate unemployment actually shifts downward for the whole path as it decreases exponentially at a larger rate $-\frac{\pi\widehat{f}}{\pi\widehat{f}+\widehat{\delta}}\xi'$ for $t > t'_{2,l}$. From the discussion above, we have $L^{l'}(t'_{2,l}) = {L^{l'}}_{hypo}(t'_{2,l}) < L^l_{hypo}(t'_{2,l} - t'_1 + t_1) < L^l(t'_{2,l} - t'_1 + t_1)$. The first inequality holds as $L^l_{hypo}(t)$ is a decreasing function of ξ, and the second inequality holds as $L^l_{hypo}(t) < L^l(t), \forall t > t_{2,l}$ and $t'_{2,l} - t'_1 > t_{2,l} - t_1$. Combining with (21), we prove $U'(t'_{2,l}) < U(t'_{2,l} - t'_1 + t_1)$.

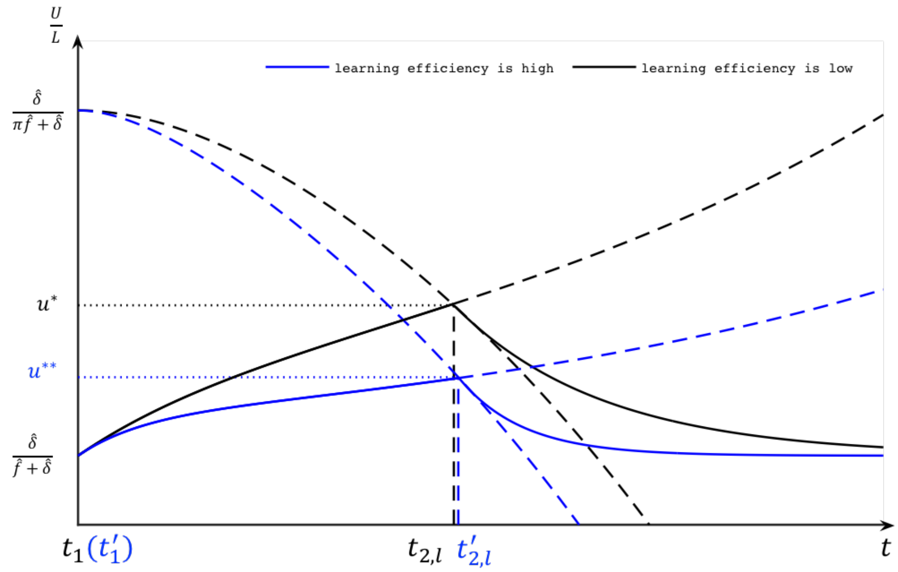


Figure B.1: How unemployment rate changes with learning efficiency

(2) *capital-goods production efficiency* When A increases, the growth rate of consumption goods g_c increases, which results in more rapid industrial upgrading. Taking the partial derivative of E_2^l and L^l with respect to g_c when $t_1 < t \leq t_{2,l}$, we have

$$\frac{\partial E_2^l}{\partial g_c} = \frac{N_1}{N_2\xi + g_c}[\frac{N_2\xi}{N_2\xi + g_c}(e^{g_c(t-t_1)} - e^{-N_2\xi(t-t_1)}) + g_c(t-t_1)e^{g_c(t-t_1)}] > 0 \tag{75}$$

$$\begin{aligned}\frac{\partial L^l}{\partial g_c} &= \frac{-N_1}{N_2(N_2\xi + g_c)}[\xi N_2(t-t_1)e^{g_c(t-t_1)} - \frac{N_2\xi}{N_2\xi + g_c}e^{g_c(t-t_1)} + \frac{N_2\xi}{N_2\xi + g_c}e^{-N_2\xi(t-t_1)}] \\ &= \frac{-\xi N_1 e^{g_c(t-t_1)}}{(N_2\xi + g_c)^2}[(t-t_1)(N_2\xi + g_c) - 1 + e^{-(N_2\xi+g_c)(t-t_1)}] < 0\end{aligned} \tag{76}$$

The last term is negative using again the inequality $e^{-x} > -x + 1$. Therefore, aggregate unemployment rate shifts upward while the number of inexperienced workers shifts downward when $t_1 < t \leq t_{2,l}$. Immediately, we have $t_{2,l}' - t_1' < t_{2,l} - t_1$ for $A' > A$, indicating that it takes less time for structural unemployment rate to reach its peak value. Meanwhile structural unemployment is larger during the industrial upgrading. Then we prove it reaches a larger peak value. We make use of the properties that L^l is decreasing with t while E_2^l is increasing with t before $t_{2,l}$.

Using (70), we have

$$\frac{\partial E_2^l(t_{2,l})}{\partial g_c} = \frac{\partial E_2^l}{\partial g_c} + \frac{\partial E_2^l}{\partial t}\frac{\partial t_{2,l}}{\partial g_c} = \frac{\pi \widehat{f}}{\pi \widehat{f} + \widehat{\delta}}(\frac{\partial L^l}{\partial g_c} + \frac{\partial L^l}{\partial t}\frac{\partial t_{2,l}}{\partial g_c}) \tag{77}$$

Rearranging (77) to find $\frac{\partial t_{2,l}}{\partial g_c}$, and substituting it into (77), lead to

$$\frac{\partial E_2^l(t_{2,l})}{\partial g_c} = \frac{\pi \widehat{f}}{\pi \widehat{f} + \widehat{\delta}} \frac{\frac{\partial E_2^l}{\partial g_c}\frac{\partial L^l}{\partial t} - \frac{\partial L^l}{\partial g_c}\frac{\partial E_2^l}{\partial t}}{\frac{\pi \widehat{f}}{\pi \widehat{f}+\widehat{\delta}}\frac{\partial L^l}{\partial t} - \frac{\partial E_2^l}{\partial t}} \Rightarrow \frac{\partial E_2^l(t_{2,l})}{\partial g_c} > 0 \Leftrightarrow \frac{\frac{\partial E_2^l}{\partial g_c}}{\frac{\partial E_2^l}{\partial t}}|_{t_{2,l}} > \frac{\frac{\partial L^l}{\partial g_c}}{\frac{\partial L^l}{\partial t}}|_{t_{2,l}} \tag{78}$$

Comparing $\frac{\frac{\partial L^l}{\partial g_c}}{\frac{\partial L^l}{\partial t}}|_{t_{2,l}}$ and $\frac{\frac{\partial E_2^l}{\partial g_c}}{\frac{\partial E_2^l}{\partial t}}|_{t_{2,l}}$. If $\frac{\frac{\partial E_2^l}{\partial g_c}}{\frac{\partial E_2^l}{\partial t}}|_{t_{2,l}} > \frac{\frac{\partial L^l}{\partial g_c}}{\frac{\partial L^l}{\partial t}}|_{t_{2,l}}$ indicates that the economy has a larger peak value of the aggregate unemployment rate when A increases. Taking the partial derivatives of E_2^l, we have

$$\begin{aligned}
\frac{\frac{\partial E_2^l}{\partial g_c}}{\frac{\partial E_2^l}{\partial t}}|_{t_{2,l}} &= \frac{\frac{N_1}{N_2\xi+g_c}[\frac{N_2\xi}{N_2\xi+g_c}(e^{g_c(t_{2,l}-t_1)} - e^{-N_2\xi(t_{2,l}-t_1)}) + g_c(t_{2,l}-t_1)e^{g_c(t_{2,l}-t_1)}]}{\frac{N_1 g_c}{N_2\xi+g_c}[g_c e^{g_c(t_{2,l}-t_1)} + N_2\xi e^{-N_2\xi(t_{2,l}-t_1)}]} \\
&= \frac{\frac{N_2\xi}{N_2\xi+g_c}(e^{g_c(t_{2,l}-t_1)} - e^{-N_2\xi(t_{2,l}-t_1)}) + g_c(t_{2,l}-t_1)e^{g_c(t_{2,l}-t_1)}}{g_c[g_c e^{g_c(t_{2,l}-t_1)} + N_2\xi e^{-N_2\xi(t_{2,l}-t_1)}]} \\
&= \frac{\frac{1}{g_c}(e^{g_c(t_{2,l}-t_1)} - e^{-N_2\xi(t_{2,l}-t_1)}) - \frac{1}{N_2\xi+g_c}(e^{g_c(t_{2,l}-t_1)} - e^{-N_2\xi(t_{2,l}-t_1)}) + (t_{2,l}-t_1)e^{g_c(t_{2,l}-t_1)}}{g_c e^{g_c(t_{2,l}-t_1)} + N_2\xi e^{-N_2\xi(t_{2,l}-t_1)}}
\end{aligned} \tag{79}$$

Similarly, taking the partial derivatives of L^l

$$\begin{aligned}
\frac{\frac{\partial L^l}{\partial g_c}}{\frac{\partial L^l}{\partial t}}|_{t_{2,l}} &= \frac{\frac{\xi N_1 N_2}{N_2(N_2\xi+g_c)^2}[((t_{2,l}-t_1)(N_2\xi+g_c)-1)e^{g_c(t_{2,l}-t_1)} + e^{-N_2\xi(t_{2,l}-t_1)}]}{\frac{\xi N_1 N_2 g_c}{N_2(N_2\xi+g_c)}[e^{g_c(t_{2,l}-t_1)} - e^{-N_2\xi(t_{2,l}-t_1)}]} \\
&= \frac{\frac{1}{N_2\xi+g_c}[((t_{2,l}-t_1)(N_2\xi+g_c)-1)e^{g_c(t_{2,l}-t_1)} + e^{-N_2\xi(t_{2,l}-t_1)}]}{g_c[e^{g_c(t_{2,l}-t_1)} - e^{-N_2\xi(t_{2,l}-t_1)}]} \\
&= \frac{-\frac{1}{N_2\xi+g_c}(e^{g_c(t_{2,l}-t_1)} - e^{-N_2\xi(t_{2,l}-t_1)}) + (t_{2,l}-t_1)e^{g_c(t_{2,l}-t_1)}}{g_c[e^{g_c(t_{2,l}-t_1)} - e^{-N_2\xi(t_{2,l}-t_1)}]}
\end{aligned} \tag{80}$$

From equations (79) and (80), $\frac{\frac{\partial E_2^l}{\partial g_c}}{\frac{\partial E_2^l}{\partial t}}|_{t_{2,l}} > \frac{\frac{\partial L^l}{\partial g_c}}{\frac{\partial L^l}{\partial t}}|_{t_{2,l}}$ is equivalent to

$$\frac{\frac{\partial E_2^l}{\partial g_c}}{\frac{\partial E_2^l}{\partial t}}|_{t_{2,l}} > \frac{\frac{\partial L^l}{\partial g_c}}{\frac{\partial L^l}{\partial t}}|_{t_{2,l}}$$

$$\Leftrightarrow [e^{g_c(t_{2,l}-t_1)} - e^{-N_2\xi(t_{2,l}-t_1)}]^2 > [e^{-N_2\xi(t_{2,l}-t_1)} + ((N_2\xi + g_c)(t_{2,l}-t_1) - 1)e^{g_c(t_{2,l}-t_1)}]e^{-N_2\xi(t_{2,l}-t_1)}$$

$$\Leftrightarrow e^{g_c(t_{2,l}-t_1)} > [(N_2\xi + g_c)(t_{2,l}-t_1) + 1]e^{-N_2\xi(t_{2,l}-t_1)}$$

$$\Leftrightarrow e^{(N_2\xi+g_c)(t_{2,l}-t_1)} > [(N_2\xi + g_c)(t_{2,l}-t_1) + 1] \tag{81}$$

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 lower after the duration $t_{2,l} - t_1$. If not, $E_{2,l}$ must be larger at any time in $(t_1', t_1' + t_{2,l} - t_1]$. It induces that L^l decreases at a larger rate during this period. Given that all workers are inexperienced initially $L_l'(t_1') = L$, L^l must come to a lower value: $L_l'(t_1' + t_{2,l} - t_1) < L^l(t_{2,l})$, and thus aggregate unemployment rate must be lower after the duration $t_{2,l} - t_1$ as all inexperienced workers are now in industry 2. This leads to the contradiction.

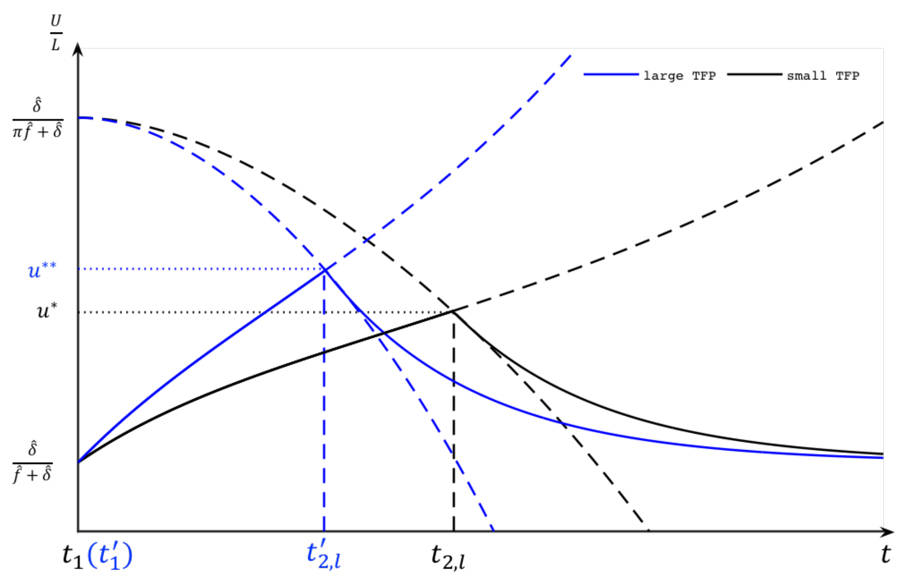


Figure B.2: How unemployment rate changes with capital-goods production efficiency

(3) *Mismatch* We prove that when π decreases, aggregate unemployment rate shifts upward after t_1. Equivalently, $\frac{\pi\widehat{f}}{\pi\widehat{f}+\widehat{\delta}}$ decreases. First prove that $\frac{\partial E_2^l}{\partial\frac{\pi\widehat{f}}{\pi\widehat{f}+\widehat{\delta}}}|_t < 0$ when $t_1 < t \leq t_{2,l}$. As $\frac{N_1}{N_2} = \frac{L}{\lambda-1}$, we have

$$\frac{\partial E_2^l}{\partial\frac{\pi\widehat{f}}{\pi\widehat{f}+\widehat{\delta}}} = \frac{L}{\lambda-1}\frac{\partial N_2}{\partial\frac{\pi\widehat{f}}{\pi\widehat{f}+\widehat{\delta}}}[(\frac{g_c}{N_2\xi+g_c})^2(e^{g_c(t-t_1)} - e^{-N_2\xi(t-t_1)}) + \frac{g_c N_2\xi(t-t_1)}{N_2\xi+g_c}e^{-N_2\xi(t-t_1)}] < 0 \quad (82)$$

The inequality holds as $\frac{\partial N_2}{\partial\frac{\pi\widehat{f}}{\pi\widehat{f}+\widehat{\delta}}} < 0$. Using equations (21) and (82), we obtain

$$U = \frac{\widehat{\delta}}{\widehat{f}+\widehat{\delta}}L + (\frac{\frac{\widehat{f}}{\widehat{f}+\widehat{\delta}}}{\frac{\pi\widehat{f}}{\pi\widehat{f}+\widehat{\delta}}} - 1)E_2^l \Rightarrow \frac{\partial U}{\partial\frac{\pi\widehat{f}}{\pi\widehat{f}+\widehat{\delta}}} = -\frac{\frac{\widehat{f}}{\widehat{f}+\widehat{\delta}}}{(\frac{\pi\widehat{f}}{\pi\widehat{f}+\widehat{\delta}})^2}E_2^l + (\frac{\frac{\widehat{f}}{\widehat{f}+\widehat{\delta}}}{\frac{\pi\widehat{f}}{\pi\widehat{f}+\widehat{\delta}}} - 1)\frac{\partial E_2^l}{\partial\frac{\pi\widehat{f}}{\pi\widehat{f}+\widehat{\delta}}} < 0 \quad (83)$$

It indicates that when mismatch is severer, aggregate unemployment rate shifts upward for the first half. To explain for the second half, first define U^* as follows

$$U^* \equiv \frac{\widehat{\delta}}{\widehat{f}+\widehat{\delta}}L + (\frac{\widehat{f}}{\widehat{f}+\widehat{\delta}} - \frac{\pi\widehat{f}}{\pi\widehat{f}+\widehat{\delta}})L^l, \quad (84)$$

From (70), $t_{2,l}$ is also the first time when $U = U^*$. Taking the partial derivatives gives

$$\frac{\partial U^*}{\partial\frac{\pi\widehat{f}}{\pi\widehat{f}+\widehat{\delta}}} = (\frac{\widehat{f}}{\widehat{f}+\widehat{\delta}} - \frac{\pi\widehat{f}}{\pi\widehat{f}+\widehat{\delta}})\frac{\partial L^l}{\partial\frac{\pi\widehat{f}}{\pi\widehat{f}+\widehat{\delta}}} - L^l \quad (85)$$

$$\frac{\partial L^l}{\partial\frac{\pi\widehat{f}}{\pi\widehat{f}+\widehat{\delta}}} = -\frac{L}{\lambda-1}\frac{\xi g_c}{(N_2\xi+g_c)^2}[e^{g_c(t-t_1)} - (1+(N_2\xi+g_c)(t-t_1))e^{-N_2\xi(t-t_1)}]\frac{\partial N_2}{\partial\frac{\pi\widehat{f}}{\pi\widehat{f}+\widehat{\delta}}} \quad (86)$$

$$\frac{\partial N_2}{\partial\frac{\pi\widehat{f}}{\pi\widehat{f}+\widehat{\delta}}} = -(\frac{N_2}{\frac{\pi\widehat{f}}{\pi\widehat{f}+\widehat{\delta}}})^2\frac{1}{\lambda-1} \quad (87)$$

Combining equations (85), (86), (87) and (23), we have

$$\frac{\partial U^*}{\partial\frac{\pi\widehat{f}}{\pi\widehat{f}+\widehat{\delta}}} = \frac{L\xi g_c}{(\lambda-1)^2}\frac{(\frac{N_2}{\frac{\pi\widehat{f}}{\pi\widehat{f}+\widehat{\delta}}})^2}{(N_2\xi+g_c)^2}[e^{g_c(t-t_1)} - (1+(N_2\xi+g_c)(t-t_1))e^{-N_2\xi(t-t_1)}](\frac{\widehat{f}}{\widehat{f}+\widehat{\delta}} - \frac{\pi\widehat{f}}{\pi\widehat{f}+\widehat{\delta}})$$
$$+\frac{L}{\lambda-1}[\frac{\xi N_2}{(N_2\xi+g_c)}e^{g_c(t-t_1)} + \frac{g_c}{N_2\xi+g_c}e^{-N_2\xi(t-t_1)}] - \frac{\lambda L}{\lambda-1} \quad (88)$$

It implies that

$$\frac{\partial U^*}{\partial \frac{\pi f}{\pi \hat{f}+\hat{\delta}}} < 0 \Leftrightarrow \lambda > (\frac{N_2}{\frac{\pi f}{\pi \hat{f}+\hat{\delta}}})^2 \frac{\frac{1}{\lambda-1}\xi g_c}{(N_2\xi+g_c)^2}[e^{g_c(t-t_1)} - (1+(N_2\xi+g_c)(t-t_1))e^{-N_2\xi(t-t_1)}](\frac{\hat{f}}{\hat{f}+\hat{\delta}} - \frac{\pi\hat{f}}{\pi\hat{f}+\hat{\delta}})$$
$$+[\frac{\xi N_2}{(N_2\xi+g_c)}e^{g_c(t-t_1)} + \frac{g_c}{N_2\xi+g_c}e^{-N_2\xi(t-t_1)}] \quad (89)$$

The Right hand side of equation (89) is an increasing function of t. To demonstrate that $\frac{\partial U^*}{\partial \frac{\pi f}{\pi \hat{f}+\hat{\delta}}} < 0$ holds for $t \in (t_1, t_{2,l}]$, we simply need to check it holds at time $t_{2,l}$. Combining equations (22),(23) and (70), we have

$$\lambda = \frac{1}{\frac{\pi f}{\pi \hat{f}+\hat{\delta}}}\frac{N_2 g_c}{N_2\xi+g_c}(e^{g_c(t_{2,l}-t_1)} - e^{-N_2\xi(t_{2,l}-t_1)}) + [\frac{\xi N_2}{(N_2\xi+g_c)}e^{g_c(t_{2,l}-t_1)} + \frac{g_c}{N_2\xi+g_c}e^{-N_2\xi(t_{2,l}-t_1)}] \quad (90)$$

Combining (89) and (90) lead to

$$\frac{\partial U^*}{\partial \frac{\pi f}{\pi \hat{f}+\hat{\delta}}}|_{t_{2,l}} < 0 \Leftrightarrow (\frac{N_2}{\frac{\pi f}{\pi \hat{f}+\hat{\delta}}})^2 \frac{\frac{1}{\lambda-1}\xi g_c}{(N_2\xi+g_c)^2}[e^{g_c(t_{2,l}-t_1)} - (1+(N_2\xi+g_c)(t_{2,l}-t_1))e^{-N_2\xi(t_{2,l}-t_1)}](\frac{\hat{f}}{\hat{f}+\hat{\delta}} - \frac{\pi\hat{f}}{\pi\hat{f}+\hat{\delta}})$$
$$< \frac{1}{\frac{\pi f}{\pi \hat{f}+\hat{\delta}}}\frac{N_2 g_c}{N_2\xi+g_c}(e^{g_c(t_{2,l}-t_1)} - e^{-N_2\xi(t_{2,l}-t_1)}) \Leftarrow (\frac{N_2}{\frac{\pi f}{\pi \hat{f}+\hat{\delta}}})^2 \frac{\frac{1}{\lambda-1}\xi g_c}{(N_2\xi+g_c)^2}(\frac{\hat{f}}{\hat{f}+\hat{\delta}} - \frac{\pi\hat{f}}{\pi\hat{f}+\hat{\delta}}) < \frac{1}{\frac{\pi f}{\pi \hat{f}+\hat{\delta}}}\frac{N_2 g_c}{N_2\xi+g_c}$$
$$\Leftrightarrow \frac{N_2\xi}{N_2\xi+g_c}(\frac{\hat{f}}{\hat{f}+\hat{\delta}} - \frac{\pi\hat{f}}{\pi\hat{f}+\hat{\delta}}) < (\lambda-1)\frac{\pi\hat{f}}{\pi\hat{f}+\hat{\delta}}$$
$$\Leftarrow \frac{\hat{f}}{\hat{f}+\hat{\delta}} < \lambda\frac{\pi\hat{f}}{\pi\hat{f}+\hat{\delta}} \quad (91)$$

The last inequality holds according to our assumption. It must be satisfied to ensure that mismatch is not so severe that inexperienced workers never move into industry 2. Similarly, by using $U^*(t_{2,l}) = U(t_{2,l})$, we get

$$\frac{\partial U(t_{2,l})}{\partial \frac{\pi \hat{f}}{\pi \hat{f}+\hat{\delta}}} = \frac{\frac{\partial U^*}{\partial t}|_{t_{2,l}}\frac{\partial U}{\partial \frac{\pi f}{\pi \hat{f}+\hat{\delta}}}|_{t_{2,l}} - \frac{\partial U^*}{\partial \frac{\pi f}{\pi \hat{f}+\hat{\delta}}}|_{t_{2,l}}\frac{\partial U}{\partial t}|_{t_{2,l}}}{\frac{\partial U^*}{\partial t}|_{t_{2,l}} - \frac{\partial U}{\partial t}|_{t_{2,l}}} < 0 \quad (92)$$

Aggregate unemployment rate reaches a larger peak value with greater mismatch. We know from equations (86) and (87) that $\frac{\partial L^l}{\partial \frac{\pi f}{\pi \hat{f}+\hat{\delta}}} > 0$, and from (82) that $\frac{\partial E_2^l}{\partial \frac{\pi f}{\pi \hat{f}+\hat{\delta}}} < 0$. Thus, when π decreases, L^l shifts downward and E_2^l shifts upward. the expansion duration $t_{2,l} - t_1$ for (70) to hold becomes shorter. The time to reach the peak value of the aggregate unemployment rate moves forward.

For $\pi' < \pi$, assume the new expansion duration is $t'_{2,l} - t'_1$. From the discussion above, we have $U^{*\prime}(t_{2,l} - t_1 + t'_1) = U'(t_{2,l} - t_1 + t'_1)$ as $t'_{2,l} - t'_1 < t_{2,l} - t_1$. It follows that

$$U'(t_{2,l} - t_1 + t'_1) - U(t_{2,l}) = U^{*\prime}(t_{2,l} - t_1 + t'_1) - U^*(t_{2,l}) \tag{93}$$

$U^*(t)$ is continuously differentiable with respect to t. Combining with (93), the inequality $\frac{\partial U^*}{\partial \frac{\pi \hat{f}}{\pi \hat{f} + \hat{\delta}}}|_{t_{2,l}} < 0$ implies that $U'(t_{2,l} - t_1 + t'_1) > U(t_{2,l})$. That is, aggregate unemployment rate is still larger after the same waiting time $t_{2,l} - t_1$ when π gets smaller. Then aggregate unemployment rate decreases exponentially at a 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 whole path, aggregate unemployment rate shifts upward.

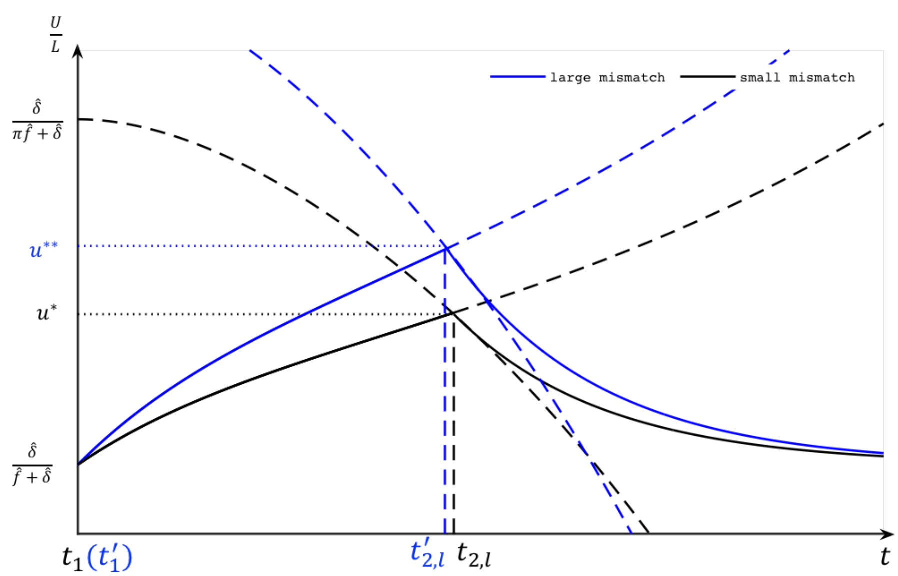


Figure B.3: How unemployment rate changes with the degree of mismatch

B.6 Supplementary proof of infinite-industry model

Optimal Steady State: For the optimal steady state, solve the following static optimization problem:

$$\max \int_{i\in[0,1]} \lambda^{I(i)} s(I(i),i)di \tag{94}$$

$$s.t. \int_{i\in[0,1]} \alpha(I(i))s(I(i),i)di \leq \Omega \tag{95}$$

where $\alpha(I(i)) = 0$ if $I(i) = 0$, and $\alpha(I(i)) = a^{I(i)}$ if $I(i) > 0$. $I(i)$ indicates that individual i choose industry $I(i)$ to work in. $s(I(i),i)$ takes the value $\frac{\widehat{f}}{\widehat{f}+\widehat{\delta}}$ if i is experienced in industry $I(i)$, or it takes the value $\frac{\pi\widehat{f}}{\pi\widehat{f}+\widehat{\delta}}$ if i is inexperienced in industry $I(i)$. So (94) is the final output in consumption good sector, and (95) is the constraint for the capital flows Ω. Using the multiplier μ_Ω, we have

$$\begin{aligned} L &= \int_{i\in[0,1]} \lambda^{I(i)} s(I(i),i)di - \mu_\Omega \int_{i\in[0,1]} \alpha(I(i))s(I(i),i)di + \mu_\Omega\Omega \\ &= \int_{i\in[0,1]} (\lambda^{I(i)} - \mu_\Omega \alpha(I(i)))s(I(i),i)di + \mu_\Omega\Omega \end{aligned}$$

The first order condition with respect to $I(i)$ becomes

$$I(i) \in \text{argmax}\{(\lambda^{I(i)} - \mu_\Omega \alpha(I(i)))s(I(i),i)\}$$

(1) For those who are inexperienced in industry n, $n \geq 1$, and experienced in industry $n+1$, it's strictly better to allocate them in industry $n+1$ than n if and only if

$$(\lambda^n - \mu_\Omega a^n)\frac{\pi\widehat{f}}{\pi\widehat{f}+\widehat{\delta}} < (\lambda^{n+1} - \mu_\Omega a^{n+1})\frac{\widehat{f}}{\widehat{f}+\widehat{\delta}} \Leftrightarrow \mu_\Omega < (\frac{\lambda}{a})^n \frac{\lambda\frac{\widehat{f}}{\widehat{f}+\widehat{\delta}} - \frac{\pi\widehat{f}}{\pi\widehat{f}+\widehat{\delta}}}{a\frac{\widehat{f}}{\widehat{f}+\widehat{\delta}} - \frac{\pi\widehat{f}}{\pi\widehat{f}+\widehat{\delta}}} \equiv \mu_n^{LH}$$

(2) For those who are experienced (inexperienced) in industry n, $n \geq 1$, and experienced (inexperienced) in industry $n+1$, it's strictly better to allocate them in industry $n+1$ than n if and only if

$$(\lambda^n - \mu_\Omega a^n) < (\lambda^{n+1} - \mu_\Omega a^{n+1}) \Leftrightarrow \mu_\Omega < (\frac{\lambda}{a})^n \frac{\lambda-1}{a-1} \equiv \mu_n^{LL}$$

(3) For those who are experienced in industry n, $n \geq 1$, and inexperienced in industry $n+1$, it's strictly better to allocate them in industry $n+1$ than n if and only if

$$(\lambda^n - \mu_\Omega a^n)\frac{\widehat{f}}{\widehat{f}+\widehat{\delta}} < (\lambda^{n+1} - \mu_\Omega a^{n+1})\frac{\pi\widehat{f}}{\pi\widehat{f}+\widehat{\delta}} \Leftrightarrow \mu_\Omega < (\frac{\lambda}{a})^n \frac{\lambda\frac{\pi\widehat{f}}{\pi\widehat{f}+\widehat{\delta}} - \frac{\widehat{f}}{\widehat{f}+\widehat{\delta}}}{a\frac{\pi\widehat{f}}{\pi\widehat{f}+\widehat{\delta}} - \frac{\widehat{f}}{\widehat{f}+\widehat{\delta}}} \equiv \mu_n^{HL}$$

(4) It's strictly better to allocate workers in industry 1 than industry 0 if and only if

$$1 < \lambda - a\mu_\Omega \Leftrightarrow \mu_\Omega < \frac{\lambda - 1}{a} \equiv \mu_0$$

With our assumption 2, we have

$$\begin{cases} \mu_{n+1}^{LH} < \mu_n^{HL}, \mu_n^{HL} < \mu_n^{LL} < \mu_n^{LH}, \forall n \geq 1 \\ \mu_1^{LL} < \mu_0 \end{cases} \tag{96}$$

From (96), by induction we know that if $\mu_n^{HL} \leq \mu_\Omega \leq \mu_n^{LH}$,

$$(\lambda^n - \mu_\Omega a^n)\frac{\pi\widehat{f}}{\pi\widehat{f}+\widehat{\delta}} > (\lambda^i - \mu_\Omega a^i)\frac{\widehat{f}}{\widehat{f}+\widehat{\delta}}, \forall i < n$$

$$(\lambda^{n+1} - \mu_\Omega a^{n+1})\frac{\pi\widehat{f}}{\pi\widehat{f}+\widehat{\delta}} > (\lambda^i - \mu_\Omega a^i)\frac{\widehat{f}}{\widehat{f}+\widehat{\delta}} \forall i > n+1$$

For $\mu_n^{HL} \leq \mu_\Omega \leq \mu_n^{LH}$, all workers stay in industry n or $n+1$. Then the solution to this static problem can be summarized as follows:

(1) if $\mu_\Omega > \mu_0$: all workers stay in industry 0.

(2) if $\mu_\Omega = \mu_0$: workers stay in industry 0 and industry 1.

(3) if $\mu_1^{LL} < \mu_\Omega < \mu_0$: all workers stay in industry 1.

(4) For $n \geq 2$, if $\mu_\Omega = \mu_n^{LH}$: experienced workers in industry n or inexperienced workers in industry $n+1$ stay in industry n. Workers who are inexperienced in industry n and experienced in industry $n+1$ stay in industry n and industry $n+1$.

(5) For $n \geq 2$, if $\mu_n^{LL} < \mu_\Omega < \mu_n^{LH}$: experienced workers in industry n or inexperienced workers in industry $n+1$ stay in industry n. Workers who are inexperienced in industry n and experienced in industry $n+1$ stay in industry $n+1$.

(6) For $n \geq 1$, if $\mu_\Omega = \mu_n^{LL}$: workers of the same experience type in industry n and $n+1$ stay in industry n and industry $n+1$. Workers who are inexperienced in industry n and experienced in industry $n+1$ stay in industry $n+1$. Workers who are experienced in industry n and inexperienced in industry $n+1$ stay in industry n.

(7) For $n \geq 1$, if $\mu_n^{HL} < \mu_\Omega < \mu_n^{LL}$: workers who are experienced in industry n and inexperienced in industry $n+1$ stay in industry n. Others stay in industry $n+1$.

(8) For $n \geq 1$, if $\mu_\Omega = \mu_n^{HL}$: workers who are experienced in industry n and inexperienced in industry $n+1$ stay in industry n and industry $n+1$. Others stay in industry $n+1$.

(9) For $n \geq 1$, if $\mu_{n+1}^{LH} < \mu_\Omega < \mu_n^{HL}$: all workers stay in industry $n+1$.

Let $L_{(n,n+1)}^{(i,j)}$ denote the number of workers who are i-type of experience in industry n and j-type of experience in industry $n+1$. The function of required capital flow $\Omega(C, \{L_{(n,n+1)}^{(i,j)}\})$ to produce output C becomes

$$\Omega(C, \{L_{(n,n+1)}^{(i,j)}\}) = \begin{cases} \frac{a^{n+1}\frac{\widehat{f}}{\widehat{f}+\widehat{\delta}} - a^n \frac{\pi\widehat{f}}{\pi\widehat{f}+\widehat{\delta}}}{\lambda^{n+1}\frac{\widehat{f}}{\widehat{f}+\widehat{\delta}} - \lambda^n \frac{\pi\widehat{f}}{\pi\widehat{f}+\widehat{\delta}}}(C - C_{(n,n+1)}^0) + E_{(n,n+1)}^0, \text{ if } C \in [C_{(n,n+1)}^0, C_{(n,n+1)}^1) \\ \frac{a^{n+1}-a^n}{\lambda^{n+1}-\lambda^n}(C - C_{(n,n+1)}^1) + E_{(n,n+1)}^1, \text{ if } C \in [C_{(n,n+1)}^1, C_{(n,n+1)}^2) \\ \frac{a^{n+1}\frac{\pi\widehat{f}}{\pi\widehat{f}+\widehat{\delta}} - a^n \frac{\widehat{f}}{\widehat{f}+\widehat{\delta}}}{\lambda^{n+1}\frac{\pi\widehat{f}}{\pi\widehat{f}+\widehat{\delta}} - \lambda^n \frac{\widehat{f}}{\widehat{f}+\widehat{\delta}}}(C - C_{(n,n+1)}^2) + E_{(n,n+1)}^2, \text{ if } C \in [C_{(n,n+1)}^2, C_{(n+1,n+2)}^0) \end{cases} \tag{97}$$

where

$$C^0_{(n,n+1)} \equiv \lambda^n [\frac{\widehat{f}}{\widehat{f}+\widehat{\delta}}(L^{(h,l)}_{(n,n+1)} + L^{(h,h)}_{(n,n+1)}) + \frac{\pi\widehat{f}}{\pi\widehat{f}+\widehat{\delta}}(L^{(l,l)}_{(n,n+1)} + L^{(l,h)}_{(n,n+1)})],$$
$$E^0_{(n,n+1)} \equiv a^n [\frac{\widehat{f}}{\widehat{f}+\widehat{\delta}}(L^{(h,l)}_{(n,n+1)} + L^{(h,h)}_{(n,n+1)}) + \frac{\pi\widehat{f}}{\pi\widehat{f}+\widehat{\delta}}(L^{(l,l)}_{(n,n+1)} + L^{(l,h)}_{(n,n+1)})]$$
$$C^1_{(n,n+1)} \equiv \lambda^n [\frac{\widehat{f}}{\widehat{f}+\widehat{\delta}}(L^{(h,l)}_{(n,n+1)} + L^{(h,h)}_{(n,n+1)} + \lambda L^{(l,h)}_{(n,n+1)}) + \frac{\pi\widehat{f}}{\pi\widehat{f}+\widehat{\delta}}L^{(l,l)}_{(n,n+1)}]$$
$$E^1_{(n,n+1)} \equiv a^n [\frac{\widehat{f}}{\widehat{f}+\widehat{\delta}}(L^{(h,l)}_{(n,n+1)} + L^{(h,h)}_{(n,n+1)} + a L^{(l,h)}_{(n,n+1)}) + \frac{\pi\widehat{f}}{\pi\widehat{f}+\widehat{\delta}}L^{(l,l)}_{(n,n+1)}]$$
$$C^2_{(n,n+1)} \equiv \lambda^n [\frac{\widehat{f}}{\widehat{f}+\widehat{\delta}}(L^{(h,l)}_{(n,n+1)} + \lambda L^{(h,h)}_{(n,n+1)} + \lambda L^{(l,h)}_{(n,n+1)}) + \frac{\pi\widehat{f}}{\pi\widehat{f}+\widehat{\delta}}\lambda L^{(l,l)}_{(n,n+1)}]$$
$$E^2_{(n,n+1)} \equiv a^n [\frac{\widehat{f}}{\widehat{f}+\widehat{\delta}}(L^{(h,l)}_{(n,n+1)} + a L^{(h,h)}_{(n,n+1)} + a L^{(l,h)}_{(n,n+1)}) + \frac{\pi\widehat{f}}{\pi\widehat{f}+\widehat{\delta}}a L^{(l,l)}_{(n,n+1)}]$$

The Dynamic Path: Like the benchmark model, when ξ is small, the necessary conditions are

$$\dot{\mu} = (\rho - A)\mu,$$
$$C \approx \hat{C} \in \text{argmax}\{\frac{C^{1-\sigma}-1}{1-\sigma} + \mu(AK - \Omega(C, \{L^{(i,j)}_{(n,n+1)}\})\}$$

where μ is the multiplier for the capital evolution equation. Define

$$M_{0,n} \equiv \frac{\lambda^{n+1}\frac{\pi\widehat{f}}{\pi\widehat{f}+\widehat{\delta}} - \lambda^n \frac{\widehat{f}}{\widehat{f}+\widehat{\delta}}}{a^{n+1}\frac{\pi\widehat{f}}{\pi\widehat{f}+\widehat{\delta}} - a^n \frac{\widehat{f}}{\widehat{f}+\widehat{\delta}}}, \quad M_{1,n} \equiv \frac{\lambda^{n+1} - \lambda^n}{a^{n+1} - a^n}$$

There are four stages for industrial upgrading from industry n to industry $n+1$:

State I: All workers stay in industry n.

For $\mu \in ((C^0_{(n,n+1)})^{-\sigma} M_{1,n}, (C^0_{(n,n+1)})^{-\sigma} M_{0,n-1}]$,

$$\hat{C}^{-\sigma} - \mu \frac{\partial\Omega(\hat{C}, L_l, L - L_l)}{\partial C_-}|_{\hat{C}=C^0_{(n,n+1)}} \geq 0,\ \hat{C}^{-\sigma} - \mu \frac{\partial\Omega(\hat{C}, L_l, L - L_l)}{\partial C_+}|_{\hat{C}=C^0_{(n,n+1)}} \leq 0.$$

At this stage, all workers search in industry n, and there is no industrial upgrading in this stage while capital is accumulated. Inexperienced workers in industry n gradually become experienced workers by on-the-job learning, so consumption grows at a speed lower than the constant rate g_c. As all workers are initially inexperienced in industry $n+1$, $L^{(l,h)}_{(n,n+1)} = 0$. Therefore

$$C(t) = C^0_{(n,n+1)}(t) = C^1_{(n,n+1)}(t)$$

Let L^i_n denote the number of workers who are i-type of experience in industry n and also stay in industry n. Let t^{all}_n denote the starting time of this stage: $\mu(t^{all}_n) = (C^0_{(n,n+1)})^{-\sigma} M_{0,n-1}$. Let t^l_{n+1} denote the ending time of this stage: $\mu(t^l_{n+1}) = (C^0_{(n,n+1)})^{-\sigma} M_{1,n}$. As $\dot{L}^l_n = -\xi \frac{\pi \widehat{f}}{\pi \widehat{f} + \widehat{\delta}} L^l_n$, we have

$$\begin{cases} L^l_n(t) = L^l_n(t^{all}_n) e^{-\xi \frac{\pi \widehat{f}}{\pi \widehat{f} + \widehat{\delta}}(t - t^{all}_n)} \\ L^h_n(t) = L - L^l_n(t) \\ L^s_k(t) = 0, \forall k \neq n, \forall s \in \{h, l\} \end{cases} \tag{98}$$

State II: Inexperienced workers in industry n search in industry $n+1$.

For $\mu \in ((C^2_{(n,n+1)})^{-\sigma} M_{1,n}, (C^1_{(n,n+1)})^{-\sigma} M_{1,n}]$,

$$\hat{C}^{-\sigma} = \mu / M_{1,n} \Rightarrow \frac{\dot{\hat{C}}}{\hat{C}} = \frac{A - \rho}{\sigma} = g_c \ \Rightarrow \ \hat{C}(t) = \hat{C}(t^l_{n+1}) e^{g_c (t - t^l_{n+1})} \tag{99}$$

From our definitions, we have

$$\dot{L}^{(l,l)}_{(n,n+1)} = -\xi \frac{\pi \widehat{f}}{\pi \widehat{f} + \widehat{\delta}} L^{(l,l)}_{(n,n+1)} \Rightarrow L^{(l,l)}_{(n,n+1)}(t) = L^{(l,l)}_{(n,n+1)}(t^l_{n+1}) e^{-\xi \frac{\pi \widehat{f}}{\pi \widehat{f} + \widehat{\delta}}(t - t^l_{n+1})}, \tag{100}$$

and

$$\dot{L}^{(l,h)}_{(n,n+1)} = \xi \frac{\pi \widehat{f}}{\pi \widehat{f} + \widehat{\delta}} L^l_{n+1} = \xi \frac{\pi \widehat{f}}{\pi \widehat{f} + \widehat{\delta}} \left(\frac{C - C^1_{(n,n+1)}}{\lambda^{n+1} \frac{\pi \widehat{f}}{\pi \widehat{f} + \widehat{\delta}} - \lambda^n \frac{\pi \widehat{f}}{\pi \widehat{f} + \widehat{\delta}}} \right) \tag{101}$$

$$C^1_{(n,n+1)} = \lambda^n \frac{\widehat{f}}{\widehat{f} + \widehat{\delta}} (L - L^{(l,l)}_{(n,n+1)} - L^{(l,h)}_{(n,n+1)}) + \lambda^n \frac{\pi \widehat{f}}{\pi \widehat{f} + \widehat{\delta}} L^{(l,l)}_{(n,n+1)} + \lambda^{n+1} \frac{\widehat{f}}{\widehat{f} + \widehat{\delta}} L^{(l,h)}_{(n,n+1)} \tag{102}$$

Combining equations (99) - (102) and

$$\hat{C}(t^l_{n+1}) = \lambda^n(\tfrac{\pi\widehat{f}}{\pi\widehat{f}+\widehat{\delta}}L^l_n(t^l_{n+1}) + \tfrac{\widehat{f}}{\widehat{f}+\widehat{\delta}}L^h_n(t^l_{n+1})),\ L^{(l,l)}_{(n,n+1)}(t^l_{n+1}) = L^l_n(t^l_{n+1}),$$

we have

$$\begin{aligned}\dot{L}^{(l,h)}_{(n,n+1)} = & \frac{\xi(\frac{\pi\widehat{f}}{\pi\widehat{f}+\widehat{\delta}}L^l_n(t^l_{n+1}) + \frac{\widehat{f}}{\widehat{f}+\widehat{\delta}}L^h_n(t^l_{n+1}))}{\lambda - 1}e^{g_c(t-t^l_{n+1})} - \xi\frac{\widehat{f}}{\widehat{f}+\widehat{\delta}}\frac{L}{\lambda-1} + \\ & \xi\frac{\frac{\widehat{f}}{\widehat{f}+\widehat{\delta}} - \frac{\pi\widehat{f}}{\pi\widehat{f}+\widehat{\delta}}}{\lambda-1}L^l_n(t^l_{n+1})e^{-\xi\frac{\pi\widehat{f}}{\pi\widehat{f}+\widehat{\delta}}(t-t^l_{n+1})} - \xi\frac{\widehat{f}}{\widehat{f}+\widehat{\delta}}L^{(l,h)}_{(n,n+1)}\end{aligned} \tag{103}$$

Solving the differential equation gives

$$\begin{cases} L^l_{n+1}(t) = \frac{\pi\widehat{f}+\widehat{\delta}}{\pi\widehat{f}}[\frac{Z_n g_c}{g_c+\xi\frac{\widehat{f}}{\widehat{f}+\widehat{\delta}}}e^{g_c(t-t^l_{n+1})} - \frac{\frac{\pi\widehat{f}}{\pi\widehat{f}+\widehat{\delta}}L^l_n(t^l_{n+1})}{\lambda-1}e^{-\xi\frac{\pi\widehat{f}}{\pi\widehat{f}+\widehat{\delta}}(t-t^l_{n+1})} - (\frac{\frac{\widehat{f}}{\widehat{f}+\widehat{\delta}}L^h_n(t^l_{n+1})}{\lambda-1} - \frac{\xi Z_n\frac{\widehat{f}}{\widehat{f}+\widehat{\delta}}}{g_c+\xi\frac{\widehat{f}}{\widehat{f}+\widehat{\delta}}})e^{-\xi\frac{\widehat{f}}{\widehat{f}+\widehat{\delta}}(t-t^l_{n+1})}] \\ L^h_{n+1}(t) = \frac{\xi Z_n}{g_c+\xi\frac{\widehat{f}}{\widehat{f}+\widehat{\delta}}}e^{g_c(t-t^l_{n+1})} + \frac{L^l_n(t^l_{n+1})}{\lambda-1}e^{-\xi\frac{\pi\widehat{f}}{\pi\widehat{f}+\widehat{\delta}}(t-t^l_{n+1})} + (\frac{L^h_n(t^l_{n+1})}{\lambda-1} - \frac{\xi Z_n}{g_c+\xi\frac{\widehat{f}}{\widehat{f}+\widehat{\delta}}})e^{-\xi\frac{\widehat{f}}{\widehat{f}+\widehat{\delta}}(t-t^l_{n+1})} - \frac{L}{\lambda-1} \\ L^l_n(t) = e^{-\frac{\pi\widehat{f}}{\pi\widehat{f}+\widehat{\delta}}\xi(t-t^l_{n+1})}L^l_n(t^l_{n+1}) - L^l_{n+1}(t) \\ L^h_n(t) = L - e^{-\frac{\pi\widehat{f}}{\pi\widehat{f}+\widehat{\delta}}\xi(t-t^l_{n+1})}L^l_n(t^l_{n+1}) - L^h_{n+1}(t) \\ L^s_k(t) = 0, \forall k \neq n \text{ and } n+1, \forall s \in \{h,l\} \end{cases}$$

where Z_n is given by

$$Z_n \equiv \frac{\frac{\pi\widehat{f}}{\pi\widehat{f}+\widehat{\delta}}L^l_n(t^l_{n+1}) + \frac{\widehat{f}}{\widehat{f}+\widehat{\delta}}L^h_n(t^l_{n+1})}{\lambda - 1}.$$

Let $t_n^{all\_high}$ denote the ending time of this stage: $\mu(t_n^{all\_high}) = (C^2_{(n,n+1)})^{-\sigma}M_{1,n}$. At this stage, inexperienced workers in industry n move into the new industry $n+1$ and search in both industries, while experienced workers in industry n search only in industry n. With on-the-job learning, a fraction of inexperienced workers in industry n become experienced in industry $n+1$. Consumption grows at the constant rate g_c.

State III: All inexperienced workers in industry n stay in industry $n+1$.

For $\mu \in ((C^2_{(n,n+1)})^{-\sigma} M_{0,n}, (C^2_{(n,n+1)})^{-\sigma} M_{1,n}]$,

$$\hat{C}^{-\sigma} - \mu \frac{\partial \Omega(\hat{C}, L_l, L - L_l)}{\partial C_-}|_{\hat{C}=C^2_{(n,n+1)}} \geq 0, \hat{C}^{-\sigma} - \mu \frac{\partial \Omega(\hat{C}, L_l, L - L_l)}{\partial C_+}|_{\hat{C}=C^2_{(n,n+1)}} \leq 0$$

Experienced workers in industry n do not move into industry $n+1$. This is again due to the jump in marginal cost. All workers that are inexperienced in industry n search in industry $n+1$ and gradually become experienced in industry $n+1$. Let t^h_{n+1} denote the end of this stage: $\mu = (C^2_{(n,n+1)})^{-\sigma} M_{0,n}$. Though there is no industrial upgrading, consumption grows due to the increase in experienced workers, though at a speed lower than g_c:

$$C(t) = C^2_{(n,n+1)}(t)$$

Using $\dot{L}^l_{n+1} = \dot{L}^{(l,l)}_{(n,n+1)} = -\xi \frac{\pi \widehat{f}}{\pi \widehat{f} + \widehat{\delta}} L^l_{n+1}$, we find

$$\begin{cases} L^l_{n+1}(t) = L^l_{n+1}(t_n^{all\_high}) e^{-\xi \frac{\pi \widehat{f}}{\pi \widehat{f} + \widehat{\delta}}(t - t_n^{all\_high})} \\ L^h_{n+1}(t) = L^l_{n+1}(t_n^{all\_high}) + L^h_{n+1}(t_n^{all\_high}) - L^l_{n+1}(t) \\ L^h_n(t) = L^h_n(t_n^{all\_high}), L^l_n(t) = 0 \\ L^s_k(t) = 0, \forall k \neq n \text{ and } n+1, \forall s \in \{h, l\} \end{cases} \tag{104}$$

State IV: Experienced workers in industry n search in industry $n+1$.

For $\mu \in ((C^0_{(n+1,n+2)})^{-\sigma} M_{0,n}, (C^2_{(n,n+1)})^{-\sigma} M_{0,n}]$,

$$\hat{C}^{-\sigma} = \mu / M_{0,n} \Rightarrow \frac{\dot{\hat{C}}}{\hat{C}} = \frac{A - \rho}{\sigma} = g_c \;\Rightarrow\; \hat{C}(t) = \hat{C}(t^h_{n+1}) e^{g_c(t - t^h_{n+1})} \tag{105}$$

As $L^h_{n+1} = L^{(h,h)}_{(n,n+1)} + L^{(l,h)}_{(n,n+1)}$ holds at this stage, rewrite $C^2_{(n,n+1)}$

$$C^2_{(n,n+1)} = \lambda^n \frac{\widehat{f}}{\widehat{f} + \widehat{\delta}} (L - L^h_{n+1} - L^{(l,l)}_{(n,n+1)}) + \lambda^{n+1} \frac{\widehat{f}}{\widehat{f} + \widehat{\delta}} L^h_{n+1} + \lambda^{n+1} \frac{\pi \widehat{f}}{\pi \widehat{f} + \widehat{\delta}} L^{(l,l)}_{(n,n+1)} \tag{106}$$

As all inexperienced workers in industry n stay in industry $n+1$,

$$\dot{L}^{(l,l)}_{(n,n+1)} = -\xi \frac{\pi \widehat{f}}{\pi \widehat{f} + \widehat{\delta}} L^{(l,l)}_{(n,n+1)} \Rightarrow L^{(l,l)}_{(n,n+1)}(t) = L^{(l,l)}_{(n,n+1)}(t^h_{n+1}) e^{-\xi \frac{\pi \widehat{f}}{\pi \widehat{f}+\widehat{\delta}}(t-t^h_{n+1})} \tag{107}$$

The evolution of L^h_{n+1} follows

$$\dot{L}^h_{n+1} = \dot{L}^{(l,h)}_{(n,n+1)} + \dot{L}^{(h,h)}_{(n,n+1)} = \xi \frac{\pi \widehat{f}}{\pi \widehat{f} + \widehat{\delta}} L^l_{n+1} = \xi \frac{\pi \widehat{f}}{\pi \widehat{f} + \widehat{\delta}} (L^{(l,l)}_{(n,n+1)} + \frac{C - C^2_{(n,n+1)}}{\lambda^{n+1} \frac{\pi \widehat{f}}{\pi \widehat{f}+\widehat{\delta}} - \lambda^n \frac{\widehat{f}}{\widehat{f}+\widehat{\delta}}}) \tag{108}$$

Combining (105) -(107) and substituting using

$$\hat{C}(t^h_{n+1}) = \lambda^n \frac{\widehat{f}}{\widehat{f} + \widehat{\delta}} L^h_n(t^h_{n+1}) + \lambda^{n+1} \frac{\widehat{f}}{\widehat{f} + \widehat{\delta}} L^h_{n+1}(t^h_{n+1}) + \lambda^{n+1} \frac{\pi \widehat{f}}{\pi \widehat{f} + \widehat{\delta}} L^l_{n+1}(t^h_{n+1}), \tag{109}$$

we transform the differential equation as

$$\dot{L}^h_{n+1} = \xi Q_n e^{g_c(t-t^h_{n+1})} - \xi N_1 - \xi N_2 L^h_{n+1} \tag{110}$$

where

$$Q_n \equiv \frac{(L^h_n(t^h_{n+1}) \frac{\widehat{f}}{\widehat{f}+\widehat{\delta}} + \lambda L^l_{n+1}(t^h_{n+1}) \frac{\pi \widehat{f}}{\pi \widehat{f}+\widehat{\delta}} + \lambda L^h_{n+1}(t^h_{n+1}) \frac{\widehat{f}}{\widehat{f}+\widehat{\delta}}) \frac{\pi \widehat{f}}{\pi \widehat{f}+\widehat{\delta}}}{\lambda \frac{\pi \widehat{f}}{\pi \widehat{f}+\widehat{\delta}} - \frac{\widehat{f}}{\widehat{f}+\widehat{\delta}}}.$$

The solution to this equation gives

$$\begin{cases} L^l_{n+1}(t) = \frac{\pi \widehat{f}+\widehat{\delta}}{\pi \widehat{f}} [\frac{Q_n g_c}{N_2 \xi + g_c} e^{g_c(t-t^h_{n+1})} - (N_1 + N_2 L^h_{n+1}(t^h_{n+1}) - \frac{Q_n \xi N_2}{g_c + \xi N_2}) e^{-N_2 \xi (t-t^h_{n+1})}] \\ L^h_{n+1}(t) = (L^h_{n+1}(t^h_{n+1}) + \frac{N_1}{N_2} - \frac{Q_n \xi}{N_2 \xi + g_c}) e^{-N_2 \xi (t-t^h_{n+1})} + \frac{Q_n \xi}{N_2 \xi + g_c} e^{g_c(t-t^h_{n+1})} - \frac{N_1}{N_2} \\ L^h_n(t) = L - L^l_{n+1}(t) - L^h_{n+1}(t), L^l_n(t) = 0 \\ L^s_k(t) = 0, \forall k \neq n \text{ and } n+1, \forall s \in \{h, l\} \end{cases}$$

Experienced workers in industry n now move into industry $n+1$. They are initially inexperienced in industry $n+1$ and become experienced through on-the-job learning. Consumption grows at a constant rate g_c. When all experienced workers in industry n search in industry $n+1$ at time t^{all}_{n+1}, the economy goes to the stage I of industrial upgrading from industry $n+1$ to industry $n+2$.

Unemployment:

Combining the expressions of $\{L_n^l(t)\}_{n=2}^{\infty}$ at all stages, we obtain the explicit expression of the aggregate unemployment $U(t)$ shown in (27) and (28). To verify that this is the right solution, and to prove that $U(t)$ is a decreasing function at time $t \in (t_n^{all}, t_{n+1}^h]$ and an increasing function at time $t \in (t_{n+1}^h, t_{n+1}^{all}]$, we need to check the conditions that $\dot{L}_{n+1}^l(t) > 0$ at stage II and stage IV. Using the expressions of $L_{n+1}^l(t)$, we only need to prove that $\dot{L}_{n+1}^l(t_{n+1}^l) > 0$ and $\dot{L}_{n+1}^l(t_{n+1}^h) > 0$. At stage II, $\dot{L}_{n+1}^l(t) > 0$ is equivalent to

$$g_c C(t) > \xi \frac{\pi \widehat{f}}{\pi \widehat{f} + \widehat{\delta}} L_{n+1}^l (\lambda^{n+1} \frac{\widehat{f}}{\widehat{f} + \widehat{\delta}} - \lambda^n \frac{\pi \widehat{f}}{\pi \widehat{f} + \widehat{\delta}}) + \xi \frac{\pi \widehat{f}}{\pi \widehat{f} + \widehat{\delta}} L_n^l (\lambda^n \frac{\widehat{f}}{\widehat{f} + \widehat{\delta}} - \lambda^n \frac{\pi \widehat{f}}{\pi \widehat{f} + \widehat{\delta}})$$

Thus $\dot{L}_{n+1}^l(t_n^{all\_high}) > 0$ is equivalent to

$$g_c C(t_n^{all\_high}) > \xi \frac{\pi \widehat{f}}{\pi \widehat{f} + \widehat{\delta}} L_{n+1}^l (t_n^{all\_high}) (\lambda^{n+1} \frac{\widehat{f}}{\widehat{f} + \widehat{\delta}} - \lambda^n \frac{\pi \widehat{f}}{\pi \widehat{f} + \widehat{\delta}})$$
$$\Rightarrow g_c (L_n^h(t) \frac{\widehat{f}}{\widehat{f} + \widehat{\delta}} + \lambda L_{n+1}^l(t) \frac{\pi \widehat{f}}{\pi \widehat{f} + \widehat{\delta}} + \lambda L_{n+1}^h(t) \frac{\widehat{f}}{\widehat{f} + \widehat{\delta}})|_{t=t_n^{all\_high}} > \xi \frac{\pi \widehat{f}}{\pi \widehat{f} + \widehat{\delta}} L_{n+1}^l(t) (\lambda \frac{\widehat{f}}{\widehat{f} + \widehat{\delta}} - \frac{\pi \widehat{f}}{\pi \widehat{f} + \widehat{\delta}})|_{t=t_n^{all\_high}}$$

Similarly, $\dot{L}_{n+1}^l(t_{n+1}^{all}) > 0$ is equivalent to

$$g_c (\lambda L_{n+1}^l(t) \frac{\pi \widehat{f}}{\pi \widehat{f} + \widehat{\delta}} + \lambda L_{n+1}^h(t) \frac{\widehat{f}}{\widehat{f} + \widehat{\delta}})|_{t=t_{n+1}^{all}} > \xi \frac{\pi \widehat{f}}{\pi \widehat{f} + \widehat{\delta}} L_{n+1}^l(t) (\lambda \frac{\widehat{f}}{\widehat{f} + \widehat{\delta}} - \frac{\widehat{f}}{\widehat{f} + \widehat{\delta}})|_{t=t_{n+1}^{all}}$$

As $\lambda > \frac{\pi \widehat{f} + \widehat{\delta}}{\pi \widehat{f} + \pi \widehat{\delta}}$, it implies that

$$g_c (L_{n+1}^l(t) \frac{\pi \widehat{f}}{\pi \widehat{f} + \widehat{\delta}} + L_{n+1}^h(t) \frac{\widehat{f}}{\widehat{f} + \widehat{\delta}})|_{t=t_{n+1}^{all}} > \xi \frac{\pi \widehat{f}}{\pi \widehat{f} + \widehat{\delta}} L_{n+1}^l(t) (\frac{\widehat{f}}{\widehat{f} + \widehat{\delta}} - \frac{\pi \widehat{f}}{\pi \widehat{f} + \widehat{\delta}})|_{t=t_{n+1}^{all}}$$

Using the expression of $L_{n+1}^l(t)$, we have

$$\dot{L}_{n+1}^l(t_{n+1}^l) > 0 \Leftrightarrow g_c (L_n^l(t) \frac{\pi \widehat{f}}{\pi \widehat{f} + \widehat{\delta}} + L_n^h(t) \frac{\widehat{f}}{\widehat{f} + \widehat{\delta}})|_{t=t_{n+1}^l} > \xi \frac{\pi \widehat{f}}{\pi \widehat{f} + \widehat{\delta}} L_n^l(t) (\frac{\widehat{f}}{\widehat{f} + \widehat{\delta}} - \frac{\pi \widehat{f}}{\pi \widehat{f} + \widehat{\delta}})|_{t=t_{n+1}^l}$$
$$\dot{L}_{n+1}^l(t_{n+1}^h) > 0 \Leftrightarrow g_c (L_n^h(t) \frac{\widehat{f}}{\widehat{f} + \widehat{\delta}} + \lambda L_{n+1}^l(t) \frac{\pi \widehat{f}}{\pi \widehat{f} + \widehat{\delta}} + \lambda L_{n+1}^h(t) \frac{\widehat{f}}{\widehat{f} + \widehat{\delta}})|_{t=t_{n+1}^h} > \xi \frac{\pi \widehat{f}}{\pi \widehat{f} + \widehat{\delta}} L_{n+1}^l(t) (\lambda \frac{\widehat{f}}{\widehat{f} + \widehat{\delta}} - \frac{\widehat{f}}{\widehat{f} + \widehat{\delta}})|_{t=t_{n+1}^h}$$

For the industrial upgrading from industry 1 to industry 2, $L_1^h(t_2^h) = L$ and $L_1^l(t_2^h) = 0$ as workers are all experienced in industry 1. Then it is straightforward to check that $\dot{L}_2^l(t_2^h) > 0$. From the discussions above, $\dot{L}_n^l(t_n^h) > 0$ implies $\dot{L}_n^l(t_n^{all}) > 0$, $\dot{L}_n^l(t_n^{all}) > 0$ implies $\dot{L}_{n+1}^l(t_{n+1}^l) > 0$ and then implies $\dot{L}_{n+1}^l(t_n^{all\_high}) > 0$, and $\dot{L}_{n+1}^l(t_n^{all\_high}) > 0$ implies $\dot{L}^l_{n+1}(t^h_{n+1}) > 0$. Using this induction, we know $\dot{L}^l_{n+1}(t) > 0$ holds at stage II and stage IV in the industrial upgrading from industry n to industry $n+1$.

Life Span Of Industries:

Industry $n+1$ appears at stage II in the industrial upgrading from industry n to $n+1$ when inexperienced workers in industry n move into industry $n+1$. It disappears at the end of stage IV in the industrial upgrading from industry $n+1$ to $n+2$ when all experienced workers in industry $n+1$ move into $n+2$. Thus the life span of industry $n+1$ is $t^{all}_{n+2} - t^l_{n+1}$. $t^{all}_{n+2} - t^l_{n+1} = (t^l_{n+3} - t^l_{n+1}) - (t^l_{n+3} - t^{all}_{n+2})$. We have

$$(t^l_{n+3} - t^l_{n+1}) = \frac{\log(\frac{C^1_{(n+2,n+3)}}{C^1_{(n,n+1)}})}{g_c} \rightarrow \frac{2\log(\lambda)}{g_c}$$

as $L^l_n(t^{all}_n)$ and $L^h_n(t^{all}_n)$ converge given the same structure of our industrial upgrading.

In stage I, the following equation holds,

$$(C^0_{(n+2,n+3)}(t^l_{n+3}))^{-\sigma} M_{1,n+2} = (C^0_{(n+2,n+3)}(t^{all}_{n+2}))^{-\sigma} M_{0,n+1} e^{-(A-\rho)(t^l_{n+3} - t^{all}_{n+2})}$$

When ξ is small, $C^0_{(n+2,n+3)}(t^l_{n+3}) \approx C^0_{(n+2,n+3)}(t^{all}_{n+2})$, and then

$$t^l_{n+3} - t^{all}_{n+2} \approx \frac{\log(\frac{a}{\lambda}\frac{a-1}{\lambda-1}\frac{\lambda\frac{\pi\hat{f}}{\pi\hat{f}+\hat{\delta}} - \frac{\hat{f}}{\hat{f}+\hat{\delta}}}{a\frac{\pi\hat{f}}{\pi\hat{f}+\hat{\delta}} - \frac{\hat{f}}{\hat{f}+\hat{\delta}}})}{A-\rho}$$

In the long run, the life span of industries is given by (26), which is a decreasing function of π.

B.7 Extensions

B.7.1 Productivity

We extend our analysis to allow for different productivity of workers. We show that the cyclicality pattern of the aggregate unemployment rate still holds.

Case I: A simple extension is to assume that the technology to produce the intermediate good x_n for $n \geq 1$ follows:

$$F_n(k, l_h^n, l_l^n) = \lambda^n \min\{\frac{k}{a^n}, l_h^n\} + \delta_L \lambda^n \min\{\frac{k}{\delta_L a^n}, l_l^n\}$$

where $\frac{1}{a} < \delta_L < 1$. l_h^n and l_l^n denote the number of experienced and the inexperienced workers in industry n employed by a particular firm. The inexperienced workers in industry n have lower productivity as $\delta_L < 1$ but their production requires less capital. Industry n is still more capital intensive thant industry $n-1$ for low-skilled workers in industry n as $\delta_L > \frac{1}{a}$. In an instantaneous equilibrium, what matters for the output is the "expected" productivity of the workers. For the experienced workers in industry n, their "expected" productivity is $\frac{\widehat{f}}{\widehat{f}+\widehat{\delta}}$ and for the inexperienced workers in industry n, their "expected" productivity is $\delta_L \frac{\pi \widehat{f}}{\pi \widehat{f}+\widehat{\delta}}$. Producing one unit of intermediate goods in industry n requires constant $\frac{a^n}{\lambda^n}$ units of capital inflow. Let's define $\tilde{\pi}$ as

$$\frac{\tilde{\pi}\widehat{f}}{\tilde{\pi}\widehat{f}+\widehat{\delta}} = \delta_L \frac{\pi\widehat{f}}{\pi\widehat{f}+\widehat{\delta}} \tag{111}$$

Let's define $\tilde{\xi} \equiv \xi/\delta_L$. The equilibrium path of $\{C(t), K(t), \Omega(t), L_{(n,n+1)}^{(i,j)}(t)\}$ for this new economy is the same as their equilibrium path for our benchmark model with mismatch friction $\tilde{\pi}$ and learning rate $\tilde{\xi}$. To check it, we can check that the minimal required capital to produce output C is $\Omega(C, \{L_{(n,n+1)}^{(i,j)}\}; \tilde{\pi})$ and the evolution equations of $L_{(n,n+1)}^{(i,j)}$ become

$$\dot{L}_{(n,n+1)}^{(i,h)} = \tilde{\xi}\frac{\tilde{\pi}\widehat{f}}{\tilde{\pi}\widehat{f}+\widehat{\delta}} L_{(n,n+1)}^{(i,l)}, \forall i \in \{l, h\}$$

The planner's problem is equivalently the same as our benchmark case once we reset the mismatch rate to be $\tilde{\pi}$ and learning rate to be $\tilde{\xi}$. Given the dynamics of labor markets $L_{(n,n+1)}^{(i,j)}(t)$, we prove the cyclicality of the aggregate unemployment rate. But here the aggregate unemployment rate is $U(t) = \frac{\widehat{\delta}}{\widehat{f}+\widehat{\delta}} \sum_{n=0}^{\infty} L_n^h + \frac{\widehat{\delta}}{\pi\widehat{f}+\widehat{\delta}} \sum_{n=0}^{\infty} L_n^l \neq \frac{\widehat{\delta}}{\widehat{f}+\widehat{\delta}} \sum_{n=0}^{\infty} L_n^h + \frac{\widehat{\delta}}{\tilde{\pi}\widehat{f}+\widehat{\delta}} \sum_{n=0}^{\infty} L_n^l$. Compared with the benchmark case, $U(t)$ has relatively smaller fluctuation.

Case II: For an alternative setting of productivity differences, we assume that the production function of intermediate good x_n for $n \geq 1$ follows:

$$F_n(k, l_h^n, l_l^n) = \lambda^n \min\{\frac{k}{a^n}, l_h^n\} + \delta_L \lambda^n \min\{\frac{k}{a^n}, l_l^n\}$$

Now the inexperienced workers require the same amount of capital investment as experienced workers but have lower output $\delta_L \lambda^n$. We define $\tilde{\pi}$ using (111). As in our benchmark setting, we impose the following assumption:

Assumption 3: *The following is satisfied* [18]*:*

$$\text{(i) } f = \kappa \widehat{f}, \delta = \kappa \widehat{\delta},\ \kappa \to \infty;$$

$$\text{(ii) } 0 < A - \rho < \sigma A;$$

$$\text{(iii) } \tilde{\xi} < \tilde{\tilde{\xi}}^{inf};$$

$$\text{(iv) } \delta_L > \frac{\lambda}{a}, \frac{\pi \widehat{f}}{\pi \widehat{f} + \widehat{\delta}} > \max\{\frac{a}{(a\delta_L - \lambda)\lambda + 1}, (\frac{1}{\delta_L} - 1)\} \frac{\widehat{f}}{\widehat{f} + \widehat{\delta}}$$

For the optimal steady state, we follow the same approach and show that the function of required capital flow $\tilde{\Omega}(C, \{L_{(n,n+1)}^{(i,j)}\})$ becomes

$$\tilde{\Omega}(C, \{L_{(n,n+1)}^{(i,j)}\}) = \begin{cases} \frac{a^{n+1}\frac{\widehat{f}}{\widehat{f}+\widehat{\delta}} - a^n \frac{\pi \widehat{f}}{\pi \widehat{f}+\widehat{\delta}}}{\lambda^{n+1}\frac{\widehat{f}}{\widehat{f}+\widehat{\delta}} - \lambda^n \frac{\tilde{\pi}\widehat{f}}{\tilde{\pi}\widehat{f}+\widehat{\delta}}}(C - \tilde{C}_{(n,n+1)}^0) + E_{(n,n+1)}^0, \text{ if } C \in [\tilde{C}_{(n,n+1)}^0, \tilde{C}_{(n,n+1)}^1) \\ \frac{a^{n+1} - a^n}{\lambda^{n+1} - \lambda^n}(C - \tilde{C}_{(n,n+1)}^1) + E_{(n,n+1)}^1, \text{ if } C \in [\tilde{C}_{(n,n+1)}^1, \tilde{C}_{(n,n+1)}^2) \\ \frac{a^{n+1}\frac{\pi \widehat{f}}{\pi \widehat{f}+\widehat{\delta}} - a^n \frac{\widehat{f}}{\widehat{f}+\widehat{\delta}}}{\lambda^{n+1}\frac{\tilde{\pi}\widehat{f}}{\tilde{\pi}\widehat{f}+\widehat{\delta}} - \lambda^n \frac{\widehat{f}}{\widehat{f}+\widehat{\delta}}}(C - \tilde{C}_{(n,n+1)}^2) + E_{(n,n+1)}^2, \text{ if } C \in [\tilde{C}_{(n,n+1)}^2, \tilde{C}_{(n+1,n+2)}^0) \end{cases}$$

where

$$\tilde{C}_{(n,n+1)}^0 \equiv \lambda^n [\frac{\widehat{f}}{\widehat{f} + \widehat{\delta}}(L_{(n,n+1)}^{(h,l)} + L_{(n,n+1)}^{(h,h)}) + \frac{\tilde{\pi}\widehat{f}}{\tilde{\pi}\widehat{f} + \widehat{\delta}}(L_{(n,n+1)}^{(l,l)} + L_{(n,n+1)}^{(l,h)})],$$

$$\tilde{C}_{(n,n+1)}^1 \equiv \lambda^n [\frac{\widehat{f}}{\widehat{f} + \widehat{\delta}}(L_{(n,n+1)}^{(h,l)} + L_{(n,n+1)}^{(h,h)} + \lambda L_{(n,n+1)}^{(l,h)}) + \frac{\tilde{\pi}\widehat{f}}{\tilde{\pi}\widehat{f} + \widehat{\delta}} L_{(n,n+1)}^{(l,l)}]$$

$$\tilde{C}_{(n,n+1)}^2 \equiv \lambda^n [\frac{\widehat{f}}{\widehat{f} + \widehat{\delta}}(L_{(n,n+1)}^{(h,l)} + \lambda L_{(n,n+1)}^{(h,h)} + \lambda L_{(n,n+1)}^{(l,h)}) + \frac{\tilde{\pi}\widehat{f}}{\tilde{\pi}\widehat{f} + \widehat{\delta}} \lambda L_{(n,n+1)}^{(l,l)}]$$

[18] Condition (iv) is sufficient to ensure that $\frac{\lambda}{a} \frac{\lambda \frac{\widehat{f}}{\widehat{f}+\widehat{\delta}} - \frac{\tilde{\pi}\widehat{f}}{\tilde{\pi}\widehat{f}+\widehat{\delta}}}{a \frac{\widehat{f}}{\widehat{f}+\widehat{\delta}} - \frac{\pi \widehat{f}}{\pi \widehat{f}+\widehat{\delta}}} < \frac{\lambda \frac{\tilde{\pi}\widehat{f}}{\tilde{\pi}\widehat{f}+\widehat{\delta}} - \frac{\widehat{f}}{\widehat{f}+\widehat{\delta}}}{a \frac{\pi \widehat{f}}{\pi \widehat{f}+\widehat{\delta}} - \frac{\widehat{f}}{\widehat{f}+\widehat{\delta}}}$.

For the dynamic path, we follow the steps in our benchmark model and show that there are still four stages for industrial upgrading from industry n to industry $n+1$:

(1) When $\mu \in ((\tilde{C}^0_{(n,n+1)})^{-\sigma} M_{1,n}, (\tilde{C}^0_{(n,n+1)})^{-\sigma} \tilde{M}_{0,n-1}]$, all workers stay in industry n, where

$$\tilde{M}_{0,n} \equiv \frac{\lambda^{n+1} \frac{\tilde{\pi}\widehat{f}}{\tilde{\pi}\widehat{f}+\widehat{\delta}} - \lambda^n \frac{\widehat{f}}{\widehat{f}+\widehat{\delta}}}{a^{n+1} \frac{\pi\widehat{f}}{\pi\widehat{f}+\widehat{\delta}} - a^n \frac{\widehat{f}}{\widehat{f}+\widehat{\delta}}}$$

And the equations of L^s_k follow (98).

(2) When $\mu \in ((\tilde{C}^2_{(n,n+1)})^{-\sigma} M_{1,n}, (\tilde{C}^1_{(n,n+1)})^{-\sigma} M_{1,n}]$, inexperienced workers in industry n move into industry $n+1$. The evolution equations of $L^{(l,l)}_{(n,n+1)}$ and $L^{(l,h)}_{(n,n+1)}$ follow:

$$\begin{cases} \dot{L}^{(l,l)}_{(n,n+1)} = -\xi \frac{\pi\widehat{f}}{\pi\widehat{f}+\widehat{\delta}} L^{(l,l)}_{(n,n+1)} \\ \dot{L}^{(l,h)}_{(n,n+1)} = \frac{\xi}{\delta_L}(\tilde{Z}_n e^{g_c(t-t^l_{n+1})} - \frac{\widehat{f}}{\widehat{f}+\widehat{\delta}} \frac{L}{\lambda-1} + \frac{\frac{\widehat{f}}{\widehat{f}+\widehat{\delta}} - \frac{\pi\widehat{f}}{\pi\widehat{f}+\widehat{\delta}}}{\lambda-1} L^l_n(t^l_{n+1}) e^{-\xi \frac{\pi\widehat{f}}{\pi\widehat{f}+\widehat{\delta}}(t-t^l_{n+1})} - \frac{\widehat{f}}{\widehat{f}+\widehat{\delta}} L^{(l,h)}_{(n,n+1)}) \end{cases}$$

Then we have

$$\begin{cases} L^l_{n+1}(t) = \frac{\tilde{\pi}\widehat{f}+\widehat{\delta}}{\tilde{\pi}\widehat{f}}[\tilde{Z}_n \frac{g_c + \frac{\xi}{\delta_L}(1-\delta_L)\frac{\widehat{f}}{\widehat{f}+\widehat{\delta}}}{g_c + \frac{\xi}{\delta_L}\frac{\widehat{f}}{\widehat{f}+\widehat{\delta}}} e^{g_c(t-t^l_{n+1})} + \frac{\frac{\widehat{f}}{\widehat{f}+\widehat{\delta}}(1-\delta_L) - \frac{\tilde{\pi}\widehat{f}}{\tilde{\pi}\widehat{f}+\widehat{\delta}}}{\lambda-1} L^l_n(t^l_{n+1}) e^{-\xi \frac{\pi\widehat{f}}{\pi\widehat{f}+\widehat{\delta}}(t-t^l_{n+1})} - (1-\delta_L)\frac{\widehat{f}}{\widehat{f}+\widehat{\delta}} \frac{L}{\lambda-1} \\ \qquad - \frac{\widehat{f}}{\widehat{f}+\widehat{\delta}}(\frac{L^h_n(t^l_{n+1})\delta_L}{\lambda-1} - \frac{\xi\tilde{Z}_n}{g_c + \frac{\xi}{\delta_L}\frac{\widehat{f}}{\widehat{f}+\widehat{\delta}}}) e^{-\frac{\xi}{\delta_L}\frac{\widehat{f}}{\widehat{f}+\widehat{\delta}}(t-t^l_{n+1})}] \\ L^h_{n+1}(t) = \frac{\xi\tilde{Z}_n}{g_c + \frac{\widehat{f}}{\widehat{f}+\widehat{\delta}}\frac{\xi}{\delta_L}} e^{g_c(t-t^l_{n+1})} + \frac{L^l_n(t^l_{n+1})\delta_L}{\lambda-1} e^{-\xi \frac{\pi\widehat{f}}{\pi\widehat{f}+\widehat{\delta}}(t-t^l_{n+1})} + (\frac{L^h_n(t^l_{n+1})\delta_L}{\lambda-1} - \frac{\xi\tilde{Z}_n}{g_c + \frac{\widehat{f}}{\widehat{f}+\widehat{\delta}}\frac{\xi}{\delta_L}}) e^{-\frac{\xi}{\delta_L}\frac{\widehat{f}}{\widehat{f}+\widehat{\delta}}(t-t^l_{n+1})} - \frac{L\delta_L}{\lambda-1} \end{cases}$$

where $\tilde{Z}_n \equiv \frac{\frac{\tilde{\pi}\widehat{f}}{\tilde{\pi}\widehat{f}+\widehat{\delta}} L^l_n(t^l_{n+1}) + \frac{\widehat{f}}{\widehat{f}+\widehat{\delta}} L^h_n(t^l_{n+1})}{\lambda-1}$. Under our assumption, $\dot{L}^l_{n+1}(t) > 0, \forall t \in [t^l_{n+1}, t^{all\_high}_n)$ if $\dot{L}^l_{n+1}(t^l_{n+1}) > 0$.

(3) When $\mu \in ((\tilde{C}^2_{(n,n+1)})^{-\sigma} \tilde{M}_{0,n}, (\tilde{C}^2_{(n,n+1)})^{-\sigma} M_{1,n}]$, all experienced workers in industry n stay in industry n and all inexperienced workers in industry n stay in industry $n+1$. The equations of L^s_k follow (104).

(4) When $\mu \in ((\tilde{C}^0_{(n+1,n+2)})^{-\sigma} \tilde{M}_{0,n}, (\tilde{C}^2_{(n,n+1)})^{-\sigma} \tilde{M}_{0,n}]$, experienced workers in industry n move into industry $n+1$. And the equations of $L^i_{n+1}(t)$ change into

$$\begin{cases} L^l_{n+1}(t) = \frac{\pi\widehat{f}+\widehat{\delta}}{\pi\widehat{f}}[\frac{\tilde{Q}_n g_c}{\tilde{N}_2\xi+g_c}e^{g_c(t-t^h_{n+1})} - (\tilde{N}_1+\tilde{N}_2 L^h_{n+1}(t^h_{n+1}) - \frac{\tilde{Q}_n\xi\tilde{N}_2}{g_c+\xi\tilde{N}_2})e^{-\tilde{N}_2\xi(t-t^h_{n+1})}] \\ L^h_{n+1}(t) = (L^h_{n+1}(t^h_{n+1}) + \frac{\tilde{N}_1}{\tilde{N}_2} - \frac{\tilde{Q}_n\xi}{\tilde{N}_2\xi+g_c})e^{-\tilde{N}_2\xi(t-t^h_{n+1})} + \frac{\tilde{Q}_n\xi}{\tilde{N}_2\xi+g_c}e^{g_c(t-t^h_{n+1})} - \frac{\tilde{N}_1}{\tilde{N}_2} \end{cases}$$

where

$$\tilde{Q}_n \equiv \frac{(L^h_n(t^h_{n+1})\frac{\widehat{f}}{\widehat{f}+\widehat{\delta}} + \lambda L^l_{n+1}(t^h_{n+1})\frac{\tilde{\pi}\widehat{f}}{\tilde{\pi}\widehat{f}+\widehat{\delta}} + \lambda L^h_{n+1}(t^h_{n+1})\frac{\widehat{f}}{\widehat{f}+\widehat{\delta}})\frac{\pi\widehat{f}}{\pi\widehat{f}+\widehat{\delta}}}{\lambda\frac{\tilde{\pi}\widehat{f}}{\tilde{\pi}\widehat{f}+\widehat{\delta}} - \frac{\widehat{f}}{\widehat{f}+\widehat{\delta}}}$$

$$\tilde{N}_1 \equiv \frac{L\frac{\widehat{f}}{\widehat{f}+\widehat{\delta}}\frac{\pi\widehat{f}}{\pi\widehat{f}+\widehat{\delta}}}{\lambda\frac{\tilde{\pi}\widehat{f}}{\tilde{\pi}\widehat{f}+\widehat{\delta}} - \frac{\widehat{f}}{\widehat{f}+\widehat{\delta}}}, \tilde{N}_2 \equiv \frac{(\lambda-1)\frac{\widehat{f}}{\widehat{f}+\widehat{\delta}}\frac{\pi\widehat{f}}{\pi\widehat{f}+\widehat{\delta}}}{\lambda\frac{\tilde{\pi}\widehat{f}}{\tilde{\pi}\widehat{f}+\widehat{\delta}} - \frac{\widehat{f}}{\widehat{f}+\widehat{\delta}}}$$

We still have $\dot{L}^l_{n+1}(t) > 0, \forall t \in [t^h_{n+1}, t^{all}_{n+1})$ if $\dot{L}^l_{n+1}(t^h_{n+1}) > 0$. To verify this is the right solution and the cyclicality of the aggregate unemployment rate, we use again the induction method. As all workers are assumed to be experienced in industry 1, it's straightforward to check that $\dot{L}^l_2(t^h_2) > 0$. Then it implies $\dot{L}^l_2(t^{all}_2) > 0$ or

$$g_c(\lambda L^l_2(t)\frac{\tilde{\pi}\widehat{f}}{\tilde{\pi}\widehat{f}+\widehat{\delta}} + \lambda L^h_2(t)\frac{\widehat{f}}{\widehat{f}+\widehat{\delta}})|_{t=t^{all}_2} > \xi\frac{\pi\widehat{f}}{\pi\widehat{f}+\widehat{\delta}}L^l_2(t)(\lambda\frac{\widehat{f}}{\widehat{f}+\widehat{\delta}} - \frac{\widehat{f}}{\widehat{f}+\widehat{\delta}})|_{t=t^{all}_2}$$

which implies that

$$g_c(L^l_2(t)\frac{\tilde{\pi}\widehat{f}}{\tilde{\pi}\widehat{f}+\widehat{\delta}} + L^h_2(t)\frac{\widehat{f}}{\widehat{f}+\widehat{\delta}})|_{t=t^{all}_2} > \xi\frac{\pi\widehat{f}}{\pi\widehat{f}+\widehat{\delta}}L^l_2(t)(\frac{\widehat{f}}{\widehat{f}+\widehat{\delta}} - \frac{\tilde{\pi}\widehat{f}}{\tilde{\pi}\widehat{f}+\widehat{\delta}})|_{t=t^{all}_2}$$

Then it implies that $\dot{L}^l_3(t^l_3) > 0$ as it is equivalent to

$$g_c(L^l_2(t)\frac{\tilde{\pi}\widehat{f}}{\tilde{\pi}\widehat{f}+\widehat{\delta}} + L^h_2(t)\frac{\widehat{f}}{\widehat{f}+\widehat{\delta}})|_{t=t^l_3} > \xi\frac{\pi\widehat{f}}{\pi\widehat{f}+\widehat{\delta}}L^l_2(t)(\frac{\widehat{f}}{\widehat{f}+\widehat{\delta}} - \frac{\tilde{\pi}\widehat{f}}{\tilde{\pi}\widehat{f}+\widehat{\delta}})|_{t=t^l_3}$$

Then it implies $\dot{L}^l_3(t^{all\_high}_2) > 0$ or

$$g_c(L^h_2(t)\frac{\widehat{f}}{\widehat{f}+\widehat{\delta}} + \lambda L^l_3(t)\frac{\tilde{\pi}\widehat{f}}{\tilde{\pi}\widehat{f}+\widehat{\delta}} + \lambda L^h_3(t)\frac{\widehat{f}}{\widehat{f}+\widehat{\delta}})|_{t=t^{all\_high}_2} > \xi\frac{\pi\widehat{f}}{\pi\widehat{f}+\widehat{\delta}}L^l_3(t)(\lambda\frac{\widehat{f}}{\widehat{f}+\widehat{\delta}} - \frac{\tilde{\pi}\widehat{f}}{\tilde{\pi}\widehat{f}+\widehat{\delta}})|_{t=t^{all\_high}_2}$$

which implies $\dot{L}^l_3(t^h_3) > 0$ or

$$g_c(L_2^h(t)\frac{\widehat{f}}{\widehat{f}+\widehat{\delta}}+\lambda L_3^l(t)\frac{\tilde{\pi}\widehat{f}}{\tilde{\pi}\widehat{f}+\widehat{\delta}}+\lambda L_3^h(t)\frac{\widehat{f}}{\widehat{f}+\widehat{\delta}})|_{t=t_3^h} > \xi\frac{\pi\widehat{f}}{\pi\widehat{f}+\widehat{\delta}}L_3^l(t)(\lambda\frac{\widehat{f}}{\widehat{f}+\widehat{\delta}}-\frac{\widehat{f}}{\widehat{f}+\widehat{\delta}})|_{t=t_3^h}$$

From $\dot{L}_3^l(t_3^h) > 0$, we repeat the above analysis and so on and so forth. It shows again that the pattern of labor market dynamics is the same as our benchmark case without productivity difference. Then it immediately verifies the cyclical aggregate unemployment rate.

B.7.2 Matching function

This section shows that the assumption of a constant job finding rate can be rationalized as an equilibrium outcome in the Diamond-Mortensen-Pissarides model setting. Suppose that the matching function is concave and constant returns to scale in each industry, we prove that the job finding rate is unique and constant during the industry upgrading process. Consider an economy with a unit mass of identical households. Each household is initially endowed with capital K, inexperienced labor L^l and experienced labor L^h. There are two sectors, one producing capital goods and another producing consumption goods, and their production functions are the same as in the benchmark model. Labor markets are frictional: new matches between i-experience-typed unemployed workers in industry j and job vacancies in that industry are determined by the matching function $A_j^i m(U_j^i, V_j^i)$, where U_j^i is the number of i-experience-typed unemployed workers searching in industry j, and V_j^i is the number of job vacancies for i-experience-typed workers in industry j. The function $m(\cdot)$ is strictly increasing and strictly concave in both arguments, and it is constant return to scale. A_j^i denotes the matching efficiency. Assume that $A_j^i = A$ if $i \neq l$ or $j \neq 2$ and $A_2^l = \pi A$. That is, matching efficiency is lower for inexperienced 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 experience type i in industry j. Matches are destroyed exogenously at rate δ in all industries. Posting vacancies is costly, and costs c in terms of the final consumption good in any industry for all workers independent of experienced types. The planner's problem becomes

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

subject to

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

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 $\kappa \to \infty$. To find an optimal steady state, the static optimization becomes

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

subject to

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

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

$$K_1 + K_2 + K_3 \leq \Omega.$$

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

Proposition 10. *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 when industry j and industry $j+1$ coexist.*

Proof. With some abuse of notation, set the Lagrange as follows

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

$$+\mu_k(\Omega - a \sum_{i\in\{l,h\}} E^i - a^2 \sum_{i\in\{l,h\}} E_2^i - K_3) + \sum_{i\in\{l,h\}} \mu_i (L_i - \sum_{j=0}^{2} (E_j^i + U_j^i))$$

$$+ \sum_{i\in\{l,h\}} \sum_{j=0}^{2} \mu_j^i (\widehat{A}_j^i q(\theta_j^i) U_j^i - \widehat{\delta} E_j^i)$$

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

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

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

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

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

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

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

Combining these conditions, we get

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

The strict 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 0 is uniquely determined and constant. In terms of industry 1, we have

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

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

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

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

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

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

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 experienced workers in industry 1 and 2. For inexperienced workers, more specifically, experienced in industry 1 and inexperienced in industry 2, we can similarly prove $\theta_1^l = \theta_2^l$, and derive the following equation

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

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 inexperienced workers in industry 2 is π times that of inexperienced workers in industry 1 because of mismatch. Finally, it is straightforward to check that in both cases, 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.

The intuition is as follows. The planner optimally assigns the number of vacancies for each worker who moves from industry n to $n+1$ 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 the industrial upgrading process. Therefore, the job finding rate is endogenously constant, consistent with the benchmark model. What varies is the share of workers in industry n and in $n+1$. With a larger E, workers move from industry n to $n+1$ to absorb extra capital investments and produce more consumption goods.

■

B.8 Numerical Solution

For the four-industry model, the first order condition of the planner's problem is

$$C \in arg \max\{\frac{C^{1-\sigma}-1}{1-\sigma} + \mu(AK - \Omega(C, L^l, L-L^l)) - \Lambda G(C, L^l, L-L^l)\}$$

We approximate it as follows

$$C \approx \hat{C} \in arg \max\{\frac{C^{1-\sigma}-1}{1-\sigma} + \mu(AK - \Omega(C, L^l, L-L^l))\}$$

Using this approximation, we find a closed-form solution of the equilibrium path. Alternatively, we can solve the planner's problem numerically by using the shooting algorithm. To show that our analytical solution is close to the right solution, we consider an extreme case in this section where we set the on-the-job learning rate $\xi = 2.0$ which is 20 times the consumption growth rate g_c. For all other parameters, we set them the same as our benchmark calibration. The results are shown in Figure B.4.

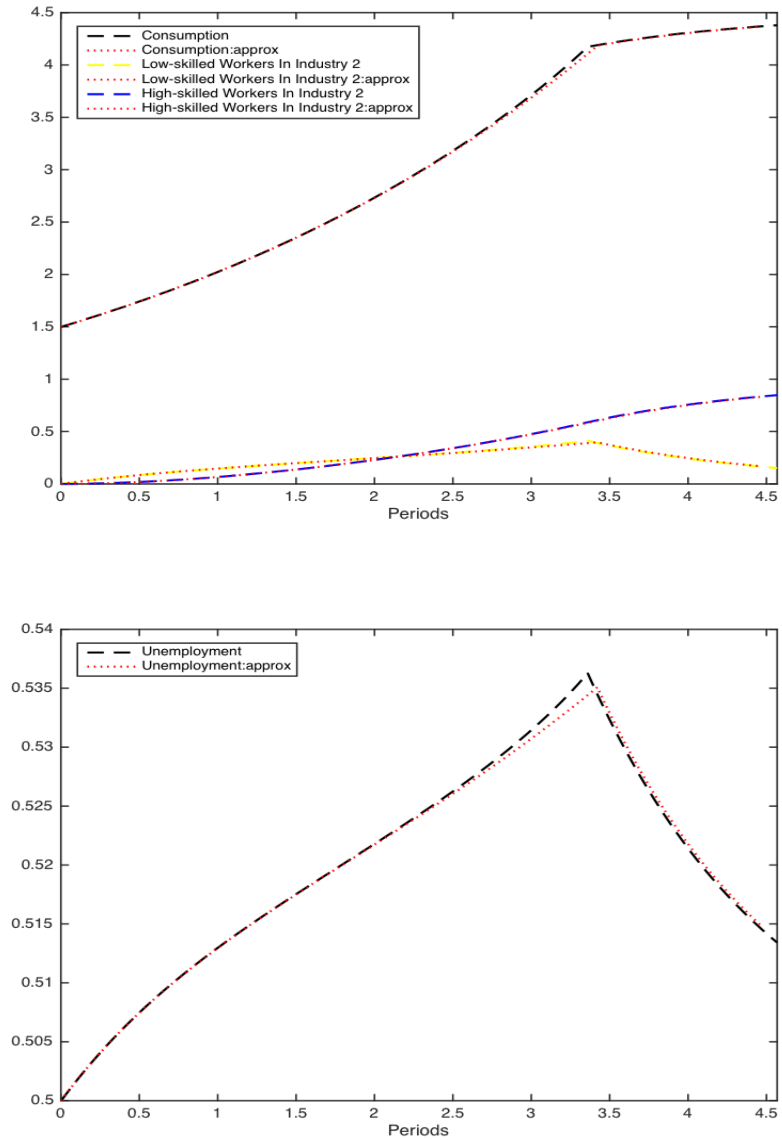


Figure B.4: Labor market dynamics: numerical solution

Even for this large learning rate, we observe little difference between our numerical solution and the analytical one by using the approximation of the first order condition. And the comparative static analysis using the numerical solutions gives the same pattern regarding the effects of learning efficiency, mismatch and capital-goods production efficiency, as shown in Figure B.5.

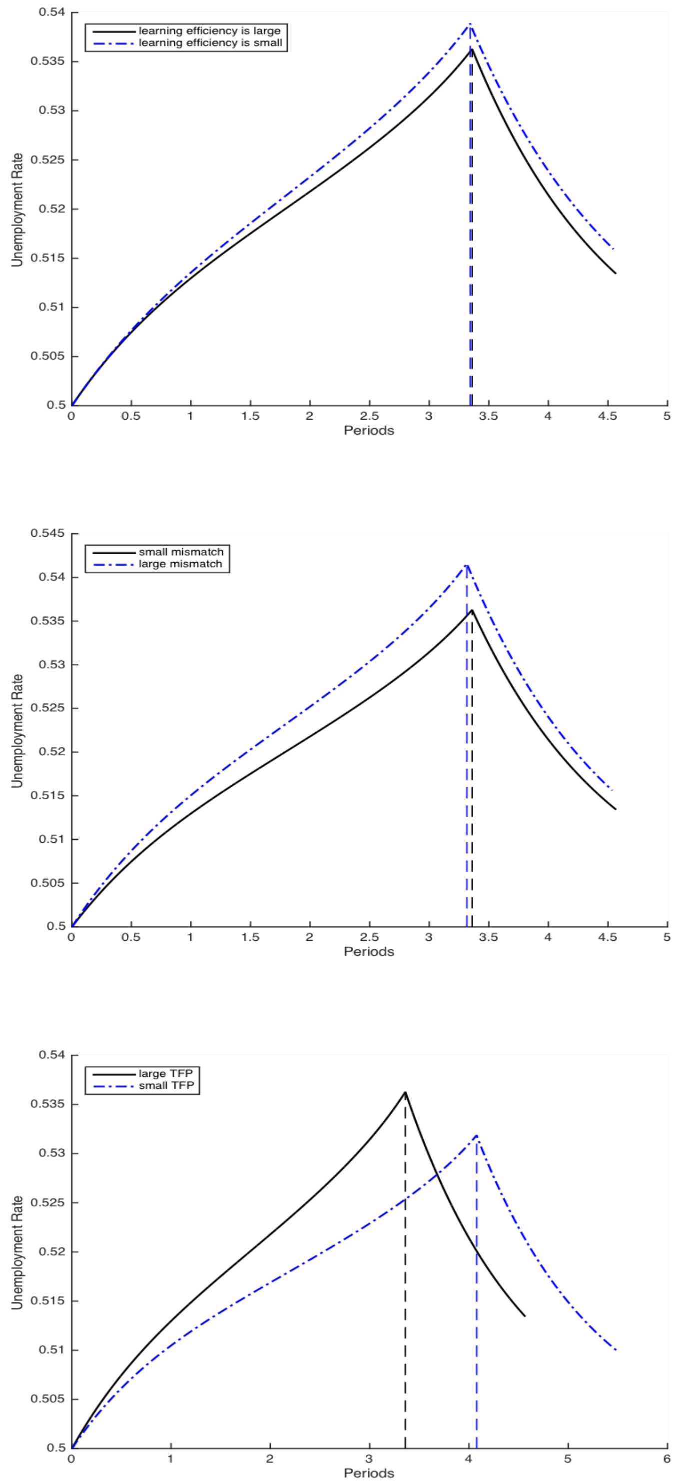


Figure B.5: Comparative static analysis: numerical solution