Labor Market Segmentation and the Dynamics of Entry Barriers: A Theoretical Framework

Yifan Wang

the University of Hong Kong (HKU)

May, 2024

Abstract

This paper examines the dynamics of labor market segmentation characterized by simultaneous high barriers to entry, intense competition, and elevated compensation in certain sectors. We develop a theoretical framework incorporating heterogeneous agents with endowment-based abilities and education-occupation matching through a novel matrix multiplication approach. The model distinguishes between "three-high" sectors featuring employment rationing and conventional sectors with wage-based clearing mechanisms. We analyze how workers navigate these segments through task upgrading and learning-by-doing, demonstrating that education-occupation mismatches can persist while increasing entry barriers. Our framework explains the coexistence of excess qualified labor supply with persistent wage premiums and provides insights into structural changes in labor market equilibria through worker mobility and skill acquisition. The model's implications are particularly relevant for understanding labor market dynamics in Asian economies experiencing "neijuan" (involution) phenomena.

1 Introduction

This paper investigates the dynamic relationship between labor market surplus and endogenous barriers to entry in a multi-sector economy. We demonstrate that education-occupation mismatches, coupled with "neijuan" (involution), generate a counterintuitive equilibrium where excess labor supply increases entry barriers rather than depressing wages. Our theoretical framework incorporates heterogeneous agents and sector-specific human capital accumulation to analyze three key mechanisms: (1) the cumulative effects of education-work mismatches at the macroeconomic level, (2) the emergence of "three-high" sectors characterized by simultaneous high competition, high entry barriers, and high compensation, and (3) the structural transformation of labor market equilibria through endogenous skill upgrading and inter-sectoral mobility. The model's innovative features include sector-specific barriers to entry, directional intra-industry mobility, and endogenous human capital accumulation through task-upgrading and learning-by-doing. Using a combination of static and dynamic analyses, we calibrate the model using industry-level data on wages and educational attainment. This framework enables counterfactual analyses of exogenous shocks and provides insights into the persistence of market segmentation in Asian economies. Our findings contribute to understanding how microeconomic choice mechanisms generate macro-level labor market dynamics and structural change.

Timeline Define the intuitive timeline or the sub-periods for an agent:

1. With endowments across different skills (occupations) settled at birth. [1]

2. With education background chosen, define the before-working ability across industries. Make things simple, assume the monotonicity constraints of education background and initial endowments. [2]

3. Either entering the optimal (most possible) industry or being unemployed due to insufficient positions. *Key question: why positions are insufficient?*

4. Through task upgrading and "learning by doing", workers decide whether to flow to other industries.

[1] I haven't define the key variations of occupations and industries so far due to the misallocation of education and information issue. For instance, if an accountant acquires the bachelor of accounting / finance, end up in Apple's budget department or other industries' same positions, then she is actually in the same sub-labor market.

[2] Why I differ endowments and education background? I wish to introduce misallocation of education, due to luck, families issues and inexplicable error terms, acting as a mechanism of broader education. Say, a talents chooses the wrong education entering the low-income industry, but still able to flow to other industries through "learning by doing".

Timeline The model follows a sequential structure for each agent:

1. **Initial Endowments**: Agents are born with heterogeneous skill endowments;[3]

2. **Education and Ability Formation**: Education background B interacts with endowments to form pre-market abilities $\mathbf{a} = f(E, B)$. The transformation follows monotonicity constraints between education quality and initial endowments[4]

3. **Initial Labor Market Entry**: Agents choose sectors based on expected utility maximization:

   - Entry into preferred sector if position available
   - Unemployment with lump-sum transfer if rejected
   - *Key question: why positions are insufficient?*

4. **Dynamic Adjustment**: In each subsequent period t, agents choose between:

   - Remain in current position
   - Search for new positions in alternative sectors
   - Accept unemployment for skill upgrading and invest in task-specific human capital through "learning by doing"

5. **Skill Evolution**: Abilities evolve according to some rules:

[3] I haven't define the key variations of occupations and industries so far due to the misallocation of education and information issue. For instance, if an accountant acquires the bachelor of accounting / finance, end up in Apple's budget department or other industries' same positions, then she is actually in the same sub-labor market.

[4] Why I differ endowments and education background? I wish to introduce misallocation of education, due to luck, families issues and inexplicable error terms, acting as a mechanism of broader education. Say, a talents chooses the wrong education entering the low-income industry, but still able to flow to other industries through "learning by doing".

2 The model

2.1 Environment

Agents Define the ability matrix for agent $i \in \mathbf{I}$ as [5]

$$\mathbf{a_i} = \Theta \cdot \mathbf{E_i} + \Phi \cdot \mathbf{B_i} + \epsilon \tag{1}$$

where

- Vector $\mathbf{E}$ denotes the basic personal endowments;
- vector $\mathbf{B}$ captures the education background;
- Θ, Φ are weight matrices capturing the contributions of different endowments/educational backgrounds to industry-specific abilities;
- the stochastic error term ϵ represents the education-occupation mismatch.

Specifically, the individual level ability vector $\mathbf{a_i} = \{a_{i1}, \ldots, a_{il}\}$, where $a_i j$ represents agent i's industry j specific technical competencies and cross-industry skill transferability, is transformed through a *Sigmoid-based transformation* to compute $\tilde{a}_{ij}$:

$$\tilde{a}_{ij} = \frac{M}{1 + e^{-k(a_{ij} - \mu)}} \tag{2}$$

where M is the upper bound, k controls slope, and μ centers the function.

Define a O-ring type aggregator

$$A_O(\mathbf{a_i}) = \left(\prod_{j=1} a_{ij}\right)\left(1 + \sum_{j=1}\sum_{m=1} \gamma_{jm} a_{ij} a_{im}\right) \tag{3}$$

where the first part $\prod_{j=1} a_{ij}$ denotes the product of basic industry level ability, emphasizing the baseline multiplicative complementarity across all abilities. The second term $\left(1 + \sum_{j=1}\sum_{m=1} \gamma_{jm} a_{ij} a_{im}\right)$ captures pairwise synergies between specific ability combinations, with $\gamma_{ij} \in [0, 1]$ measuring the strength of complementarity between abilities i and j. [6]

[5]My ability formation mechanism fundamentally aligns with the task-based approach pioneered by Acemoglu & Autor (2011), though employing different mathematical formulations. While they use a continuous task space with integration over task-specific efficiencies $\pi_{ij(t)}$, our model adopts a discrete matrix transformation framework where industry-specific abilities are determined by the interaction between endowment vector $\mathbf{E}$ and education background $\mathbf{B}$ through weight matrices. This matrix-based approach maintains the essential theoretical insight of matching individual capabilities to industry requirements while introducing explicit channels for educational investment and matching frictions, thereby extending the task-based framework to accommodate broader human capital formation processes.

[6]This specification reflects both the necessity of maintaining minimum competency across all dimen-

Decision making The agent i receives the expected utility when planning to enter industry j at time t:

$$\mathbb{E}\left(U_{ijt}\right) = \mathbb{E}\left[w_{ijt}(a_{ij}) - c_{ij}(a_{ij})\right]$$

Suppose $\mathbf{J_1} \bigcup \mathbf{J_2} = \mathbf{J}$ where $\mathbf{J_1}$ denotes the common market and $\mathbf{J_2}$ the three-high market, underlines the different wage structure:

$$w_{ijt} = \beta_j f\left(A_O, a_{ijt}\right) + \eta_j, j \in \mathbf{J_1}$$

where β_j is the industry-specific return to ability, η_j as the base wage. Let the wage in the three-high market as

$$\mathbf{E}[w_{ijt}] = p_{ijt} w_{jt}^{high} + (1 - p_{ijt})(b - c_0), j \in \mathbf{J_2}$$

where p_{ijt} refers to the probability of being admitted into the three-high market. Formally define the probability as

$$p_{ijt} = \frac{1}{1 + e^{-\lambda(F_{jt}(A_O(\mathbf{a_i})) - (1-\theta_j))}}$$

where $\lambda > 0$ is a smoothing parameter[7]. $F_{jt}(\cdot)$ is continuously differentiable and denotes the cumulative distribution function of aggregated abilities in market j at time t, and $\theta_j \in (0,1)$ represents the admission threshold for market j. Specifically, θ_j can be interpreted as the proportion of candidates that the market aims to admit (e.g., $\theta_j = 0.2$ indicates that the market targets the top 20% of candidates).[8]

sions (first term) while allowing for enhanced productivity through strategic skill combinations (second term). For instance, strong technical skills paired with management ability may generate additional value beyond their individual contributions, represented by a larger γ_{ij} for this combination.

[7]The smoothing parameter λ governs the steepness of the probability transition around the threshold $(1 - \theta_j)$, with larger values of λ ($\lambda \to \infty$) approximating a more discrete selection mechanism.

[8]The time subscript t in $F_t(\cdot)$ allows for potential temporal variation in the ability distribution, reflecting dynamic changes in the candidate pool's composition.

2.2 Logit Choice Model with Participation Conditions

We apply the logit model to formulate the expected utility index as a probabilistic choice[9], including an outside option (unemployment). The model estimates the probability that agent i will select industry j or the outside option, based on the utility value V_{ij} calculated for each alternative. The formulation Q_1 is as follows:

$$P_{ij} = \frac{\exp(V_{ij})}{\exp(V_{i0}) + \sum_{k\in\mathbf{J}} \exp(V_{ik})} \quad \text{for } j \in \{0\} \cup \mathbf{J} \tag{11}$$

$$V_{ij} = \begin{cases} b - c_0 & \text{if } j = 0 \text{ (outside option)} \\ \max\{\beta_j f(A_O, a_{ij}) + \eta_j - c_{ij}(a_{ij}), b - c_0\} & \text{if } j \in \mathbf{J_1} \\ \max\left\{p_{ijt}[w_{jt}^{high} - c_{ij}(a_{ij})] + (1 - p_{ijt})(b - c_0), b - c_0\right\} & \text{if } j \in \mathbf{J_2} \end{cases} \tag{12}$$

$$c_{ij}(a_{ij}) = \kappa_j \cdot \left(\frac{1}{A_O(a_{ij})}\right) + \delta_{ij}, \quad \forall j \in \mathbf{J} \tag{13}$$

$$p_{ijt} = \frac{1}{1 + e^{-\lambda(F_{jt}(A_O(\mathbf{a_i})) - (1-\theta_j))}} \tag{14}$$

With participation decisions determined by:

$$I_{ij} = \begin{cases} 1 & \text{if } \beta_j f(A_O, a_{ij}) + \eta_j - c_{ij}(a_{ij}) > b - c_0 \text{ and } j \in \mathbf{J_1} \\ 1 & \text{if } p_{ijt}[w_{jt}^{high} - c_{ij}(a_{ij})] + (1 - p_{ijt})(b - c_0) > b - c_0 \text{ and } j \in \mathbf{J_2} \\ 0 & \text{otherwise} \end{cases} \tag{15}$$

The outside option ($j = 0$) utility V_{i0} consists of:

- Base unemployment benefit/home production value (b)

[9] Another approach is to form the "Utility Maximization Choice Model". Write down the overall optimization problem for agent i as Q_i

$$j^* = \arg\max_{j\in\mathbf{J}} \mathbb{E}(U_{ijt}) = \begin{cases} \beta_j f(A_O, a_{ijt}) + \eta_j - c_{ij}(a_{ij}) & \text{if } j \in \mathbf{J_1} \\ p_{ijt} w_{jt}^{high} + (1 - p_{ijt})b - c_{ij}(a_{ij}) & \text{if } j \in \mathbf{J_2} \end{cases} \tag{4}$$

$$p_{ijt} = F(\text{rank}(A(a_{ij}))|N_j) \tag{5}$$

$$c_{ij}(a_{ij}) = \kappa_j \cdot \left(\frac{1}{A_O(a_{ij})}\right) + \delta_{ij} \tag{6}$$

$$\mathbf{J} = \mathbf{J_1} \cup \mathbf{J_2} \tag{7}$$

$$p_{ijt} \in [0, 1] \tag{8}$$

$$\kappa_j > 0, \delta_{ij} \geq 0 \tag{9}$$

$$w_{jt}^{high} > b > 0 \tag{10}$$

- Fixed cost c_0 of job searching

The actual labor supply for the common market is:

$$L_j^s = \sum_{i \in \mathbf{I}} P_{ij} \cdot I_{ij} \quad \text{for } j \in \mathbf{J_1} \tag{16}$$

For the three-high market ($j \in \mathbf{J_2}$), we distinguish between the effective labor supply (those who choose to participate) and the actual labor supply (those who are admitted):

$$L_j^{s,effective} = \sum_{i \in \mathbf{I}} P_{ij} \cdot I_{ij} \quad \text{for } j \in \mathbf{J_2} \quad \text{(total attempting)} \tag{17}$$

$$L_j^{s,actual} = \sum_{i \in \mathbf{I}} P_{ij} \cdot p_{ij} \cdot I_{ij} \quad \text{for } j \in \mathbf{J_2} \quad \text{(actually admitted)} \tag{18}$$

where $L_j^{s,effective} - L_j^{s,actual}$ represents the excess labor supply in the three-high market.

The total unemployment in the economy consists of voluntary and involuntary unemployment:

$$\begin{aligned} U &= U_{voluntary} + U_{involuntary} \\ &= \sum_{i \in \mathbf{I}} \left[P_{i0} + \sum_{j \in \mathbf{J_2}} P_{ij}(1 - p_{ij}) I_{ij} \right] \\ &= \underbrace{\sum_{i \in \mathbf{I}} P_{i0}}_{\text{voluntary unemployment}} + \underbrace{\sum_{i \in \mathbf{I}} \sum_{j \in \mathbf{J_2}} P_{ij}(1 - p_{ij}) I_{ij}}_{\text{involuntary unemployment}} \end{aligned} \tag{19}$$

Note that:

- P_{i0} captures those who voluntarily choose the outside option
- $P_{ij}(1 - p_{ij})I_{ij}$ for $j \in \mathbf{J_2}$ represents those who chose three-high markets but were not accepted
- The participation rate in the economy is: $1 - \frac{1}{|\mathbf{I}|} \sum_{i \in \mathbf{I}} P_{i0}$

The unemployment rate can be expressed as:

$$u = \frac{U}{|\mathbf{I}|} = \frac{1}{|\mathbf{I}|} \left[\sum_{i \in \mathbf{I}} P_{i0} + \sum_{i \in \mathbf{I}} \sum_{j \in \mathbf{J_2}} P_{ij}(1 - p_{ij}) I_{ij} \right] \tag{20}$$

2.3 Equilibrium Conditions

For any industry $j \in \mathbf{J_1}$, the choice probability matrix can be expressed as

$$P_{ij}(\eta_j, \eta_{-j})$$

where η_j is the market determined base wage in industry j and η_{-j} represents the vector of base wages in all other industries. The labor supply in industry j is the sum of individual choices $L_j^s(\eta_j, \eta_{-j}) = \sum_{i \in I} P_{ij}(\eta_j, \eta_{-j}) \cdot I_{ij}$ as discussed. The labor demand $L_j^d(\eta_j)$ is derived from firms' cost minimization problems in industry j.

Definition 1 (Market Equilibrium in Industry j). *Given other industries' base wages η_{-j}, a market equilibrium in industry $j \in J_1$ is characterized by a base wage η_j^* that satisfies:*

$$L_j^s(\eta_j^*, \eta_{-j}) = L_j^d(\eta_j^*)$$

Proposition 1 (Existence of Partial Equilibrium). *Under the following conditions:*

1. $L_j^s(\eta_j, \eta_{-j})$ *is continuous and strictly increasing in* η_j *for all* $j \in \mathbf{J_1}$
2. $L_j^d(\eta_j)$ *is continuous and strictly decreasing in* η_j *for all* $j \in \mathbf{J_1}$
3. $\lim_{\eta_j \to 0} L_j^s(\eta_j, \eta_{-j}) = 0$ *and* $\lim_{\eta_j \to \infty} L_j^d(\eta_j) = 0$ *for all* $j \in \mathbf{J_1}$

there exists a unique partial equilibrium $\{\eta_j^*\}_{j \in J_1}$ such that [10]*:*

$$L_j^s(\eta_j^*, \eta_{-j}^*) = L_j^d(\eta_j^*) \quad \forall j \in \mathbf{J_1}.$$

Definition 2 (General Equilibrium with Two-Tier Labor Markets). *Given wages $\{w_j^{high}\}_{j \in J_2}$ in the three-high sectors outside option c_0 and unemployment benefit b, a general equilibrium consists of:*

1. *Base wages $\{\eta_j^*\}_{j \in J_1}$ determined simultaneously in common markets*
2. *Choice probabilities $\{P_{ij}\}$ for all individuals i and sectors $j \in \{0\} \cup J_1 \cup J_2$*
3. *Employment indicators $\{I_{ij}\}$ for all i, j*

10

Proof. For any given vector of other industries' wages η_{-j}, conditions 1-3 ensure the existence of a unique equilibrium wage η_j^* in industry j by the intermediate value theorem. The existence of a complete partial equilibrium follows from applying Brouwer's fixed point theorem to the system of equations. □

such that:

(i) Labor Market Clearing in Common Markets:

$$\sum_{i \in I} P_{ij}(\{\eta_j^*\}_{j \in J_1}, \{w_j^{high}\}_{j \in J_2}) \cdot I_{ij} = L_j^d(\eta_j^*) \quad \forall j \in J_1$$

(ii) Optimal Individual Choices: For all i,

$$I_{ij} = \begin{cases} 1 & \text{if } \beta_j f(A_O, a_{ij}) + \eta_j^* - c_{ij}(a_{ij}) > b - c_0 \text{ and } j \in J_1 \\ 1 & \text{if } p_{ij}[w_j^{high} - c_{ij}(a_{ij})] + (1 - p_{ij})(b - c_0) > b - c_0 \text{ and } j \in J_2 \\ 0 & \text{otherwise} \end{cases}$$

(iii) Unemployment Determination:

$$\begin{aligned} U &= U_{voluntary} + U_{involuntary} \\ &= \sum_{i \in I} P_{i0} + \sum_{i \in I} \sum_{j \in J_2} P_{ij}(1 - p_{ij}) I_{ij} \end{aligned}$$

where:

- $U_{voluntary} = \sum_{i \in I} P_{i0}$ *represents individuals who choose the outside option*
- $U_{involuntary} = \sum_{i \in I} \sum_{j \in J_2} P_{ij}(1 - p_{ij}) I_{ij}$ *represents individuals who choose three-high sectors but are not accepted*

This equilibrium characterization highlights three key features:

1. Market-clearing wages in common markets coexist with rationing in three-high sectors
2. Both voluntary and involuntary unemployment emerge endogenously
3. The participation decision incorporates expectations about acceptance probabilities in three-high sectors