Evolution of the Labor Market in a Rapidly Developing Economy

A Preliminary Draft

James Liang[1]

April, 2011

Abstract

This paper builds a model for the evolution of the labor market in a rapidly growing economy in response to the arrival of a large number of foreign (high-productivity) firms. Such high-productivity firms increase demand for skill which requires both education and experience. While the return to education will rise, the wages of young college graduates may decrease. The dynamic version of the model also predicts that in the short run, there will be an over-supply of young college graduates and higher wage inequality than in the long run. Using a unique wage dataset, I found that these predictions are consistent with the recent development of the labor market in China. I also found supporting evidence from other rapidly developing countries in the last thirty years.

1 Introduction

Economists have found that human capital accumulation is a critical contributor to economic growth. Human capital is acquired not only in schools, but also on the job where firms provide training and valuable work experience. In addition, certain types of firms can raise the return to education and induce future investment in education. Foreign firms in developing countries play a particularly important role in this regard, because they usually have better technology, demand higher skills, and offer higher pay for education and experience than local firms. This paper studies the evolution of human capital in a rapidly developing country in response to the arrival of foreign firms which hire more highly skilled workers.

China is a good case study. China was completely closed for foreign direct investment (FDI) before 1978, and then in the 1980s, it adopted a policy to promote and attract FDI. As a result, FDI took off in China the

[1]I am very grateful to my advisor, Edward Lazear for his continuous guidance and support on the project. I thank Nick Bloom, Claire Lim, John Roberts, and Kathryn Shaw for their extremely helpful comments.

1990s. FDI accounted for 1% of GDP in 1990, and rapidly increased to 5% of GDP in 1995, remaining high in the 2000s. These foreign firms have brought profound changes to the local labor markets. In big cities like Shanghai and Beijing, they make up for more than 15% of skilled labor employment, and typically these foreign firms hire more skilled labor, pay higher wages and have a steeper pay scale than local firms (Zhao (2001)). In doing so, they have significantly raised the return to human capital, and thereby greatly influenced the development of the human capital in China. The return to schooling started to rise rapidly in the 1990s, from 5% in 1993 to 10% in 1999. From 1986 to 1995, college enrollment only increased from 700,000 to about 1 million, but from 1996 to 2005, college enrollment increased dramatically to more than 4 million a year.

In the early years of opening up, the wages of college graduates (both experienced and inexperienced) increased significantly (Zhang et al. (2005)). However, in recent years, the wage premium of young college graduates has been declining, even when the overall return to education continues to rise steadily. All the increase in return to education comes from the increase in the earnings of experienced graduates. Using a unique dataset from a leading online recruiting agency, I found that the college premium for the young and the old started to diverge in late 2000. The college wage premium for older graduates (age 28 or older) increased from 40% in early 2000 to 50% in late 2000, while the college premium for younger graduates (age 27 or younger) decreased from 20% to less than to 10%. The most recent graduate cohort (born after 7/1/1982) started out their careers at a wage 15% lower than the earlier cohorts, but were able to catch up in average wages after about 5 years after graduation.

I present a theoretical framework to analyze the evolution of human capital development as a result of an increase in demand for skills which require both education and experience. In this model, I will derive the impact of skill-intensive foreign firms on the evolution of wages, return to education, return to experience, and skill mix in the economy. The model predicts that with an influx of foreign firms, the overall return to education and the return to experience will increase, and as a result, will induce a large increase in investment in college education. The wage of inexperienced graduates will go up initially, but over time, it may decrease to a level lower than before the influx of FDI. I also present a dynamic model and a numerical solution which shows the evolution of the labor market over time. Initially, all wages will go up. But in the medium run, there will be an oversupply of inexperienced graduates, and the wage for inexperienced graduates will decrease while the wage for experienced graduates will continue to rise. These predictions are found to be consistent with the evidence in China and in several other rapidly developing countries.

The model also analyzes FDI's impact on local firms. Some of the local firms that are directly competing with foreign firms for high-skill workers will be worse off, but other domestic firms may benefit from a larger pool of moderately skilled workers. Moreover, the model predicts that a more developed domestic industry can help train more high-skill workers, thus benefiting the foreign firms and the whole economy.

I will set up the model in section 2, and solve and discuss the steady state equilibrium in section 3. In section 4, I will characterize the dynamic equilibrium, and present a numerical solution. In section 5, I will use a unique wage dataset to analyze the recent wage pattern in China and compare it to the simulation result. In section 6, I will present supporting evidence from other rapidly developing countries. In section 7, I will use the model to run some policy experiments numerically to analyze the policy of subsidizing college education and the policy of limiting the growth of college education.

1.1 Related Literature

On the theory side, as far as I know, this is the first attempt to analyze the labor market of a rapidly developing economy in a dynamic setup. Heckman (1998) used a dynamic general equilibrium model to explain the changing wage structure in the recent U.S. labor market. Katz and Murphy (1992) provided a framework to analyze wages and inequality in response to exogenous supply shifts and technology changes. The novelty of this model lies in endogenizing the supply of human capital under the assumption that productive human capital requires both education and work experience. This model also introduces firm heterogeneity which allows me to analyze the effects on different types of firms.

There are many empirical research papers examining the changes in wage structures in the U.S. in the last 30 years. Acemoglu (2002) analyzed the effect of skill-biased technology on the recent labor market. Freeman and Katz (1995) compared the wage structure in the U.S. to that of other OECD countries in the last few decades. On the topic in developing countries, many research papers documented the development of labor markets in emerging countries such as China (Zhang et al. (2005)), Korea (Kwark and Rhee

(1993)), and Taiwan (Baraka (1999)).

In addition, there is a lot of empirical research investigating the effect of FDI on the labor market of the host country. Freestra and Hanson (1997) found FDI can contribute a great deal to the increase of return to skill in the local market in Mexico. Zhao (2001) found that foreign companies pay almost twice the college premium as local companies in China.

2 Model Setup

Let representative Consumer's utility $U = \int_{j\in\Omega} (q_j)^\rho dj + Q$, where $0 < \rho < 1$. The first term is the utility for differentiated goods, and the second term is the utility for numeraire.

There are three types of workers: the low-ability workers denoted by l, medium-ability workers denoted by m (college-educated but inexperienced), and high-ability workers denoted by h (experienced and educated). l ability workers can get a college education to become m, but m may become h only after working for some number of periods. Each period, a portion η of the m workers turn into h workers. Also, in each period, there are $1 - \mu$ workers entering and exiting the labor force. The total labor force is normalized to be 1.

The production function for the differentiated goods is given by $q^j = z_l^j + \delta z_m^j + \delta^2 * z_h^j$, which says the production for variety j, is a linear function of the amount of labor hired for l, m, and h-type workers. h-type workers are δ times more productive than m-type workers, who are δ times more productive than l-type workers. δ is the productivity for the firm. There are a mass of n domestic firms whose δ follows a distribution with c.d.f $F(\delta)$ over the range $(1, \Delta)$. (Assume $0 < f() < \infty$). For simplicity, let's assume all firms have zero fixed cost, and the owners of the firms spend their profits on numeraires only. The number of foreign firms n_f is exogenous. The productivity of the foreign firms is at the upper bound of the productivity distribution Δ.

There is also a numeraire sector that only hires l workers. The production function is given by a constant scale function $Q = z_l'$. I normalize the price for the numeraire to be 1, and also the wage for l to be 1 (i.e. $w_l = 1$). This sector can be interpreted as the low-skill sector, such as agriculture, which has low productivity and provides no on-the-job training.

To model a worker's decision to get a college education, I assume different workers have different costs to obtain a college education, and a worker will obtain an education if the reward is greater than the cost. The supply function $G(v)$ denotes the number of workers who will get a college education as a function of v. v is the expected wage premium net the common cost of education (denoted by C) which includes tuition, etc. The supply curve can be generated by the assumption that different workers have different effort costs on the top of common cost to obtain an education (the effort cost can be interpreted as innate ability). $G(v)$ can be interpreted as the fraction of people whose effort cost of education is less than the net reward v. (Assume $G'() > 0$)

For simplicity, I assume no time discounting and risk neutrality.

With the above setup, I will derive the equilibrium condition. Then I will analyze the equilibrium and dynamics when there is an increase in the number of foreign firms n_f.

3 Steady State Equilibrium

From the utility function I can derive the demand for goods q^j given price p^j. For simplicity I will drop the index j.

The demand function is $p(q) = Aq^{-\frac{1}{\sigma}}$ where $\sigma = 1/(1-\rho)$ and A is a constant. This is a constant elasticity demand function. By the usual analysis for constant elasticity demand function, I can derive the production quantity, the labor demand and the profit for each firm. If a firm decides to hire workers of type l, with wage 1, the demand function for the workers type l is $z_l = B$, where constant $B \equiv (1-\frac{1}{\sigma})^\sigma A^\sigma$. If a firm decides to hire workers m, the demand function is $z_m = Bw_m^{-\sigma}\delta^{\sigma-1}$. If a firm decides to hire workers h, the demand function is $z_h = Bw_h^{-\sigma}(\delta^2)^{\sigma-1}$.

Hereafter, I define $w \equiv \frac{w_h}{w_m}$, which can be interpreted as the return to experience. The equilibrium is characterized by wage schedule w_m and w_h (or equivalently w_m and w). I am interested in analyzing solutions where $w > w_m > 1$. In this case, there is a positive matching between workers and firms, i.e. higher-productivity firms hire higher-productivity workers, and the differentiated good sector hires all three types of workers. Based on the production function $q = z_l + \delta z_m + \delta^2 z_h$, the hiring decision is that firms with $\delta \in (1, w_m)$ will hire l; firms with $\delta \in (w_m, w)$ will hire m; and firms with $\delta \in (w, \Delta)$ will hire h. All foreign firms have productivity Δ by assumption, so they will hire h (assuming there are enough h in the economy to be hired by foreign firms).

In the steady state, the ratio of L_h (total number of h workers) to L_m (the total number of m workers) is a constant $k \equiv \frac{\mu\eta}{1-\mu}$, because

$$L_h^t = \mu L_h^{t-1} + \mu\eta L_m^{t-1} \tag{1}$$

The steady state requires $L_h^t = L_h^{t-1} = L_h$, solving the above equation yields

$$L_h = \frac{\mu\eta}{1-\mu} L_m \tag{2}$$

Further, the total supply of m and h is given by $G(v)$

$$L_h + L_m = G(v) \tag{3}$$

So, the stock of m and the stock of h are both a constant proportion of $G(v)$, where v is expected net reward of getting a college education (amortized per period). v is given as the following expression:

$$v = \frac{1}{1+k} w_m + \frac{k}{1+k} w_h - 1 - C \tag{4}$$

where C is the common cost of education amortized per period.

The labor market clearing condition for L_m is given by the following equation. The left-hand side is the demand for m, which is generated by domestic firms whose δ is between w_m and w. The right-hand side is the supply of m, which is a portion $1/(k+1)$ of the total educated workforce $G(v)$.

$$nBw_m^{-\sigma} \int_{w_m}^{w} \delta^{\sigma-1} dF(\delta) = \frac{1}{k+1} G(v) \tag{5}$$

The labor market clearing condition for L_h is given by the following equation. The left-hand side is the demand for h, which is generated by domestic firms whose δ is between w and Δ, and n_f foreign firms. The right-hand side is the supply of h, which is a portion $k/(k+1)$ of the total educated workforce $G(v)$.

$$Bw_h^{-\sigma}[n \int_{w}^{\Delta} (\delta^2)^{\sigma-1} dF(\delta) + n_f (\Delta^2)^{\sigma-1}] = \frac{k}{k+1} G(v) \tag{6}$$

The above two equations jointly solve two unknowns w_m and w_h. v is given by equation (4), and $w \equiv \frac{w_h}{w_m}$.

Table 1: A glossary of notation

Symbols	Description	Value Used in Simulation
A, B	constant	
σ	elasticity of demand	1.3
l, m, h	low, medium, high ability workers	
w_m, w_h	wage for medium and high ability workers	
w	$\equiv w_h/w_m$	
L_m, L_h	stock of m and h workers	
$F()$	distribution of firm productivity	Uniform Distribution on [1,2]
δ	productivity of the firm	
Δ	productivity of the foreign firms	2
n	number of local firms	10
n_f	number of foreign firms	10
η	the probability of m becoming h in each period	0.05
μ	the probability of staying in the workforce in each period	0.95
k	$\mu\eta/(1-\mu)$	
C	the common cost of obtaining education	0.2
$G()$	the distribution of idiosyncratic cost of obtaining education	Normal Distribution with mean 0.4 and variance 0.1
v	the expected return to education before idiosyncrative cost	

The labor employed by the unskilled sector (producing numeraire Q) is calculated as the remaining workers in the economy using the following labor clearing condition.

$$Q = 1 - G(v) - C * G(v) - nBF(w_m) \tag{7}$$

Q is calculated as the total labor force minus the labors employed by the skilled sector, the common cost of education (the second last term), and the labor demand for l from the differentiated goods sector (the last term).

Proposition 1 The market clearing equations for m and h have a unique solution: w_m, w_h

Proof in Appendix.

I will analyze the situation where the solutions of the above two equations, satisfy $w > w_m > 1$, which leads to positive sorting such that higher productivity firms will hire higher ability workers and all three types of workers are employed by the differentiated goods sector. This is the most interesting case, otherwise if $w_m < 1$, then no l will be hired by the differentiated goods sector.

3.1 Impact on Labor Market with More Foreign Firms

I want to see how the equilibrium changes if there are more foreign firms, i.e. n_f increases. This could be due to a regulation change to allow FDI in certain sectors previously closed to foreign firms.

Proposition 2 When n_f increases, w,v,and w_h increase.

Proof in Appendix.

With more high-productivity foreign firms, return to experience (w), return to education (v), and the wage for h (w_h) will increase. But the effect on w_m is ambiguous.

Proposition 3. When n_f increases, if $G'()$ is sufficiently small then w_m increases, and if $G'()$ is sufficiently large then w_m decreases.

Proof in Appendix.

Intuitively, we already know that more foreign firms will raise the return to education and the return to experience. When the return to experience w increases, some firms cannot afford h any more and will have to hire m, thus increasing the demand for m. On the other hand, when return to education increases, the supply of m increases. When the supply curve is very elastic, the supply effect dominates and w_m decreases. This is saying, if the education

Table 2: Impact of foreign firms on existing firms when $w_m^1 < w_m^0$

Firms with $\delta \in$	Workers	Profit Before	Profit After	Comparison
$(1, w_m^1)$	l	$\frac{1}{\sigma-1}B$	same as before	same
(w_m^1, w_m^0)	$l- > m$	$\frac{1}{\sigma-1}B$	$\frac{1}{\sigma-1}B(\frac{\delta}{w_m^0})^{\sigma-1}$	better off
(w_m^0, w^0)	m	$\frac{1}{\sigma-1}B(\frac{\delta}{w_m^0})^{\sigma-1}$	$\frac{1}{\sigma-1}B(\frac{\delta}{w_m^1})^{\sigma-1}$	better off
$(w^0, w^0 w_m^0/w_m^1)$	$h- > m$	$\frac{1}{\sigma-1}B(\frac{\delta^2}{w_h^0})^{\sigma-1}$	$\frac{1}{\sigma-1}B(\frac{\delta}{w_m^1})^{\sigma-1}$	better off
$(w^0 w_m^0/w_m^1, w^1)$	$h- > m$	$\frac{1}{\sigma-1}B(\frac{\delta^2}{w_h^0})^{\sigma-1}$	$\frac{1}{\sigma-1}B(\frac{\delta}{w_m^1})^{\sigma-1}$	worse off
(w^1, Δ)	h	$\frac{1}{\sigma-1}B(\frac{\delta^2}{w_h^0})^{\sigma-1}$	$\frac{1}{\sigma-1}B(\frac{\delta^2}{w_h^1})^{\sigma-1}$	worse off

system is responsive to a higher return to education, more foreign firms in the skilled sector can reduce w_m. In the extreme case, when the supply of m is perfectly elastic at some point v_0, then when w increases, w_m will have to decrease to keep $v = v_0$. On the other hand, if $G(v)$ is perfectly inelastic, i.e. the supply of m, and h is fixed, then as w increases, the demand for m increases while the supply is unchanged, therefore bidding up w_m.

3.2 Impact on Existing Firms

In this section, I will consider how the entry of foreign firms impacts the profitability of local firms. We already know that with more foreign firms, w_h increases. The following proposition states, that when $G()$ is elastic and w_m decreases, some of the local firms can actually benefit. Table 2 shows different firms' hiring decisions and profits before and after the influx of foreign firms. In the table, w_m^0 and w_h^0 are the equilibrium wages for m and h before the influx of foreign firms, and w_m^1, w_h^1 are the wages for m and h in the new equilibrium after the influx of foreign firms.

Proposition 4 When n_f increases, if w_m decreases in the new equilibrium, then there exists a cut off $\overline{\delta} = w^0 w_m^0/w_m^1$, such that firms with $\delta < \overline{\delta}$ are better off, and firms with $\delta > \overline{\delta}$ are worse off.

Proof: The firms that are still hiring h are worse off, because w_h increases. The firms that continue to hire m are better off since w_m is reduced. The firms that switch from hiring l to hiring m are better off, because previously they make $\frac{1}{\sigma-1}B$, and now they make $\frac{1}{\sigma-1}B(\frac{\delta}{w_m^1})^{\sigma-1}$. The profit is greater if and only if $\delta > w_m^1$, which is true for them to hire m now, therefore they are better off. Now consider the firms with $\delta \in (w^0, w^1)$. They switch from hiring h to hiring m. Previously they made profit $\frac{1}{\sigma-1}B(\frac{\delta^2}{w_h^0})^{\sigma-1}$. Now they make $\frac{1}{\sigma-1}B(\frac{\delta}{w_m^1})^{\sigma-1}$. The firm is worse off if $\delta > w_h^0/w_m^1 = w^0 * w_m^0/w_m^1$, or better off if $\delta < w^0 w_m^0/w_m^1$. Q.E.D.

I just showed that when more high-productivity foreign firms come in, if the supply function of educated workers $G(v)$ is elastic enough such that w_m is lowered, some of the local firms can actually benefit. Intuitively, those firms that are not competing with foreign firms for high-ability workers can benefit from more medium-ability workers. In a sense, the high-productivity foreign firms are complements to the medium-productivity firms.

The converse is also true, i.e. when there are more medium-productivity firms and the supply of education is sufficiently elastic, the foreign firms will actually benefit. This is given by proposition 4d in the Appendix, which says, more local firms can reduce w_h, thus benefit the foreign firms. Studies on foreign direct investment have documented that a well developed local sector makes it a more attractive environment for foreign companies to invest in (Coughlin and Segev (2002)). This model provides a mechanism in which more local firms will induce more education and train high ability workers, and therefore benefit the foreign firms.

4 Dynamics and Short-Term Effects

Here I will examine the fully dynamic model. I will analyze the evolution of the labor market when the number of foreign firms increases unexpectedly. This could be due to a new policy to encourage FDI in sectors that were previously closed to FDI, as occurred in China during the 90s. Initially at $t = 0$, no foreign firm is expected to enter, i.e. $n_f \equiv 0$, and L_h^0, L_m^0, w^0, w_m^0 is a steady state equilibrium given the information at time $t = 0$. Now, at $t = 1$ with the unexpected new policy, it becomes public knowledge that n_f^t for $t >= 1$ will become positive. I will assume that in every period, only young potential workers can make the decision to get a college education, and it takes four years to graduate. As before, I assume, in each period, η portion of the m workers turn into h. I will derive the dynamic equilibrium, i.e. the wages path, w_m^t and w_h^t. The equilibrium is defined so that each person will make an optimal decision to go to college or not.

To simplify the analysis, let's assume the number of foreign firms in each period is exogenous and is given by n_f^t. Further assume that n_f^t stabilizes over time, i.e. becomes a constant after some number of periods. Formally there exist some $\tau > 0$ such that $n_f^t = n_f^{t+1}$ for all $t > \tau$.

The only decision to model is the choice to go to college. By assumption, in each period, $(1-\mu)$ workers leave the workforce, and $(1-\mu)$ young people either join the workforce or go to college. The number of people who go to college in each period is $(1-\mu)G(v^t)$. v^t is the expected future payoff of going to college at time t amortized per period. It is given by the following expression which is a linear combination of future wages weighted by the probability of earning that wage net the unskilled wage and tuition.

$$v^t = [w_m^{t+4} + \sum_{j=1}^{\infty} \mu^j (1-\eta)^j w_m^{t+j+4} + \sum_{j=1}^{\infty} \mu^j (1-(1-\eta)^j) w_h^{t+j+4}](1-\mu) - 1 - C \quad (8)$$

The equilibrium $v^t.w_m^t, w_h^t$ can be fully characterized by the following equations. The first two equations model the evolution of stock of h and m workers, and the last two equations are the per period labor market clearing conditions.

$$L_h^t = \mu L_h^{t-1} + \mu\eta L_m^{t-1} \quad (9)$$

$$L_m^t = \mu L_m^{t-1} + (1-\mu) * G(v^{t-4}) - \mu\eta L_m^{t-1} \quad (10)$$

$$nB(w_m^t)^{-\sigma}[\int_{w_m^t}^{w^t} \delta^{\sigma-1} dF(\delta)] = L_m^t \quad (11)$$

$$B(w_h^t)^{-\sigma}[n\int_{w^t}^{\Delta} (\delta^2)^{\sigma-1} dF(\delta) + n_f^t (\Delta^2)^{\sigma-1}] = L_h^t. \quad (12)$$

These five equations above fully characterized the equilibrium $\{w_m^t, w_h^t, L_h^t, L_m^t, v^t\}$. Notationally, the new steady state is denoted as $w_m^s, w_h^s, L_h^s, L_m^s, v^s$.

Proposition 5 With this setup, an equilibrium exists.

Proof in Appendix.

The proof uses Schauder's fixed point theorem, by establishing the following mapping: Given a sequence of enrollment rates $\{g^t\}$, I can compute the evolution of L_m^t and L_h^t, and then using the period-by-period demand and supply equation, I can compute w_m^t and w_h^t from which I can compute v^t, finally I can get a new set of enrollment rates $g'^t = G(v^t)$. Thus I have a mapping from $\{g^t\}$ to $\{g^t\}'$. The equilibrium is just a fixed point such that $\{g^t\} = \{g^t\}'$.

From Proposition 1, I know any convergent equilibrium will converge to the unique steady state equilibrium derived in the previous section. Notationally, as t goes to ∞, $\{w_m^t, w_h^t, L_h^t, L_m^t, v^t\}$ converges to $\{w_m^s, w_h^s, L_h^s, L_m^s, v^s\}$ which satisfies equations 5 and 6.

Proposition 6 On the equilibrium path, if $G'() > 0$, there exists some period τ, such that $L_m^\tau > L_m^s$.

Proof in Appendix.

This proposition formally states one of the main results of the transition dynamics, i.e., the stock of m workers will overshoot the long-term steady state level. In other words, there will be an over-supply of m at some point. The intuition is that, h will be in short supply for a while because it takes time to train h, therefore, w_h and expected return to education is higher in the short run than in the long run. To take advantage of this, more people will get an education and more m will pour into the market, creating a seeming oversupply of m.

4.1 A Numerical Solution

To solve the 5 equations numerically, I repeatedly apply a mapping similar to the one used in the proof of existence to find the fixed point. I solve the equations numerically under a set of parameters that resemble the Chinese economy. The parameter n_f^t, the number of foreign firms is assumed to be initially growing very fast but stabilizes over time. This assumption resembles the growth of FDI in China since the 90s. The other parameters used are described in Table 1.

Figures 1,2,3,4 from the numerical solution depict the evolution of the labor market after the arrival of the foreign firms. Figure 1 shows the evolution of the stock of m and h. It shows that the growth path of L_h is flatter than that of L_m in the earlier periods since it takes longer to train h. As predicted by Proposition 6, the stock of m will overshoot and remain significantly higher than the steady state level for many periods.

Figure 2 shows the evolution of wage schedules. The wage for young graduates w_m will go up in the short run as more foreign firms drive up the demand for m, but will quickly reverse the trend and gradually decline to the steady state value which could be lower than the value before the shock. Based on the analysis of the impact of the foreign firms on the domestic firms, we know some of the domestic firms will eventually benefit from the entry of the foreign firms as w_m decreases eventually, but in the short run,

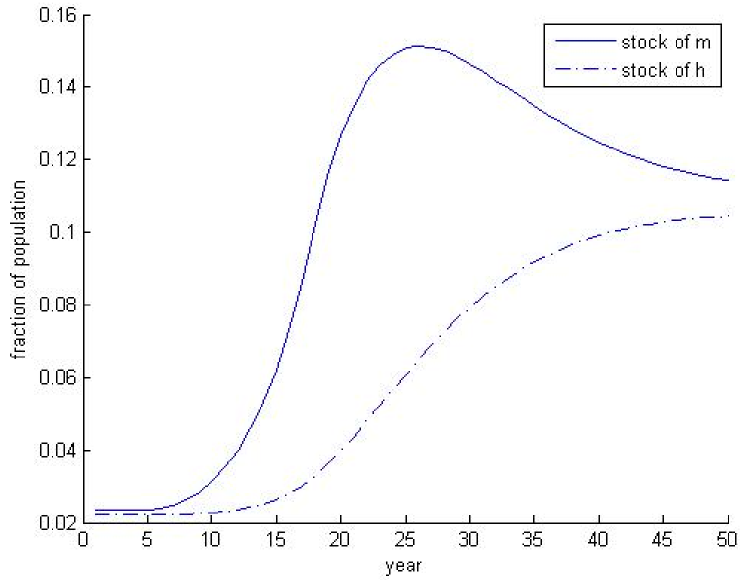

Figure 1: Evolution of Stock of h and m

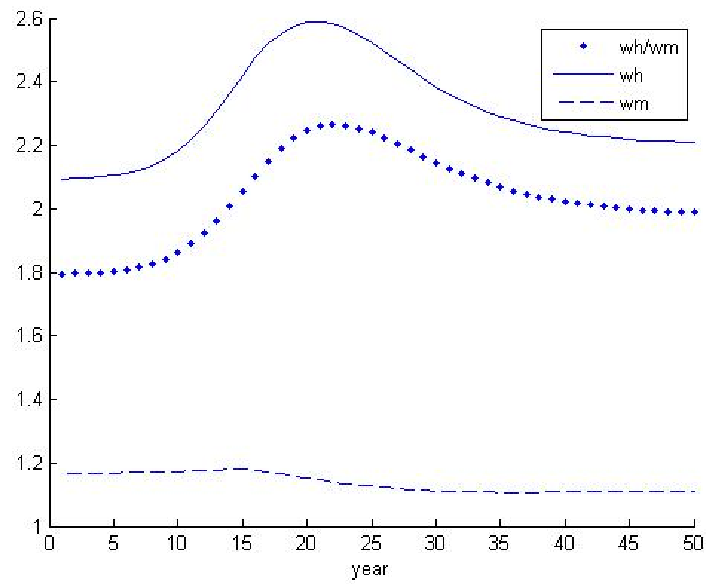

Figure 2: Evolution of Wages

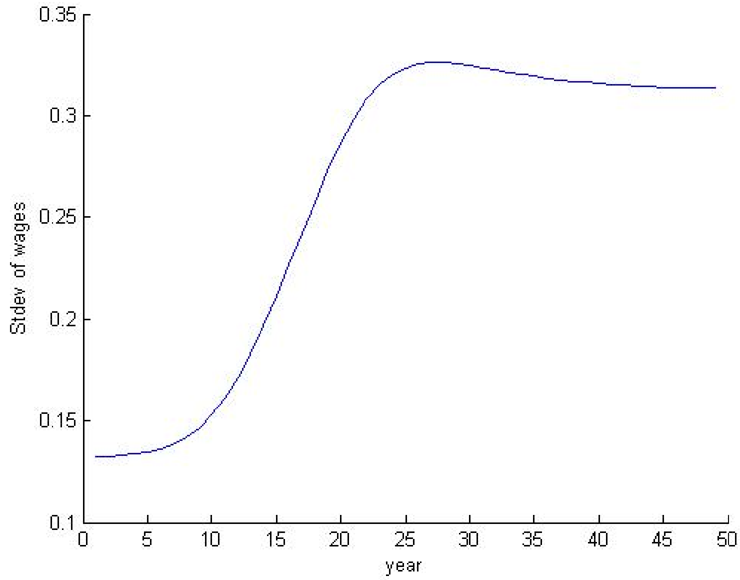

Figure 3: Evolution of Inequality (measured by std dev of wage)

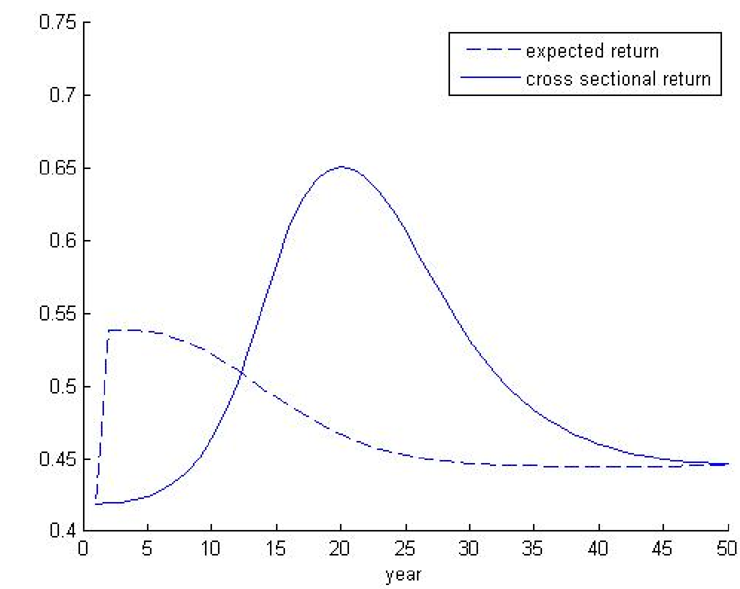

Figure 4: Evolution of Return to Education

as w_m increases, they will be feeling short-term pain.

Return to experience w goes up steeply initially, and continues to go up for a number of periods, before it peaks and comes down and gradually converges to the steady-state value which is higher than w^0. The reason that it takes longer for w to peak is that w depends on the relative ratio of L_m to L_h. In response to a higher return to education, many m pour into the work force in the first few years, and it will take them a while to become h. Therefore, the ratio of L_m to L_h, and hence w, will rise long after w_m starts to come down. Similarly, w_h will rise for a much longer period than w_m.

Figure 3 shows the evolution of cross-sectional inequality measured by the standard deviation of wages. Cross-sectional inequality increases rapidly after the shock, and reaches the peak in the medium run, and then comes down slowly in the long run. There are two reasons why inequality eventually comes down. First, return to experience (i.e. w) comes down in the long run, and second, over time more people will be educated and work in the skilled sector.

Figure 4 shows the evolution of return to education. The expected return to education (i.e. v_t) shoots up initially, but gradually declines and converges to the steady state value which is higher than v^0. The reason is that the earlier-educated cohorts can take advantage of the temporarily high wages, and thus have a higher return than the later cohorts. Consequently, the college enrollment rate also goes up sharply initially, overshooting the steady state value before convergence. The cross-sectional return to education, which is a weighted average of the w_m and w_h, takes quite a different path, and rises for a much longer period than the expected return to education.

5 Matching with Recent Wage Patterns in China

5.1 Dataset

In view of the lack of a high-quality public wage dataset, I obtained a unique wage dataset from a leading online recruiting firm in China. The dataset contains over 100,000 resumes posted to the company's website in the month of July, 2010. Each resume contains basic demographic information including age, sex, and education. For the education category, applicants select one out

Table 3: Summary Statistics

Variables	Remarks
age	Mean 28.6 , Std Dev 4.9
sex	63% male
wage	Mean 3589, Std Dev 2624 (monthly wage in RMB)
education	41% four-year college, (among them, 5% from elite colleges) 40% three-year college 17 % high school or equivalent
N	376809

of five education levels: advanced degree, 4-year college degree, 3-year college degree, high school graduate, or below high school, and applicants report the name of the education institution, majors and graduation dates.[2] In addition to basic demographic information, each resume contains information about current and previous jobs held by the applicant. For each job, applicants report employment period, company name, industry, and salary range which is selected from nine salary ranges.[3] All the following analysis will use the mid-point of the range as the reported salary. About 15% of the jobs did not include salary range information, and have been deleted from the sample. Altogether, the dataset inclues over 300,000 job observations.

Table 3 presents selected summary statistics of the dataset.

Figure 5 shows the age-earnings profiles for the four cohorts: late graduates (born after 7/1/1982), early graduates, late non-graduates and early non-graduates. For the two non-graduates cohorts, age-earnings profiles are almost the same. On the other hand, the age-earnings profile for late graduates is much steeper than that for early graduates. The late graduates cohort started their careers at significantly lower wages than the early cohort but were able to catch up after 6 years.

[2]In China, job applicants generally report education levels accurately, because many employers require copies of diplomas upon hiring

[3]Applicants may inflate current wages somewhat, but as long as the degree of inflation is roughly the same across demographic categories and level of education, the estimate of college premiums and time trend patterns will not be significantly affected.

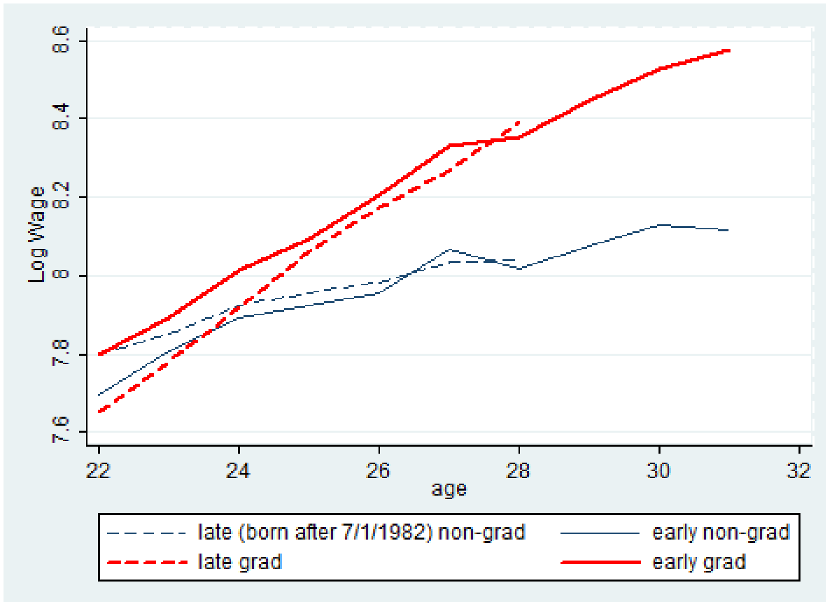

Figure 5: Cohort Wage Analysis

5.2 Wage Regression

The wage regression is a Mincerial-type regression with an interaction term between age and the college dummy variable. The following OLS regression is performed on all the observations and on different subsets of observations.

$$Log(Wage_i) = a + b * College_i + c * Exp_i + d * Exp_i^2 + e * Exp_i * College_i + u_i \quad (13)$$

$College_i$ is a dummy variable which is set to one for college graduates (both 3-year and 4-year college graduates). Exp_i is the actual age minus 22. So for a typical college graduate, this is the actual experience, but for a typical high school graduate, the actual experience is $4 + Exp_i$. The reason that I use age instead of actual experience in the regression is that I am more interested in comparing wages between graduates and non-graduates at the same age. One can easily back out the wages by actual experience and compare graduates and non-graduates wages at the same actual experience level.

The following table shows the result of the above wage regression using all observations and different subsets of observations. The first column shows the result of the regression using all observations. As shown, the coefficient on

Table 4: Wage Regressions

	Dep Var: Log Wage				
	All Obs	Early Cohort	Late Cohort	2000-2005	2006-2010
College	-0.18*	0.018*	-0.25*	0.032*	-0.21*
	(0.046)	(0.013)	(0.0059)	(0.016)	(0.0048)
Exp	0.096*	0.058*	0.090*	0.067*	0.098*
i.e. age-22	(0.0013)	(0.0035)	(0.0028)	(0.0049)	(0.0014)
Exp*Exp	-0.0057*	-0.0011*	-0.0031*	-0.0030*	-0.0059*
	(0.0001)	(0.00024)	(0.00037)	(0.00039)	(0.00011)
College*Exp	0.067*	0.042*	0.090*	0.037*	0.071*
	(0.00084)	(0.0017)	(0.0020)	(0.0030)	(0.00087)
constant	7.6	7.6	7.6	7.6	7.6
R squared	0.22	0.14	0.22	0.11	0.24
N	376809	128316	248493	35692	341117

Note: * Signficant at 5% level
Early Cohort (late cohort) are those born before (after) 7/1/1982.

the college dummy, i.e. the college premium at age 22 (or Exp_i=0) is actually negative. The coefficient on the interaction term is quite large, showing a strong complementarity between experience and education. The second and third columns compare the regressions done separately for the early cohort (born before 7/1/1982) and for the late cohort. Again as predicted, the college premium at age 22 changed from positive for the early cohort to negative for the late cohort, but the coefficient on the interaction term is much larger for the late cohort, showing an even stronger complementarity between experience and education. The fourth and fifth columns compare the results of regressions done separately for the cross-sectional job observations before and after 2005. As shown, the college premium at age 22 changed from positive to negative. Also as predicted, the coefficient on the interaction term is much larger after 2006.

5.3 Age-Earnings Profile of Different Types of Colleges

Figure 6 shows the age-earnings profiles for graduates from three types of colleges, i.e. the 3-year colleges, the 4-year colleges, and the elite colleges [4]. Consistent with the assumption that ability and experience are complements, the age-earnings profile of the 4-year college graduates is steeper than that of the 3-year college graduates, and the age-earnings profile of the elite college graduates is the steepest.

Figure 7 compares age-earnings profiles of the early cohorts (born before 7/1/1982) to those of the late cohorts across the three types of college graduates. All three age-earnings profiles shifted steeper for the later cohorts (born after 7/1/1982). The profile of the 4-year college graduates shifted more than that of the 3-year college graduates, and the wage of the young 4-year college graduates came down more than that of the young 3-year college graduates. This is consistent with the model prediction that when the wage of experienced 4-year graduates increases relative to that of experienced 3-year college graduates, the wage of young 4-year college graduates will decline relative to that of young 3-year college graduates if the supply of 4-year graduates is elastic. Also as shown, the wages of young elite graduates did not come down much, because the supply of elite college graduates is inelastic.

5.4 Robustness

First, I checked to see if the shift of earnings profile is an industry effect. In the data, each job is associated with one of the 44 industries. I broke out the dataset into sub-samples by industry, and computed the age-earnings profile of each industry. It is evident that the recent steepening of age-earnings profile is across all industries including service, manufacturing, and high-tech industries. I also broke out the sample by gender, and found the shifts of age-earnings profiles for males and females are very similar.

5.5 Comparing to the Simulation Results

Figure 8 shows the evolution of college premium for the young (age 27 or younger) and old graduates. The college wage premium for the young graduates actually declined since 2005, while the college wage premium for the old

[4] the top 30 colleges ranked by graduates earnings

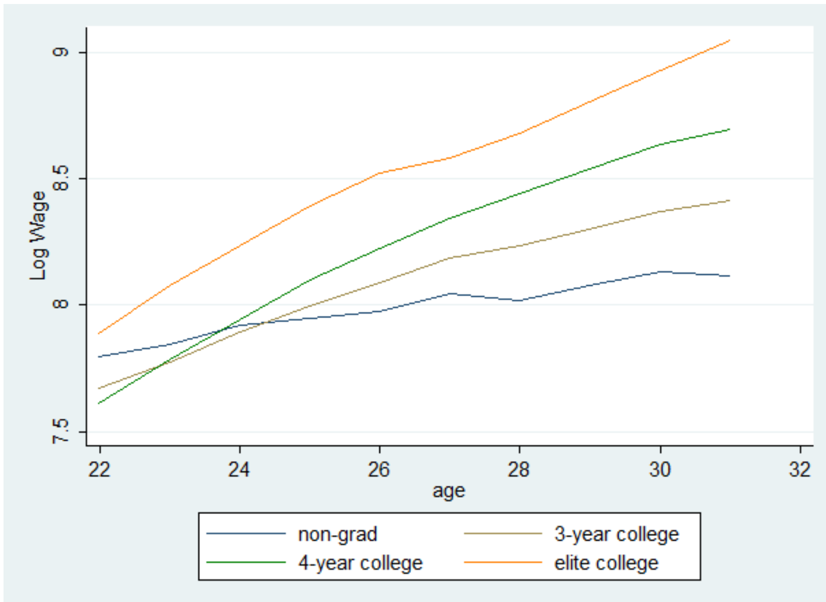

Figure 6: Age-Earnings Profile for Three Types of Colleges

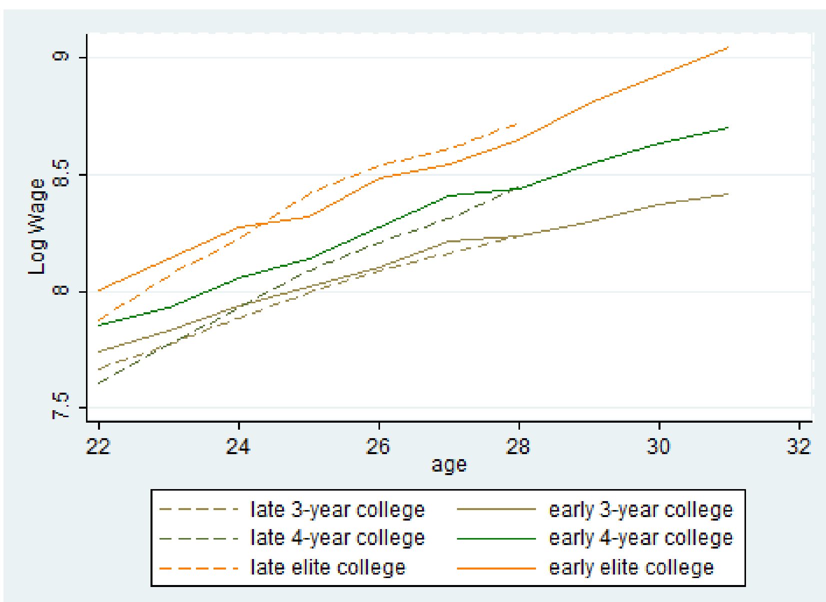

Figure 7: Cohort Comparison of Age-Earnings Profile for Different Types of Colleges

Figure 8: College Premium from Data

graduates continued to go up. This pattern matches the simulation result well as shown in figure 9.

5.6 Potential Selection Bias

The sample consists of young and middle-aged urban workers who are by no means representative of the whole Chinese population. First, unskilled migrant workers, usually with an education even lower than high school, are not in the sample. Second, the people in the sample are more educated than the general urban population; almost everybody (98%) is a high school or a college graduate. Further more, high school graduates are under-represented in the sample, which could be because high school graduates are less inclined to use the Internet for job searches. If only the high-ability high school graduates are using the Internet for job search, the high school graduates in the sample are likely to be positively selected in terms of ability. Therefore, wages for non-gradautes in the sample are probably higher than those of the general population, and the college premium estimated here is a lower bound for the college wage premium for the general population. In this light, the empirical exercise this dataset supports is actually that of comparing the

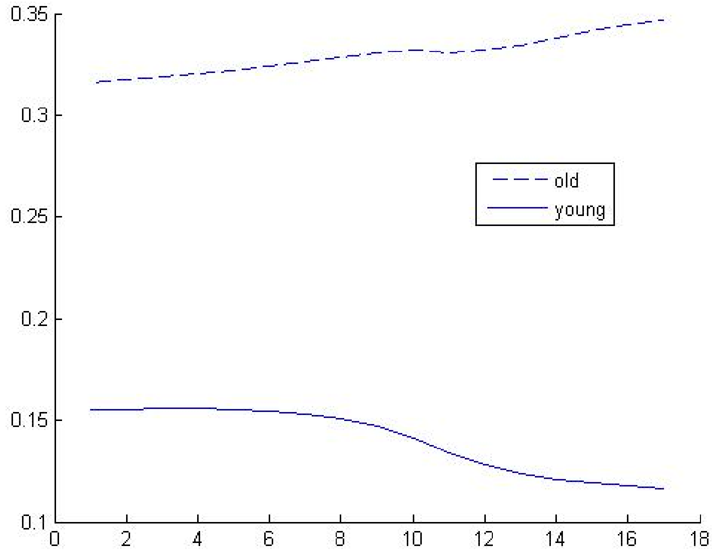

Figure 9: College Premium from Model

wages of the young graduates to the wages of relatively high-ability young urban non-graduates. Even though the college premium is estimated from a subset of the population, college premium's time pattern, particularly the divergence in college premium between the young and the experienced remain indicative of the overall college premium trend.

In further consideration of the issue of sample selection, I double-checked the college premium findings against data from the Urban Household Income and Expenditure Survey (UHIES). The UHIES survey is conducted by China National Statistics Bureau for the purpose of monitoring income and expenditure changes. The survey uses a diary record to collect individual earnings, other forms of income, household income and expenditure related data. It only includes those households whose household registrations (Hukou) are located in urban areas, so it does not include college graduates with rural backgrounds who work in cities after graduation, and it does not include migrant workers. There was a change in the survey method in 2002, so comparing data before 2002 to data after 2002 is not possible; currently only data before 2006 is available. Given that only 5 years of the data are available, I can not do synthetic cohort analysis, and am limited to cross-sectional

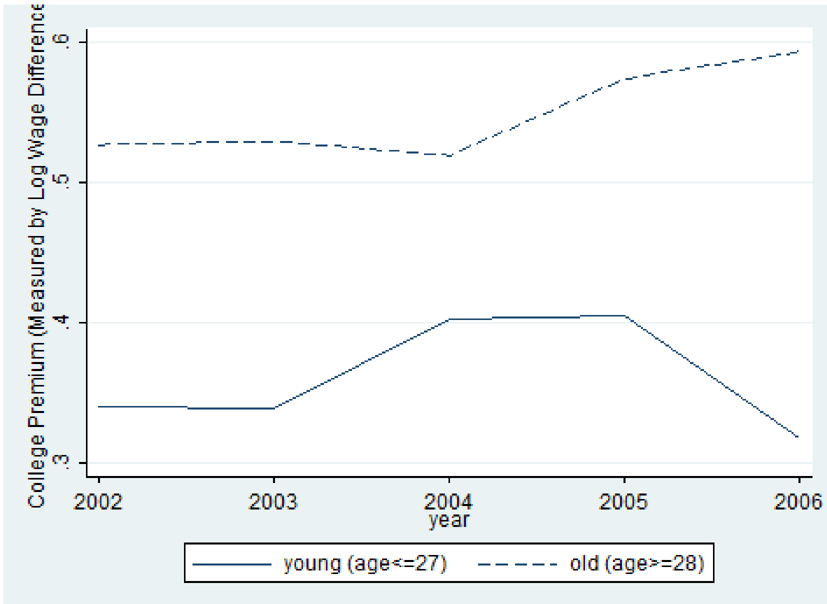

Figure 10: College Premiums from the Survey Data

analysis. Figure 10 shows the college premium trend pattern for young and experienced graduates. Similar to the findings from the resume dataset, the cross-sectional college wage premium started to diverge in 2005 and 2006.

6 Evidence from Other Rapidly Developing Countries

In this section, I will discuss evidence from other rapidly developing countries. Similar to what happened in China, I expect when something (such as a surge in FDI or exports) triggers a large increase in the demand for skill, in the medium run (i.e. 5 to 10 years), the college premium for the young graduates will come down while the college premium for the old is still going up. I will look at all the large economies that have undergone rapid industrialization in the last thirty years. There are 4 large economies (with over 20 million in population) that have achieved an annual growth rate over 7% during the period. They are Korea, Taiwan, Malaysia, and Thailand. India's growth rate was only 5-6% before 2003, and accelerated to around 9% only after 2004. I will also discuss the India case separately.

I did a search for all the papers that have reported the college wage premium by experience and education in these countries. The following is the list of papers I found:

- Taiwan
  - Baraka (1999)
- Korea
  - Kwark and Rhee (1993)
  - Choi and Jeong (2003)
- Malaysia
  - No relevant study is found
- Thailand
  - Mehta et al. (2007)

I will summarize the relevant findings on wage premium from these paper. I will also present the college enrollment and GDP growth data in relevant periods in these economies.

Figure 11 shows the growth of GDP per capita in the four economies (Korea, Taiwan, Thailand, and China). Figure 12 shows the growth of college enrollment rate in these four countries in the relevant periods. The small triangles on the curves indicate the points where these economies are at the similar stage of development as China in late 1990s when GDP per capita reached around $3000 USD, and the demand for skill started to rise rapidly. In Korea and Taiwan, after the initial stage of moving from agriculture to labor-intensive industry in the 60s, both economies started to push into exported-oriented knowledge-intensive and capital-intensive industries. Meanwhile, Korean and Taiwanese governments adopted policies to actively promote the development of such high value-added industries such as building high-tech parks and subsidizing private R&D. With the successful development of these high value-added industries, both economies continued to grow rapidly throughout the 80s, as a result, the demand for skill rose rapidly and so did the college enrollment rate. The case of Thailand is similar, only with a the starting point about 15 years behind of Korea, and about 5 years

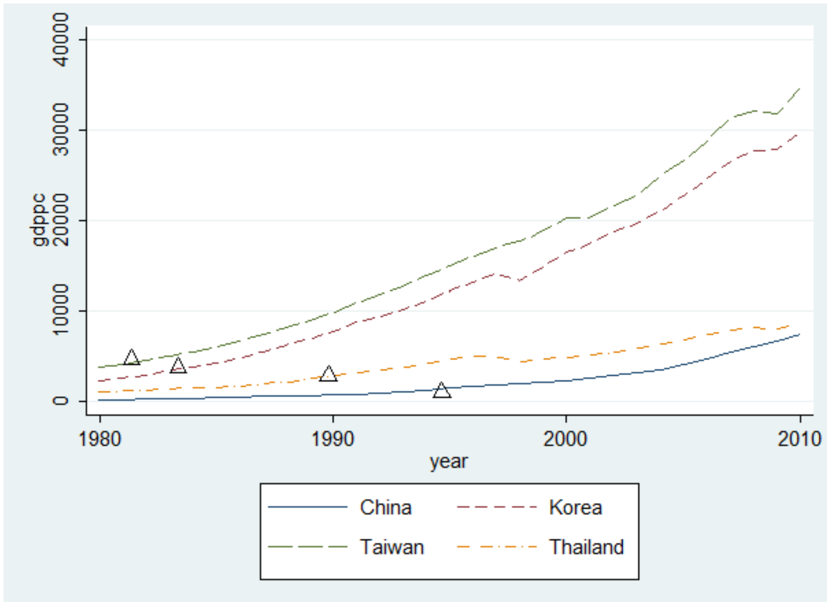

Figure 11: Growth of GDP per capita PPP

ahead of China. In the 80s and early 90s, Thailand moved away from labor intensive industries into high value-added industries, with a rapid growth in FDI, export, and per capita GDP. Similarly, the increase in demand for skill triggered a large increase in college enrollment rate in Thailand in late 1990s. The college enrollment rates of these four countries started to rise rapidly at about the same stage of development when GDP per capita grew from around $3000 to $5000 USD.

Figures 13, 14, and 15 show the college wage premium for the young and the old. In all three countries, the college premium for the young declined as the college premium for the old rose. Compared to China, the magnitude of divergence in college premiums was smaller, because the economy and the enrollment in the three countries did not grow as fast as in China in the relevant periods.

One of the papers, Mehta et al. (2007) also looked at the college wage premium in India. From 1993 to 2004, the college wage premium for both the young and the old increased by about 10%, so there was no divergence of college wage premiums in India at least until 2004. I have not found any more recent studies on Indian wage structure. India started economic liberation in early 1990s, and GDP per capita grew at 5-6% annually from 1994 to 2003,

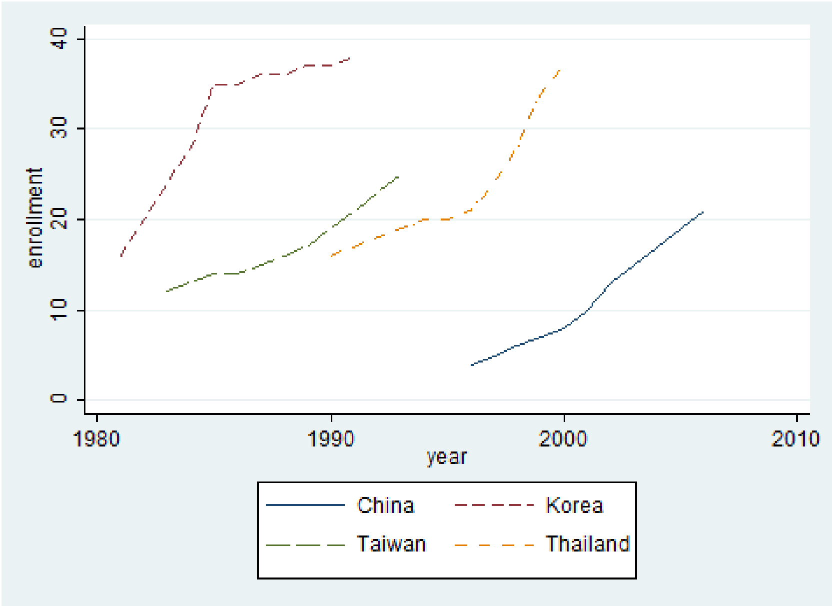

Figure 12: Growth of Enrollment Rate in %

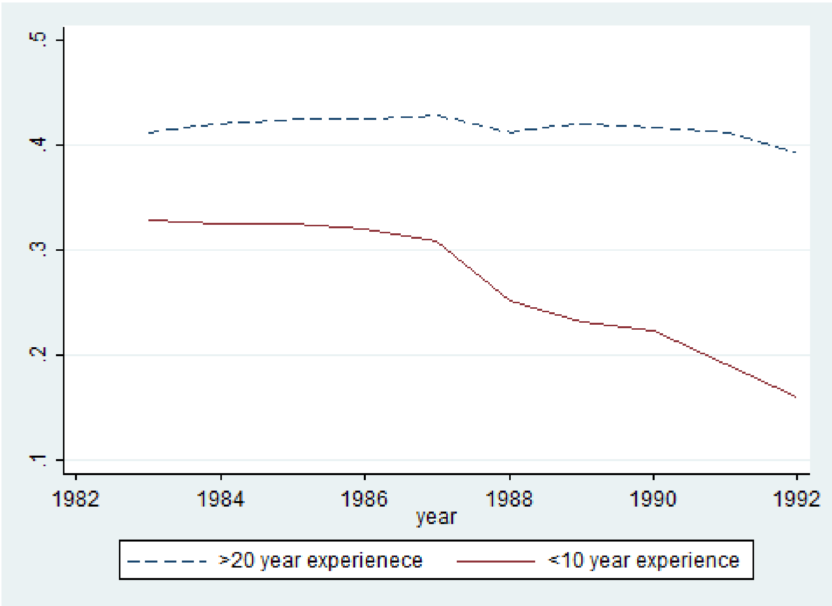

Figure 13: College Wage Premium (Young vs. Old) in Korea

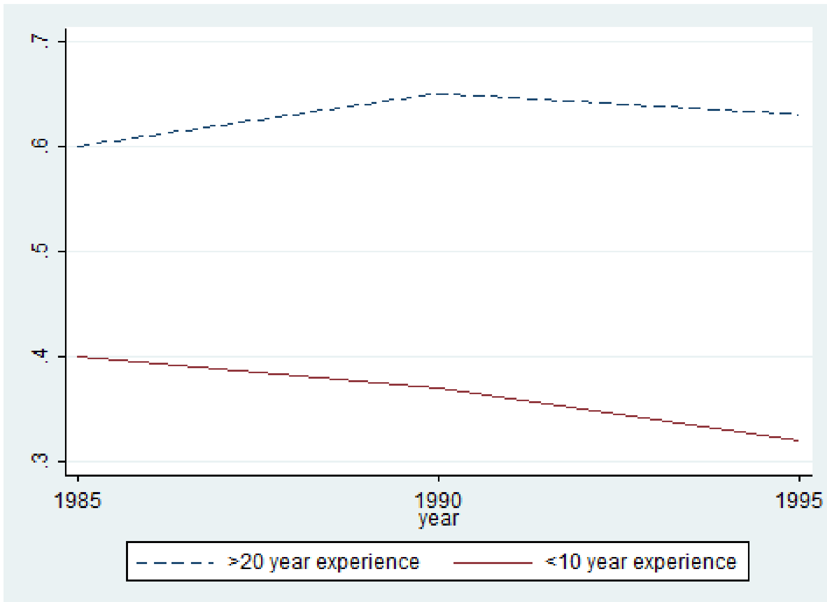

Figure 14: College Wage Premium (Young vs. Old) in Taiwan

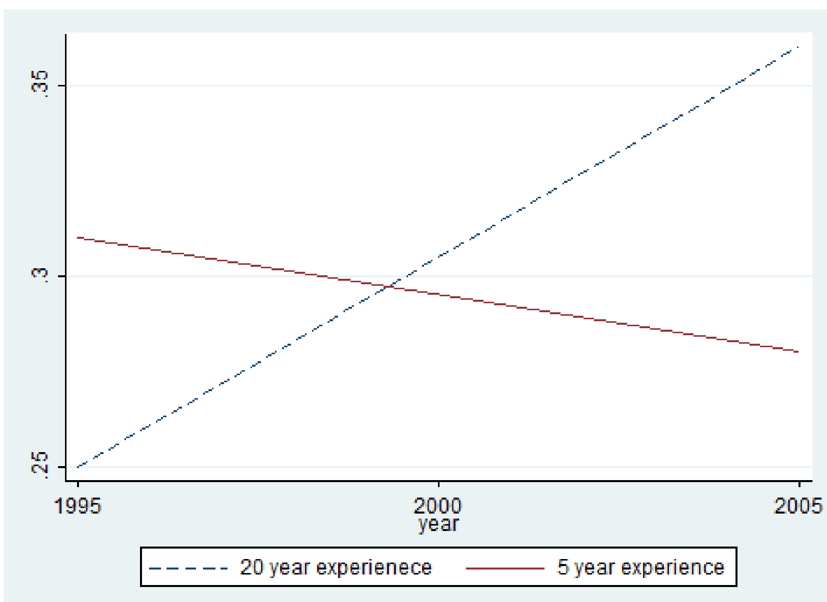

Figure 15: College Wage Premium (Young vs. Old) in Thailand

still significantly slower than the four rapidly developing countries in almost all periods. [5] College enrollment rate in India grew even more slowly, by only 6 percentage points (from 6% to 12%) in the last 20 years. In contrast, the four countries increased their college enrollment rate by at least 13 percentage points in just 10 years. Therefore, it is not surprising that we did not observe a decline of college wage premium for young graduates in India yet. However, in the later half of 2000s, India's economy really picked up speed, attaining an annual growth rate of over 9%, and by 2010 India's GDP per capita PPP reached $3000, the same level of China in 1990s, so in the near future, the familar pattern might emerge. The rising demand for skill could trigger a rapid increase in college enrollment, and the college premium for the young could decline.

7 Policy Simulations Using the Model

7.1 Policies to Subsidize College Education

In this section, I will use this simple dynamic model to analyze ways to subsidize college education. If young people can afford tuition with no credit constraint, and inequality is not a concern, then the government has no reason to subsidize college education, since the private market as analyzed in the base model delivers the optimal outcome. So the goal for subsidizing education is to improve access and reduce inequality.

In general, a tax on the whole population to subsidize college education will raise the return to education, the enrollment rate, and the GDP, but the upward distortion of the return to education will reduce welfare after taking into account the cost of student leisure. Meanwhile, since the tax is regressive, it will increase inequality between skilled and unskilled labor. This is the policy that most developing countries have pursued, but when the college enrollment increases rapidly as the economy develops, the necessary tax rate to fund education becomes increasingly burdensome, especially for the poor.

Another option is to give student loans. It will improve access; however, as indicated in this model, there is a large variance in the realized earning of graduates. Some of them become highly paid h; but many remain to be

[5] Thailand was growing at 10% before the Asian crisis, but slowed down afterwards

m who are relatively low paid, especially when w_m can go down even in the long run. Loans to m will likely be unpaid.

A better alternative is to tax h, (having a progressive tax on wages), and use the money to fund college education. I will analyze the distortion introduced by such a policy and its effect on inequality. I will not analyze in detail the gain of improving access, but will just assume there is some efficiency gain which motivates the government to subsidize college education.

To introduce taxes into the model, I only need to change the equation for v^t as follows.

$$v_t = [w_m^{t+4}+\sum_{j=1}^{\infty}\mu^j(1-\eta)^j w_m^{t+j+4}+\sum_{j=1}^{\infty}\mu^j(1-(1-\eta)^j)w_h'^{t+j+4}](1-\mu)-1-C+(1-\mu)S^t \tag{14}$$

The only difference between the above equation and the base model is that there is an extra term S^t (i.e. subsidy for enrolling in period t), and w_h^t is changed to the after-tax wage $w_h'^t = (1-\tau_h^t)w_h^t$, where τ_h^t is the tax rate for h at time t. There is a per-period budget constraint for the government:

$$\tau_h^t w_h^t L_h^t = S^t * (1-\mu) * G(v^t) \tag{15}$$

Assuming the subsidy completely eliminates the credit constraint problem.

Proposition 7 In the steady state, the new equilibrium with a tax on h to subsidize college education creates no distortion on the college enrollment decision.

Proof: I solve S^t using the budget constraint, and substitute into the expression for v^t. After simplification using steady state variables, the tax rate drops out of the equation. Therefore, the return to education is unaffected by the tax rate in steady state. Q.E.D.

The intuition is simple. Since labor supply is fixed in the model, taxing and subsidizing is just an inter-temporal redistribution for the people who go to college. Therefore the tax and subsidy cancel out and do not affect the return to education in expectation if the economy is the same in every period in the steady state. However, during the transition, this kind of inter-temporal redistribution creates distortion in the dynamic environment.

I will consider two simple schemes. One is a fixed tax rate and variable subsidy. The other is a fixed subsidy but variable taxes. In the fixed tax rate

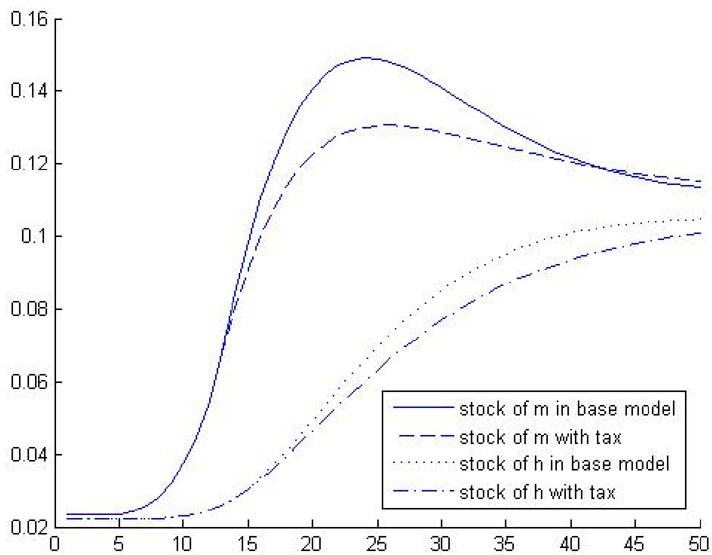

Figure 16: Compare the evolution of the stock of m and h with tax to the baseline model

scheme, compared to the later cohorts, the earlier cohorts receive less subsidy per person. This is because in earlier periods, the tax base was smaller (the stock of h was still growing). So, given the same tax rate, the total funding was less in the earlier periods, but the enrollment rate already shot up, hence the per person level of subsidy was smaller. As the earlier cohorts received less subsidy, but faced the same future tax rate, their expected return to education is lower than that of the later cohorts. By Proposition 7, we know that in the long run, the return to education of the later cohorts is undistorted. Therefore, the expected return to education and enrollment rate of the earlier cohorts are distorted downward.

Figure 16 shows the effect on skill accumulation of such a fixed tax rate scheme. Compared to the zero tax case, the growth of m and h is slower. Figure 17 and Figure 18 show the effect on mean wage and inequality. As shown, inequality is reduced but there is an efficiency loss in the short run. Keep in mind, in these graphs, I am not showing any potential efficiency gain from improving access.

Now let's consider the scheme of having a fixed subsidy (S^t is a constant),

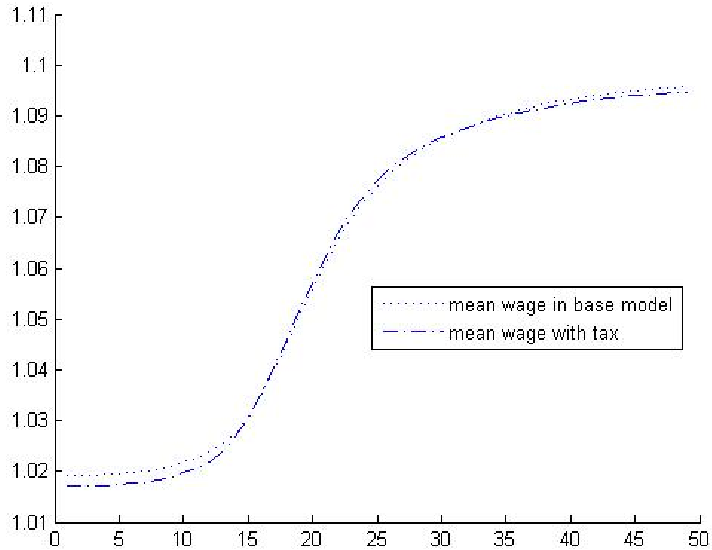

Figure 17: Compare average wage with tax to the baseline model

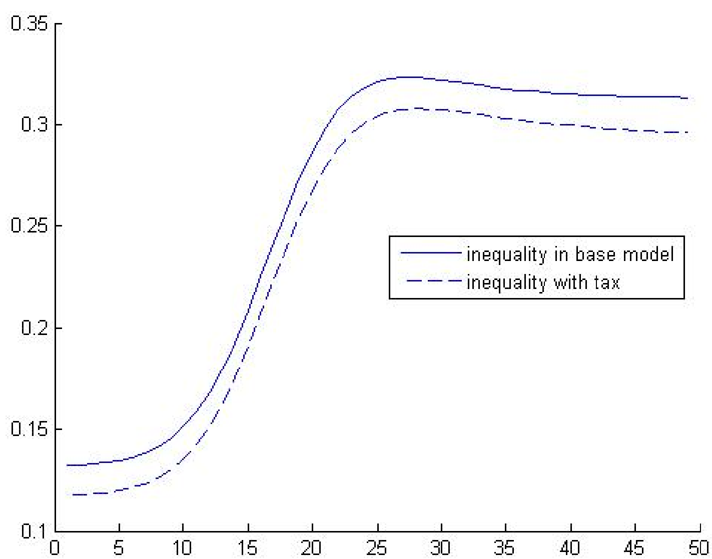

Figure 18: Compare variance of after-tax wage to the baseline model

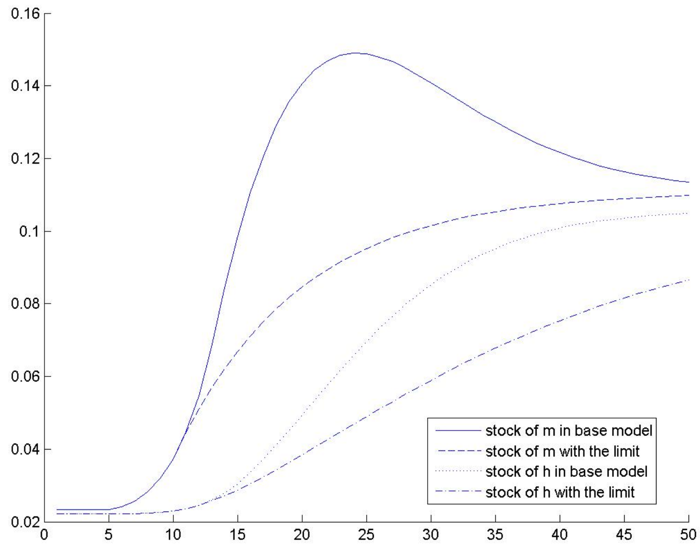

Figure 19: Compare the evolution of the stock of m and h with a limit on enrollment to the baseline model

but variable tax rate. The effect of this scheme is similar. We know in the long run, the expected return to education for later cohorts is undistorted. Compared to the later cohorts, the earlier cohorts are given the same level of subsidy, but will face a higher future tax rate. This is because in early periods, the tax base was smaller but the enrollment rate already shot up, so to fund the subsidy, the tax rate had to be higher. Therefore, the return to education of the earlier cohorts is again distorted downward, creating an efficiency loss.

I have just illustrated that while a subsidy funded by a progressive tax can reduce inequality and improve access, it creates some ineffiency in the short run by reducing the incentive to go to college. Therefore, as the demand for education increases rapidly, the government should not try to fund it through higher taxes if the goal is to maximize efficiency, but rather allow more private colleges.

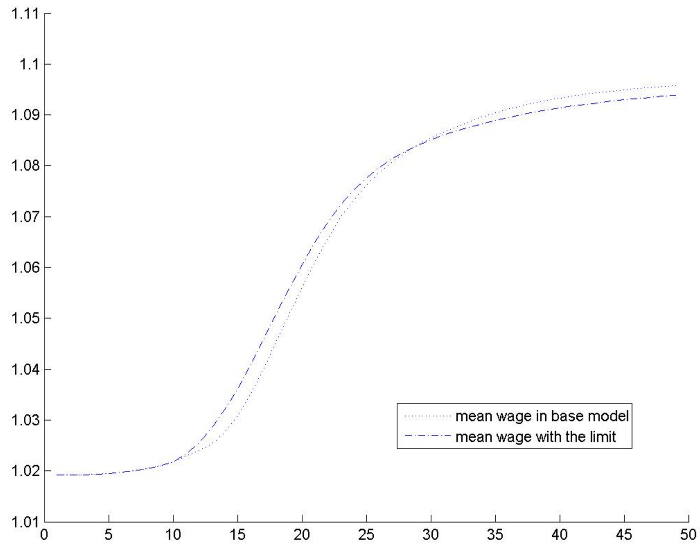

Figure 20: Compare average wage with a limit on enrollment to the baseline model

7.2 Too Many Graduates in China?

Currently, in view of the declining wages of recent graduates, there are heated debates in China about whether China already has too many college graduates, and whether the government should limit growth in college education. This model shows that the phenomenon can be explained by a fully rational model where the enrollment rate and stock of m overshoot, and w_m goes down quite significantly in the short run.

I would argue, on the contrary, that China probably still has too few college graduates due to capacity constraints. The growth of public colleges has been limited by available funding, while the growth of private colleges, though has been very fast, probably still has not caught up with the even faster demand growth. Still about half of the students who took the notoriously high-pressure entrance exam failed to get into any college including 3-year colleges. Any policy that limits the growth of college education, such as restricting entry of private colleges, will cause a loss in overall welfare, and more inequality in the long run.

To analyze the welfare effect of limiting college education, I have solved

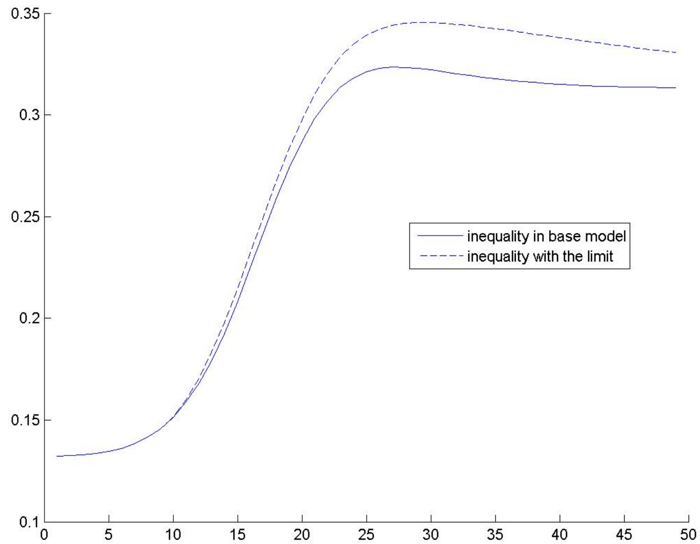

Figure 21: Compare variance of wages with a limit on enrollment to the baseline model

the equilibrium numerically with a limit on the enrollment rate not to overshoot the long term steady state rate (this is effectively setting $G(v) = G(v^s)$ for $v > v^s$). The evolution of the stock of h and m is shown in figure 19. As shown in figure 20, with limited enrollment, mean wage increases in the short run but decreases in the long run. More interestingly, as shown in figure 21, inequality has increased almost in all periods. Overall inequality has two components: first inequality among the graduates; and second inequality between graduates and non-graduates. For the first component, limiting enrollment will decrease return to experience in the short run, but it will increase return to experience in the long run because limiting the supply of m will exacerbate the shortage of h over time. Therefore, inequality among graduates will be lower in the short run and higher in the long run. For the second component, limiting enrollment will almost unambiguously increase inequality between graduates and non-graduates. A combination of the two kinds of inequality will generate the kind of inequality paths shown in the figure; inequality is higher in all periods in our numerical example.

In this paper, I have assumed that there is no growth in the productivity of local firms. However, in the real world, some local firms may indeed grow more like foreign firms i.e. becoming more productive and skill-intensive. This kind of growth is equivalent to adding more high-productivity foreign firms, since in our model, the only difference between foreign firms and local firms is in productivity. The predictions and implications of the model still stand as long as the productivity distribution of all firms in the economy is shifting to the right.

8 Summary

This paper builds a model for the evolution of the labor market in a rapidly growing economy. The model illustrates that with an increase in demand for skill which requires both experience and education, the wages of young graduates may decrease while return to skill increases. The dynamic version of the model also predicts that in the short run, there will be an over-supply of inexperienced college graduates and a higher wage inequality than in the long run. These predictions are consistent with the recent wage data of the labor market in China, and with evidence from several other rapidly developing countries in the last thirty years. From a policy point of view, the model predicts that there are some negative effects of limiting the growth of college education and subsidizing college education with progressive taxes during the transition.

References

Acemoglu, Daron, "Technical Change, Inequality, and the Labor Market," *Journal of Economic Literature*, 2002.

Baraka, Jessica, "Returns to education in Taiwan," *Working Paper*, 1999.

Choi, Kang-Shik and Jin-Ho Jeong, "The Rising Returns to Education and Skill-Biased Technological Change in Korea," *Working Paper*, 2003.

Coughlin, Cletus C. and Eran Segev, "Foreign Direct Investment in China:A Spatial Econometric Study," *World Economy*, 2002.

Freeman, Richard and Lawrence Katz, *Differences and Changes in Wage Structures*, The University of Chicago Press, 1995.

Freestra, R. and G. Hanson, "Foregin investement and relative wages: Evidence from Mexico's maquiladoras," *Journal of International Economics*, 1997.

Heckman, J., "Explaining Rising Wage Inequality: Explorations with a Dynamic General Equilibrium Model of Labor Earnings with Heterogeneous Agents," *Review of Economic Dynamics*, 1998.

Katz, L. and K.M. Murphy, "Changes in relative wages, 1963-1987: Supply and demand factors," *Quarterly Journal of Economics*, 1992.

Kwark, Noh-Sun and Changyong Rhee, "Educational Wage Differentials in Korea," *Seoul Journal of Economics*, 1993.

Mehta, Aashish, Jesus Felipe, Pilipinas Quising, and Shiela Camingue, "Changing Patterns in Mincerian Returns to Education and Employment Structure in Three Asian Countries," *Center for Global Studies, Institute for Social, Behavioral, and Economic Research, UC Santa Barbara*, 2007.

Zhang, Junsen, Yaohui Zhao, Albert Park, and Xiaoqing Song, "Economic returns to schooling in urban China, 1988 to 2001," *Journal of Comparative Economics*, 2005.

Zhao, Y., "Foreign direct investment and relative wages The case of China," *China Economic Review*, 2001.

9 Appendix

Proposition 1 The market clearing equations for m and h have a unique solution: w_m, w_h

Proof: By the definition of v, I know $w_m = \frac{(1+k)(v+1+C)}{1+kw}$, and substitute it in the market clearing equation for m, then I get the following.

$$nB[\frac{(1+k)(v+1+C)}{1+kw}]^{-\sigma} \int_{\frac{(1+k)(v+1+C)}{1+kw}}^{w} \delta^{\sigma-1} dF(\delta) = \frac{1}{k+1} G(v) \tag{16}$$

This equation provides a functional relationship between v and w. The left hand side is decreasing in v, and the right hand side is increasing in v.

When $v = 0$, the left hand side is positive, the right hand side is 0. When v is very big such that $w_m = \frac{(1+k)(v+1+C)}{1+kw} >= w$, the left hand side is $<= 0$, the right hand side is > 0. Therefore given w, the above equation has a unique solution; in other words, v and w is a one to one functional relationship. Moreover, increasing w shifts the demand curve up, while leaving the supply curve unchanged, therefore the functional relationship can be denoted by an increasing function $w = w^*(v)$.

By the definition of v, I know $w_h = \frac{(1+k)(v+1+C)}{1/w+k}$. I substitute w by $w^*(v)$ to get $w_h = \frac{(1+k)(v+1+C)}{1/w^*(v)+k}$ which is an increasing function of v, denote it as $w_h^*(v)$). Substituting $w_h = w_h^*(v)$ and $w = w^*(v)$ into the clearing condition for h, I get the following equation.

$$B(w_h^*(v))^{-\sigma}[n\int_{w^*(v)}^{\Delta}(\delta^2)^{\sigma-1}dF(\delta) + n_f(w_h^*(v))(\Delta^2)^{\sigma-1}] = \frac{k}{k+1}G(v) \quad (17)$$

Given $w_h^*()$ and $w^*()$ are increasing and n_f is decreasing, the whole demand side is decreasing in v, and the right hand side is increasing. Further, when $v = 0$, the left hand side is positive, the right hand side is 0; when v is big enough such that $w^*(v) >= \Delta$, then the left hand side is $<= 0$, and the right hand side is positive. Therefore a unique solution v exists, and $w = w^*(v)$, $w_h = \frac{(1+k)(v+1+C)}{1/w+k}$, and $w_m = \frac{(1+k)(v+1+C)}{1+kw}$, and the solution w is between (w_m, Δ). Q.E.D.

Proposition 2 When n_f increases, w, v, w_h increase, but w_m is ambiguous.

Proof: I want to show that w, w_h, v all increase. First I will show that w increases. By contradiction, suppose not, if w decreases, observing the equation:

$$nB[\frac{(1+k)(v+1+C)}{1+kw}]^{-\sigma}\int_{\frac{(1+k)(v+1+C)}{1+kw}}^{w}\delta^{\sigma-1}dF(\delta) = \frac{1}{k+1}G(v) \quad (18)$$

Decreasing w shifts the left hand side down, thus decreasing v. Then, $w_h = \frac{(1+k)(v+1+C)}{1/w+k}$ also decreases, but this would violate the market clearing condition for h as follows.

$$B(w_h)^{-\sigma}[n\int_{w}^{\Delta}(\delta^2)^{\sigma-1}dF(\delta) + n_f(\Delta^2)^{\sigma-1}] = \frac{k}{k+1}G(v) \quad (19)$$

The left hand side increases from the previous equilibrium since w_h decreases, w decreases, and n_f increases, but the right hand side decreases from the previous equilibrium as v decreases. This is a contradiction, so w must increase.

I know w increases. By looking at the supply/demand equation for m, I know that as w increases, the left hand side shift up, so v also increases. Then $w_h = \frac{(1+k)(v+1+C)}{k+1/w}$ must increase too. So I have shown that w, v, and w_h all increases. However, $w_m = w_h/w$ is ambiguous Q.E.D.

Proposition 3 With more foreign firms, if $G'()$ is sufficiently small then w_m increases, and if $G'()$ is sufficiently large then w_m decreases.

Proof: To analyze this further I try to calculate dw_m/dw. I will implicitly differentiate the demand equation for m,

$$nBw_m^{-\sigma} \int_{w_m}^{w} \delta^{\sigma-1} dF(\delta) = \frac{1}{k+1} G(v) \tag{20}$$

where $v = w_m(1+kw)/(1+k) - 1 - C$

This equation establishes a functional relationship between w and w_m. I already know with the arrival of the foreign firms, w increases, therefore to sign dw_m, all I need to do is to sign dw/dw_m. Implicitly differentiating the above equation, I can obtain an expression for dw_m/dw.

$nBw_m^{-\sigma-1}(-\sigma) \int_{w_m}^{w} \delta^{\sigma-1} dF(\delta) + nBw_m^{-\sigma}(-w_m^{\sigma-1} f(w_m)) + nBw_m^{-\sigma} w^{\sigma-1} f(w) \frac{dw}{dw_m} = \frac{1}{k+1} G'(v)[\frac{1+kw}{1+k} + \frac{w_m k}{1+k} \frac{dw}{dw_m}]$

After re-arranging and simplification, I get an expression for $\frac{dw}{dw_m}$

$\frac{dw}{dw_m} = \frac{-\sigma nBw_m^{-\sigma-1} \int_{w_m}^{w} \delta^{\sigma-1} dF(\delta) - nBw_m^{-1} f(w_m) - \frac{1}{k+1} G'(v) \frac{kw_m}{1+k}}{\frac{1}{k+1} \frac{1+kw}{1+k} G'(v) - nBw_m^{-\sigma} w^{-\sigma-1} f(w)}$

By examing the above expression, I know that when $G'()$ is sufficiently small, i.e. the supply is very elastic, $dw/dw_m > 0$, so when w increases, w_m increases. On the other hand when $G'()$ is sufficiently large, i.e. the supply is very elastic, then $dw/dw_m < 0$, so when w increases, w_m decreases.

Proposition 4 If w_m decreases in the new equilibrium, then there exists a cut off $\overline{\delta} = w^0 w_m^0 / w_m^1$, such that when $\delta < \overline{\delta}$, firms are better off, and when $\delta > \overline{\delta}$ firms are worse off.

Proof:

The firms that are still hiring h are worse off, because w_h increases. The firms that continue to hire m are better off since w_m is lowered. The firms that switch from hiring l to hiring m are better off, because previously they made $\frac{1}{\sigma-1}B$, and now they make $\frac{1}{\sigma-1}B(\frac{\delta}{w_m^1})^{\sigma-1}$. The profit is greater iff

$\delta > w_m^1$, which is true for them to hire m now, therefore they are better off.

Now consider the firms with $\delta \in (w^0, w^1)$. they switch from hiring h to hiring m. Previously they made $\frac{1}{\sigma-1}B(\frac{\delta^2}{w_h^0})^{\sigma-1}$. Now they make $\frac{1}{\sigma-1}B(\frac{\delta}{w_m^1})^{\sigma-1}$. The firm is worse off if $\delta > w_h^0/w_m^1 = w^0 * w_m^0/w_m^1$, or better off if $\delta < w^0 w_m^0/w_m^1$. Q.E.D.

Proposition 4a With the arrival of the medium productivity firms, v, w_m all increase, and w decreases.

Proof:

Let's say there are n' more medium productivity firms whose δ is d, and for simplicity, let's assume $w_m < d < w$, such that these medium productivity firms will hire m workers in the new equilibrium. Therefore, the two equations that pin down w_m and w are:

$$B[nw_m^{-\sigma} \int_{w_m}^{w} \delta^{\sigma-1} dF(\delta) + n'\Delta^{\sigma-1}] = \frac{1}{k+1} G(v) \tag{21}$$

This is the supply/demand equation for m in the new equilibrium, the only difference is the term of $n'\Delta^{\sigma-1}$ for the new domestic firms. The supply/demand equation for h is unchanged.

$$Bw_h^{-\sigma}[n \int_{w}^{\Delta} (\delta^2)^{\sigma-1} dF(\delta) + n_f(\Delta^2)^{\sigma-1}] = \frac{k}{k+1} G(v) \tag{22}$$

I will show w will decrease by contradiction. Suppose not, w increases. Substitute $w_h = \frac{(1+k)(v+1+C)}{1/w+k}$ into the above equation, and I get:

$$B[\frac{(1+k)(v+1+C)}{1/w+k}]^{-\sigma}[n \int_{w}^{\Delta} (\delta^2)^{\sigma-1} dF(\delta) + n_f(\frac{(1+k)(v+1+C)}{1/w+k})(\Delta^2)^{\sigma-1}] = \frac{k}{k+1} G(v) \tag{23}$$

If w increases, it will shift the left hand side down, thus v will decrease. Then $w_m = \frac{(1+k)(v+1+C)}{1+kw}$ will decrease, but this would violate the supply-demand equation for m. In the equation, the left hand side increases from the previous equilibrium as w_m decreases and there is a new positive term for the demand from the new firms, but the right hand side decreases as v decreases, which a contradiction. So w must decrease.

Again using the clearing condition for h, a decreasing w shifts the left hand side up, therefore v increases, and $w_m = \frac{(1+k)(v+1+C)}{1+kw}$ must also increase. Q.E.D.

Proposition 4b With the arrival of the medium productivity firms, w_h is ambiguous. When $G()$ is sufficiently inelastic, w_h increases; when $G()$ is sufficiently elastic, w_h decreases.

Proof: Substitute $v = \frac{w_h(k+1/w)}{1+k} - 1 - C$ into the demand equation for h,

$$B(w_h)^{-\sigma}[n \int_w^{\Delta} (\delta^2)^{\sigma-1} dF(\delta) + n_f] = \frac{k}{k+1} G(\frac{w_h(k+1/w)}{1+k} - 1 - C). \quad (24)$$

This equation establishes a functional relationship between w and w_h, I already know that with the arrival of the medium firm, w decreases, then to sign dw_h, all I need to do is to sign dw/dw_h. Implicitly differentiating the above equation by w, I will get an expression dw_h/dw. It can be shown that when $G'()$ is sufficiently small, i.e. the supply is very inelastic, then $dw_h/dw < 0$, therefore with the new arrivals, w decreases, then w_h must increase. On the other hand when $G'()$ is sufficiently big, i.e. the supply is very elastic, then $dw_h/dw > 0$, therefore with the new arrivals, w decreases, then w_h must decrease.

Intuitively, when w decreases, there are two effects. The first effect is that it increases the demand on the left hand side of the equation, this tends to increase w_h; the second effect is shifting supply curve up this tends to reduce w_h. When the supply curve is very elastic, the second effect dominates, w_h decreases. This says that, if the education system is really good at producing m workers, more local firms that are moderately productive can reduce w_h thus benefit more high ability firms as well. In the extreme case, when supply of m is perfectly elastic at some point v_0, then when w decreases, w_h also has to decrease to keep $v = v_0$. On the other hand, if $G(v)$ is perfectly inelastic, i.e. the supply of m, and h is fixed, then as w decreases, the demand for h increases, therefore bidding up w_h.

Proposition 4d When there are more local firms of medium productivity and supply of education is sufficiently elastic, there exists a productivity cutoff, such that all local firms with a higher productivity will be better off. All existing foreign firms are better off.

From Proposition 4b, I know w_m increases, but w_h decreases.

As w_h decreases, clearly every foreign firm which hires h is more profitable.

The domestic firms that are still hiring m are worse off, because w_m increases. The firms that continue to hire h are better off since w_h is lowered. The firms switching from hiring m to hiring l are worse off, because now they make $\frac{1}{\sigma-1}B$, and previously they made $\frac{1}{\sigma-1}B(\frac{\delta}{w_m^0})^{\sigma-1}$. The profit is smaller

Table 5: Impact of additional medium productivity firm on existing firms when $w_h^1 < w_h^0$, and $w_m^1 > w_m^0$

Firms with $\delta \in$	Workers	Profit Before	Profit After	Comparison
$(1, w_m^0)$	l	$\frac{1}{\sigma-1}B$	same	same
(w_m^0, w_m^1)	$m->l$	$\frac{1}{\sigma-1}B(\frac{\delta}{w_m^0})^{\sigma-1}$	$\frac{1}{\sigma-1}B$	worse off
(w_m^1, w^1)	m	$\frac{1}{\sigma-1}B(\frac{\delta}{w_m^0})^{\sigma-1}$	$\frac{1}{\sigma-1}B(\frac{\delta}{w_m^1})^{\sigma-1}$	worse off
$(w^1, w^1 * w_m^1/w_m^0)$	$m->h$	$\frac{1}{\sigma-1}B(\frac{\delta}{w_m^0})^{\sigma-1}$	$\frac{1}{\sigma-1}B(\frac{\delta^2}{w_h^1})^{\sigma-1}$	worse off
$(w^1 * w_m^1/w_m^0, w^0)$	$m->h$	$\frac{1}{\sigma-1}B(\frac{\delta}{w_m^0})^{\sigma-1}$	$\frac{1}{\sigma-1}B(\frac{\delta^2}{w_h^1})^{\sigma-1}$	better off
(w^0, Δ)	h	$\frac{1}{\sigma-1}B(\frac{\delta^2}{w_h^0})^{\sigma-1}$	$\frac{1}{\sigma-1}B(\frac{\delta^2}{w_h^1})^{\sigma-1}$	better off

iff $\delta < w_m^0$, which is true for them to hire l initially, therefore they are worse off.

Now consider the firms with $\delta \in (w^1, w^0)$ that switch from hiring m to hiring h. Now they make profit $\frac{1}{\sigma-1}B(\frac{\delta^2}{w_h^1})^{\sigma-1}$. Previously they made $\frac{1}{\sigma-1}B(\frac{\delta}{w_m^0})^{\sigma-1}$. The firm is better off if $\delta > w_h^1/w_m^0 = w^1 * w_m^1/w_m^0$ or worse off if $\delta < w^1 w_m^1/w_m^0$. Q.E.D

Proposition 5 A dynamic equilibrium exists.

Proof: I denote enrollment rate $g^t \equiv G(v^t)$, I will establish a mapping between g^t to $g^{t'}$, by the following procedure: from the enrollment path g^t, I can compute the evolution of L_m^t and L_h^t, and then using the period by period demand and supply equations, I can compute w_m^t and w_h^t from which I can compute v^t, finally I can get a new set of $g'^t = G^{-1}(v^t)$.

I define a Banach space for all sequences bounded by $[0, 1]$ with a sup norm. Clearly this space is compact and convex. The above-defined mapping can be easily shown to be a continuous self-mapping. Therefore I can use the Schauder Fixed Point Theorem to establish the existence of a fixed point. Q.E.D.

Lemma 1 In period by period demand/supply equations for h and m, if both L_h and L_m decrease, and one of them decreases strictly, then w_m and w_h increases strictly. (Conversely if both L_h and L_m increase, and one of them increases strictly then w_m and w_h decreases strictly.)

Proof: By contradition. If w_m decreases, then by the equation for m, w must also decrease, to keep L_m decrease, therefore $w_h = ww_m$, must also decrease. However, if w decreases, then w_h must increase to keep L_h decrease, a contradition. Therefore, w_m increases, Similarly, to show that w_h increases, I also use proof by contradiction. Suppose w_h decreases, then by the equation for h, the left hand side increases, but the right hand side (L_h) decreases, which is a contradiction. Therefore, w_h increases. The converse can be easily shown similarly. Q.E.D.

Lemma 2 Given the period by period demand and supply equations, denote W_m as the solution for the wage for m given argument L_m and L_h. Then, the partial $-\frac{\partial W_m}{\partial L_h}$ is bounded below.

Proof: The two equations are:

$$nBw_m^{-\sigma}\int_{w_m}^{w_h/w_m}\delta^{\sigma-1}dF(\delta) = L_m \tag{25}$$

$$Bw_h^{-\sigma}[n\int_{w_h/w_m}^{\Delta}(\delta^2)^{\sigma-1}dF(\delta) + n_f(\Delta^2)^{\sigma-1}] = L_h \tag{26}$$

Denote function $M(,)$ as $M(w_m, w_h) = nBw_m^{-\sigma}\int_{w_m}^{w_h/w_m}\delta^{\sigma-1}dF(\delta)$,

and denote function $H(,)$ as $H(w_m, w_h) = Bw_h^{-\sigma}[n\int_{w_h/w_m}^{\Delta}(\delta^2)^{\sigma-1}dF(\delta) + n_f(\Delta^2)^{\sigma-1}]$

and implicitly differentiate both equations on both sides by L_h, holding L_m constant. I have

$M_1\frac{\partial w_m}{\partial L_h} + M_2\frac{\partial w_h}{\partial L_h} = 0$

$H_1\frac{\partial w_m}{\partial L_h} + H_2\frac{\partial w_h}{\partial L_h} = 1$

I solve the above equations to get, $\frac{\partial W_m}{\partial L_h} = (H_1 - H_2M_1/M_2)^{-1}$.

I know $M_1 < 0$, $M_2 > 0$, $H_1 < 0$ and $H_2 > 0$, and by Lemma 1, I know $\frac{\partial W_m}{\partial L_h} < 0$

I know the absolute values of M_1, M_2, H_1 and H_2 are bounded below and above (by assumption, all the wages are bounded between $[1, \Delta^2]$, and $f() < \infty$) . Therefore the absolute value $H_1 - H_2M_1/M_2$ is bounded from above, and hence, the absolute value of $\frac{\partial W_m}{\partial L_h}$ is bounded below.

Lemma 3 If $v^t > v^s$ and both L_m^t and L_h^t converge from below then

$$\lim_{t\to\infty}\frac{L_h^s - L_h^t}{L_m^s - L_m^t} = \infty \tag{27}$$

Proof: Expanding the recursive formula for L_h I have $L_h^t = \sum_{j=1}^{t}\eta\mu^j L_m^{t-j}$, and hence $L_h^s - L_h^t = \sum_{j=1}^{t}\eta\mu^j(L_m^s - L_m^{t-j})$.

Also I know that $L_m^s - L_m^t < \mu(1-\eta)(L_m^s - L_m^{t-1})$ for all t. This is because $L_m^s - L_m^t = \mu(1-\eta)(L_m^s - L_m^{t-1}) + (1-\mu)(G(v^s) - G(v^t))$, and $G(v^t) > G(v^s)$.

Recurisively applying the above inequality, I have

$$L_m^s - L_m^{t-j} > \mu^{-j}(1-\eta)^{-j}(L_m^s - L_m^t) \tag{28}$$

I subsitute the above inequality into the expression for $L_h^s - L_h^t$ to get $L_h^s - L_h^t > (L_m^s - L_m^t)\sum_{j=1}^{t}(1-\eta)^{-j}$ therefore $\frac{L_h^s - L_h^t}{L_m^s - L_m^t}$ will go to infinity as t goes to infinity. Q.E.D.

Proposition 6 With this setup, if $G'() > 0$, there exists some period τ, in which $L_m^\tau > L_m^s$.

Proof:

By contradiction, suppose L_m never overshoots, i.e. $L_m^t <= L_m^s$ for all t, then L_h also never overshoots($L_h^t <= L_h^s$ for all t). Actually, the inequality is strict, as long as $L_h^0 < L_h^s$ (this can be shown easily by induction on equation (8)).

Then, I will show that for some τ large enough, for all $t > \tau$, $w_m^t > w_m^s$ and $w_h^t > w_h^s$. To show this, since the demand for labor n_f^t will be a constant for t large enough ($t > \tau$) by assumption, and since the supply never overshoots, ($L_m^t <= L_m^s$ and $L_h^t < L_h^s$, for $t > \tau$), by Lemma 1, the wages will always be higher than the long term values for t large enough. i.e $w_m^t > w_m^s$, $w_h^t > w_h^s$, $v^t > v^s$

Next, I will show this is not possible when $G'() > 0$.

In the first step, I will establish that

given $L_m^{t+1} < L_m^s$ for all t, I have

$$\frac{G(v^{t-4}) - G(v^s)}{L_m^s - L_m^t} < \frac{\mu(1-\eta)}{1-\mu} \tag{29}$$

This follows from simple algebra as follows:

$L_m^{t+1} = (1-\mu)G(v^{t-4}) + \mu(1-\eta)L_m^t < L_m^s$

Re-arranging, I get

$(1-\mu)G(v^{t-4}) < L_m^s - \mu(1-\eta)L_m^t = \mu(1-\eta)(L_m^s - L_m^t) + (1-\mu+\mu\eta)L_m^s$

Given $(1-\mu+\mu\eta)L_m^s = (1-\mu)G(v^s)$

I have $G(v^{t-4}) - G(v^s) < \frac{\mu(1-\eta)}{1-\mu}(L_m^s - L_m^t)$.

Second, I know that $v^{t-4} - v^s > (1-\mu)\mu^4(1-\eta)^4(w_m^t - w_m^s)$, by the formula for v_t, for all $t > \tau$, $w_m^t > w_m^s$ and $w_h^t > w_h^s$.

Third, apply the mean value theorem, $w_m^t - w_m^s = (-\frac{\partial W_m}{\partial L_m})(L_m^s - L_m^t) + (-\frac{\partial W_m}{\partial L_h})(L_h^s - L_h^t)$. Here W_m is the function that solves w_m taking argument L_m and L_h in the period by period supply and demand equations. The partials are evaluated at some values between L_m^t, L_h^t and L_m^s, L_h^s.

Combining the second and third step, I have

$v^{t-4} - v^s > (1-\mu)\mu^4(1-\eta)^4[(-\frac{\partial W_m}{\partial L_m})(L_m^s - L_m^t) + (-\frac{\partial W_m}{\partial L_h})(L_h^s - L_h^t)]$.

and again applying the mean value theorem on $G()$, I get:

$G(v^{t-4}) - G(v^s) > G'()(1-\mu)\mu^4(1-\eta)^4[(-\frac{\partial W_m}{\partial L_m})(L_m^s - L_m^t) + (-\frac{\partial W_m}{\partial L_h})(L_h^s - L_h^t)]$

Dividing both sides by $L_m^s - L_m^t$, I have

$$\frac{G(v^{t-4}) - G(v^s)}{L_m^s - L_m^t} > G'()(1-\mu)\mu^4(1-\eta)^4[(-\frac{\partial W_m}{\partial L_m}) + (-\frac{\partial W_m}{\partial L_h}).\frac{L_h^s - L_h^t}{L_m^s - L_m^t}] \tag{30}$$

By Lemma 2 $-\frac{\partial W_m}{\partial L_h}$ is uniformly bounded from below. Also by Lemma 3, $\frac{L_h^s - L_h^t}{L_m^s - L_m^t}$ goes to ∞. Observing the above equation, as long as $G'() > 0$ evaluated near v_s, $\frac{G(v^{t-4}) - G(v^s)}{L_m^s - L_m^t}$ will go to ∞. This contradicts to the first step where I have shown that $\frac{G(v^{t-4}) - G(v^s)}{L_m^s - L_m^t} < \frac{\mu(1-\eta)}{1-\mu}$ for t big enough. Therefore L_m^t will overshoot. Q.E.D