Why Do Large Firms Willingly Pay High Wages in Developing Countries?[*]

Hongbin Cai [†]   Miaojun Wang [‡]   Se Yan [§]

Abstract

Using a simple game-theoretical model, this paper provides a new explanation for why large firms in developing economies may *willingly* pay higher wages than the market wage rate. We show that large firms can strategically create entry barriers to the modern sector by setting high wage standards. They may do so to reduce competition, or to distort the benevolent government's resource allocation to their benefits. Focusing on the latter case, we also show that the size of the primitive sector will be larger than the efficient level, and public resource allocation will be biased in favor of incumbent large businesses despite the benevolent nature of the government.Using a comprehensive survey of Chinese industrial firms, we find that industrial concentration is positively correlated with the size-wage effect, which is consistent with our theoretical prediciton.

JEL Classification: J21, J31, J42, L11

Keywords: size-wage effect, entry deterrence, development

[†]Guanghua School of Management and IEPR, Peking University, Beijing, China 100871. Contact: (86)10-6276-5132, hbcai@gsm.pku.edu.cn.

[‡]Center for Research of Private Economy and College of Economics, Zhejiang University. Contact: (86)-10-87952835, wangmiaojun@gsm.pku.edu.cn.

[§]Guanghua School of Management, Peking University, Beijing, China 100871. Contact: (86)10-62757764, seyan@gsm.pku.edu.cn.

1 Introduction

There is much evidence of the employer size wage effect in that large firms pay higher wages than small firms after controlling for worker characteristics, which cannot be accounted for by conventional explanations such as efficiency wage, monitoring technology, unionization, and compensating wage differentials (e.g., Brown and Medoff, 1989; Schmidt and Zimmermann, 1991; Schaffner, 1998; Troske, 1999). Moreover, the wage premium of large firms is larger in developing countries than in developed countries (e.g., Little, Mazumdar and Page, 1987, Velenchik, 1997, Schaffner 1998, Strobl and Thornton, 2002). This latter finding is particularly puzzling, given that in many developing countries there are abundant workers seeking job opportunities. One would expect that abundant labor supply would put pressure on wage premium of large firms.

Why do large firms pay high wages in developing countries with abundant labor supply? In this paper, we present a simple model in which large firms *willingly* pay higher wages than the market wage. In the model, there are two sectors, modern and primitive, defined by different technologies used in these sectors. There are a small number of large firms that are the incumbents in the modern sector, and a large number of small entrepreneurs who can enter either sector. However, a small entrepreneur incurs some entrance cost to enter the modern sector, and the entrance cost differs among small entrepreneurs. The sequence of move in the model is as follows. At date one, the incumbent firms set a wage standard in the modern sector. At date two, small entrepreneurs choose which sector to enter and a benevolent government decides on resource allocation between the modern and primitive sectors. Government resources in a sector are critical inputs for firms in the sector. In this setting, we show that under some conditions the equilibrium outcome of the game has the following properties: (i) the incumbent firms in the modern sector set a wage standard which is higher than the market wage rate; (ii) some small entrepreneurs will not enter the modern sector although they should do so in the first-best solution; and (iii) the benevolent government's policy is biased in favor of incumbent big businesses.

The basic idea of our model is very simple. Even though paying higher wages to workers directly reduce profit, doing so allows the incumbent firms in the modern sector to limit entrance to the modern sector by small entrepreneurs with relatively high entrance costs. Reducing the size of the modern sector can be beneficial to the incumbents because the relative sizes of the two sectors affect how the benevolent government allocates resources. Even though the government will allocate less total resources to a smaller modern sector, the average resources per firm can be higher in the modern sector under reasonable conditions. As long as government resources exhibit sufficient exclusivity, the incumbent firms in the modern sector will have strong incentives to raise wage rates to induce the benevolent government to bias its resource allocation in favor of incumbent large businesses.

A key assumption of our model is that large incumbent firms in the modern sector can raise the labor costs of entrant firms in the modern sector by paying high wages, thus forcing entrant firms to pay high wages as well. Let us call this the "wage equalization effect." There are two justifications for this effect. First, there is a class of efficiency wage theories in the literature that have the feature of inter-firm relative wage equalization. For example, Summers (1988) argues that "increasing relative wage raises productivity" because, for one reason, paying wage rates higher than workers' outside opportunities reduces turnovers and thus saves on monitoring, recruiting and training costs.[1] Another justification for the wage equalization effect is provided by a large number of fair wage theories. As an early well-known example, Akerlof and Yellen (1988, 1990) propose "the fair wage/effort hypothesis" which argues that the perception of fairness by workers affected their effort.[2] When workers compare wages in different firms to determine what is the fair wage rate, high wage rates offered by some firms will raise other firms' labor costs and force them to increase their wages as well. Recently a large body of literature on behavior economics emphasizes the importance of fairness and other psycological factors in wage-setting and other organizational designing situations (see for example, Fehr, Kirchsteiger, and Riedl, 1993, and Rabin, 1998 for a survey). All these fairness considerations will tend to equalize wages across similar firms.

The wage equalization effect makes it possible that firms in the modern sector raise their own wage rates in order to raise the labor costs of potential entrants.[3] But is there any evidence that wage rates are used to limit entrance? In fact, in the analysis of the Pennington case which centered exactly on this issue, Williamson (1968) argued that this was not merely a theoretical possibility but a real threat to fair competition. In this legal case, Pennington, as one owner of a relatively small coal company, alleged that large coal companies imposed uniformly high wage rates (with the help of the United Mine Workers) in order to drive small coal producers out of business. Williamson showed that with the help of unions to coordinate high wage standards, indeed large businesses could limit entrance of small entrepreneurs. If this can happen in the United States, it can certainly take place in developing economies where the conditions are much more suitable for such predatory behavior.

It is easy to see that, with the ability to "raising rivals' costs" (Salop and Scheffman, 1983) through setting high wages, the incumbent large firms may use this strategic tool to limit market

---

[1]Specifically, Summers postulates the following equation: $e = (w - x)^a$, where e is effort, w is wage, x is the outside opportunities of workers, and $a \in [0, 1]$ measures the importance of relative wage comparison. Clearly, if some firms increase wage rates, the outside opportunities of workers in other firms increase and thus the labor costs of those firms increase.

[2]Specifically, Akerlof and Yellen postulated the following equation: $e = min(w/w^*, 1)$, where normal effort is 1 and w^* is the fair wage perceived by workers.

[3]Chu and Masson (1990) show that incumbent firms may use high wage rates to signal their strength to potential entrants in order to deter entrance.

competition if potential entrants produce competing products.[4] We do not focus on this motive, as it is self-evident.[5] Rather, we focus on a more subtle motive that is perhaps more relevant in developing countries. When studying wage differentials across sectors, one question naturally arises: why do the incumbent firms want to limit entrance to the broadly defined modern sector where many potential entrants do not compete directly with them? In this paper, we identify another motive for predation by the incumbent large firms in the modern sector even when potential entrants are not in their industries. The motive is to distort the government's resource allocation decisions.

Governments in developing countries often have quite limited resources, but face tremendous tasks to improve physical, social and economic infrastructures. Thus, government policies of allocating scarce resources have large impacts on the efficiency of firms in different sectors. For example, for a fixed amount of total education expenditure, more expenditure on higher education is likely to benefit modern firms more, while more spending on elementary education benefits traditional firms more. As another example, given a fixed budget to spend on improving the transportation system of the country, building an air transportation system that connects large cities benefits large firms more, while improving local transportation infrastructure of many medium- or small-sized cities benefits small firms more. Because such government resources are quite limited in developing countries, they are likely to have a strong degree of exclusivity in that more users will reduce the marginal product of these resources to all users. Therefore, the incumbent firms in the modern sector will feel the pressure of potential entrants to the modern sector competing with them for the use of the scarce government resources. This can motivate the large incumbent firms in the modern sector to limit entrance by setting high wage standards.

Governments in developing countries are often said to be captured or corrupted by large firms so their policies are often biased in favor of incumbent big businesses. Instead of direct capturing through lobbying and bribery, our analysis shows that large businesses can indirectly capture the government through distortional market behavior. For example, it has been an often heard criticism that developing countries tend to overemphasize higher education.[6] In this case, it is not clear how

---

[4]Compared with prices or quantities, wage rates as an instrument to deter entrance have certain advantages. First, prices, quantities and other similar predatory instruments can be detected relatively easily, which may result in resentment or even legal actions by the government. In contrast, a high wage standard is more likely to be welcomed by the government since it increases income of workers in the modern sector. Secondly, potential entrants may not enter the same industries the incumbent firms are in, so prices and quantities cannot prevent entrance. In addition, if firms can export to the world market, prices and quantities will not be effective in deterring entrance either.

[5]One possible implication of this direct anti-competitive motive may be the high wage levels in the financial sector, especially among large investment banks. High wage standards may make entrance quite costly, thus reaffirming the concentrated industry structure of the sector.

[6]A classical textbook of development economics says "LDC governments have unwisely invested too much in higher

direct capturing by large businesses causes such a common bias. By our analysis, this overspending on higher education can be a result of a purely benevolent government induced by large businesses to favor their sector. Note that in real life large businesses may not act purposefully to influence the government through setting high wages, but this does not negate the effect identified in the model. Moreover, the wage equalization effect is self-enforcing in the sense that as long as some incumbent large firms set high wages, other incumbents will see their effective labor costs rise and will follow suit. Thus, no explicit coordination among large incumber firms is needed.

The main prediction of our model is that the size-wage effect, the relative size of the primitive sector, and the degree of bias on government spending (which can be proxied by, for example, the ratio of higher versus elementary education expenditures), are positively correlated. If we consider large incumbent firms' direct competition motive to limit entrance to their own industries, then we expect the size-wage effect is positively correlated with industrial concentration. We conduct an empirical analysis of the last implication, based on a comprehensive survey of Chinese industrial firms. Our empirical finding indicates that the size-wage effect is indeed positively correlated with industrial concentration. This not only provides suggestive evidence in support of our theory, but also implies that in empirical estimation of the size-wage effect, industrial concentration should be included in order to avoid the problem of missing variables.

Our paper contributes to the literature that tries to explain the size-wage effect that cannot be counted for by the conventional explanations. One approach argues that unobservable productivity differences of workers in large and small firms, in particular, as a result of differences in general and specific human capital investments, are the main reason behind the size-wage premium (see, for example, Oi and Idson, 1999, Zabojnik and Bernhardt 2001, and Feng and Zheng 2010). Another kind of explanation is based on labor market frictions such as on-the-job searching (Burdett and Mortensen, 1998) and private information (Feng and Zheng, 2009). Our paper differs from these approaches in that large incumbent firms use high wage rates as an instrument to limit entrance to the modern sector. Our explanation is more relevant to developing countries, and thus is more of an explanation of why the wage premium of large firms is greater in developing countries than in developed countries than of an explanation of the general size-wage effect.

The rest of the paper is organized as follows. The next section sets up the model, and Section 3 presents the equilibrium analysis of the model. Section 4 contains extensions and discussions of the basic model. Section 5 provides an empirical analysis that corroborates our theory, followed by the concluding remarks in Section 6.

---

education." (Todaro, 1994, p.370)

4

2 Model

2.1 Model Setup

We consider a developing economy with one final consumption good that serves as the numeraire. Labor is abundant, so labor supply is inelastic. Workers are identical. These conditions imply that the market wage rate, denoted by w_0, is set at the subsistence level. The economy has two sectors, each consisting of firms using one of the two technologies available to produce the consumption good. Each technology is represented by a Cobb-Douglas production function: for technology $i = 1, 2$:

$$F_i = A_i l_i^\alpha g_i^\beta.$$

where F_i is the output of a representative firm using the technology i, $A_i > 0$ is a productivity parameter, l_i is the number of workers, and g_i is the amount of government resources that are available to each firm using technology i. We suppose $A_1 > A_2$, so technology 1 is more efficient than technology 2. We refer to technology 1 as "modern technology" and firms using technology 1 as "in the modern sector". Technology 2 is called "primitive technology" and firms using technology 2 are called "in the primitive sector." Note that for simplicity, capital is not explicitly shown in the production functions, but is implicitly included in A_i. Thus, we should have $\alpha + \beta < 1$.

At a wage rate w, a profit-maximizing firm in sector i will have the following optimal employment level (l_i) and the corresponding revenue (R_i) and profit (Π_i):

$$l_i = (\frac{A_i \alpha}{w})^{\frac{1}{1-\alpha}} (g_i)^{\frac{\beta}{1-\alpha}}; \quad R_i = A_i^{\frac{1}{1-\alpha}} (\frac{\alpha}{w})^{\frac{\alpha}{1-\alpha}} (g_i)^{\frac{\beta}{1-\alpha}}; \quad \Pi_i = (1-\alpha) A_i^{\frac{1}{1-\alpha}} (\frac{\alpha}{w})^{\frac{\alpha}{1-\alpha}} (g_i)^{\frac{\beta}{1-\alpha}}. \quad (1)$$

When $g_1 = g_2$, because $A_1 > A_2$, we have $l_1 > l_2$, $R_1 > R_2$ and $\Pi_1 > \Pi_2$ for any w.

In the economy, there are a small number of large entrepreneurs and a large number of small entrepreneurs. We normalize the measure of large entrepreneurs in the economy to one and let the measure of small entrepreneurs be M with $M >> 1$. We assume that one entrepreneur only sets up one firm, because each entrepreneur has just enough attention or capital to run one business. Since the modern technology is more efficient, each large entrepreneur sets up a modern firm and becomes the incumbent of the modern sector. With an entrance cost k, a small entrepreneur can also enter the modern sector and use the modern technology (see, Lucas, 1978). We suppose that the entrance cost k is uniformly distributed on $[0, K]$ with the density $m = M/K$. We call a small entrepreneur (firm) with k entrance cost a k-type entrepreneur (firm). We can interpret the entry cost as the cost of accessing or learning to use the modern technology. This cost varies across small entrepreneurs due to their different human capital or other resources that are complementary to the modern technology.

We denote the numbers of small entrepreneurs who choose to enter the modern and the primitive sectors as y_1 and $y_2 = M - y_1$ respectively. If a k-type entrepreneur finds that it is more profitable to enter the modern sector, this entrepreneur and the small entrepreneurs with lower entry costs will all enter the modern sector. So there exists a cut-off entry cost $\bar{k}$ such that small entrepreneurs enter the modern sector if and only if their entry costs are less than $\bar{k}$. Therefore we have $y_1 = m\bar{k}$ and $y_2 = M - y_1 = m(K - \bar{k})$.

In developing countries, government investments are usually critical for economic growth, but government resources are often quite limited. Allocating more resources to any of these two sectors will improve the efficiency of the firms in that sector. Thus, governments need to decide how to efficiently allocate these scarce resources between the modern and the primitive sectors. In our model, we suppose that the total amount of resources the government has is G, out of which G_1 is allocated to the modern sector and G_2 to the primitive sector, where $G_1 + G_2 = G$.[7] Then the amount of government resources a representative modern firm gets is $g_1 = G_1/(1 + y_1)$, and similarly $g_2 = G_2/y_2$. Note that in our specification of production functions, the average government resources in the sector i, g_i, affects the efficiency of firms in sector i. This assumes that there is strong congestion effect, or "exclusivity," about government resources: firms' production efficiencies are increasing in the government resources allocated to their sector, but decreasing in the number of the firms in their sector. Many government expenditures have the features of public goods, but few have the feature of pure non-exclusivity. As long as there is a sufficiently strong degree of exclusivity, our qualitative results still hold.

In the basic model, we consider a benevolent government who maximizes the total outputs of the economy. In Section 4 we briefly discuss what happens when the government is not benevolent. Specifically, the government in the basic model maximizes the following function:

$$U = (1 + y_1)R_1 + y_2 R_2 - y_1^2/(2m). \tag{2}$$

where $y_1^2/(2m)$ is the total entrance costs incurred by the small entrepreneurs entering the modern sector with entrance costs between $[0, \bar{k} = y_1/m]$, and R_1 and R_2 are the outputs of representative firms in the modern and the primitive sectors respectively.

We consider the following game. At date 1, the incumbent modern firm sets a wage standard w_1 for the modern sector. At date 2, small entrepreneurs choose to enter one of the two sectors, and at

---

[7]Here we make the simplifying assumption that the government's total resource available G is exogenously fixed. It is easy to generalize the model to the case in which G is endogenously determined by the total tax revenue of the economy. The analysis would be much more cumbersome, without gaining much additional insight, thus we will not pursue the extension.

the same time, the government makes a resource allocation decision.[8]

By our assumption of the wage equalization effect, once the incumbent modern firms set a wage standard w_1, all new entrants to the modern sector have to pay the same wage rate. This is of course a strong assumption. What we really need is that if the incumbent firms in the modern sector pay a wage rate of w_1, firms that pay their workers less than w_1 suffer productivity losses and their effective labor costs increase. In Section 4, we discuss what happens if we relax the wage equalization effect. We assume the wage-equalization effect only applies to the modern sector but not the primitive sector. One expects that worker turnovers tend to happen among similar firms and people make fairness comparisons with others in similar situations. Moreover, firms with different production technologies and organizational structures may respond differently to high outside wage rates. For example, firms in the primitive sector may not be affected by higher wages in the modern sector because they can monitor their workers much more efficiently due to their small sizes and simple organizational structures.

The wage equalization effect is self-enforcing: any incumbent modern firm will not find it profitable to deviate from the optimal wage standard w_1. Therefore, it is not necessary for the incumbent firms to explicitly coordinate their wage levels. But explicit coordinations, such as those through industry associations, certainly help enforce an equalized market wage standard. Other mechanisms such as reputations, unions, and minimum wage regulations may also help enforce the market wage levels (Williamson, 1968).

With the wage rate w_0 in the primitive sector and w_1 in the modern sector, the outputs of the representative firms in the two sectors, R_1 and R_2, in Equation (1), are given by

$$R_1 = A_1^{\frac{1}{1-\alpha}}(\frac{\alpha}{w_1})^{\frac{\alpha}{1-\alpha}}(\frac{G_1}{1+y_1})^{\frac{\beta}{1-\alpha}}; R_2 = A_2^{\frac{1}{1-\alpha}}(\frac{\alpha}{w_0})^{\frac{\alpha}{1-\alpha}}(\frac{G_2}{y_2})^{\frac{\beta}{1-\alpha}}. \tag{3}$$

2.2 The First Best Solution

The first best solution of the model is a triplet (w_1^*, y_1^*, G_1^*) that yields the greatest total output. Since the total output in the modern sector is decreasing in w_1, so in the first best solution it must be that $w_1^* = w_0$. Then (y_1^*, G_1^*) solves the following problem:

$$\max_{y_1,G_1} \quad U = (1+y_1)A_1^{\frac{1}{1-\alpha}}(\frac{\alpha}{w_0})^{\frac{\alpha}{1-\alpha}}(\frac{G_1}{1+y_1})^{\frac{\beta}{1-\alpha}} + y_2 A_2^{\frac{1}{1-\alpha}}(\frac{\alpha}{w_0})^{\frac{\alpha}{1-\alpha}}(\frac{G_2}{y_2})^{\frac{\beta}{1-\alpha}} - \frac{y_1^2}{2m},$$

where $G_2 = G - G_1$ and $y_2 = M - y_1$. Assuming interior solution, we obtain the first-order-condition

---

[8]The critical assumption of the timing of the model is that large entrepreneurs move before small entrepreneurs and the government. Changing the order of the moves by small entrepreneurs and the government does not alter the model results.

with respect to y_1:

$$(1 - \frac{\beta}{1-\alpha})A_1^{\frac{1}{1-\alpha}}(\frac{\alpha}{w_0})^{\frac{\alpha}{1-\alpha}}(\frac{G_1}{1+y_1})^{\frac{\beta}{1-\alpha}} - \frac{y_1}{m} = (1 - \frac{\beta}{1-\alpha})A_2^{\frac{1}{1-\alpha}}(\frac{\alpha}{w_0})^{\frac{\alpha}{1-\alpha}}(\frac{G_2}{y_2})^{\frac{\beta}{1-\alpha}}. \qquad (4)$$

The left hand side of the above equation is the marginal product of y_1, which equals the marginal output in the modern sector (net of the marginal congestion effect measured by $\beta/(1-\alpha)$) minus the marginal entrance cost (y_1/m). The right hand side is the marginal cost of y_1 measured by the loss of marginal revenue to the primitive sector.

The first-order condition with respect to G_1 is

$$A_1^{\frac{1}{1-\alpha}}(\frac{\alpha}{w_0})^{\frac{\alpha}{1-\alpha}}[\frac{G_1}{1+y_1}]^{\frac{\beta}{1-\alpha}-1} = A_2^{\frac{1}{1-\alpha}}(\frac{\alpha}{w_0})^{\frac{\alpha}{1-\alpha}}[\frac{G_2}{y_2}]^{\frac{\beta}{1-\alpha}-1}. \qquad (5)$$

This can be simplified as $A_1 g_1^{\alpha+\beta-1} = A_2 g_2^{\alpha+\beta-1}$, or $g_1 = g_2(A_1/A_2)^{1/(1-\alpha-\beta)}$. Thus the government will optimally allocate more resource per firm in the modern sector than in the primitive sector, because the marginal productivity of government resource is higher in the modern sector.

Since the objective function U is concave in (y_1, G_1), the second order conditions are satisfied. Therefore, if the solution of the two first-order conditions exists, that defines a pair of optimal choices: the optimal size of the modern sector $(1+y_1^*)$ and the optimal government resource allocation (G_1^*). We have the following assumption to ensure the existence of the solution.

Assumption (A1) $K > \frac{1-\alpha-\beta}{1-\alpha}[A_1^{\frac{1}{1-\alpha}} - A_2^{\frac{1}{1-\alpha-\beta}}A_1^{\frac{-\beta}{(1-\alpha-\beta)(1-\alpha)}}][\frac{\alpha}{w_0}]^{\frac{\alpha}{1-\alpha}}[\frac{G}{1+M}]^{\frac{\beta}{1-\alpha}}.$

Proposition 1 below characterizes the first best solution (all technical proofs are in the Appendix).

Proposition 1 *In the first best solution, the wage rate in the modern sector is $w_1^* = w_0$. Under Assumption (A1), the optimal size of the modern sector $(1+y_1^*)$ and the optimal government resource allocation (G_1^*) are characterized by the unique solution to Equations (4) and (5).*

Assumption (A1) ensures that not all small entrepreneurs should enter the modern sector in the first best solution, even though the modern technology is more efficient than the primitive one. This is because: (i) the complementary production input, government resources, are limited, too many entrants would reduce the efficiency level of all modern firms; and (ii) entry to the modern sector is costly. The condition of Assumption (A1) means that (i) the entrance cost is substantial; or (ii) the efficiency gap of the two technologies is not too large; or (iii) government resources are scarce and important. These tend to be true in many developing economies. In this paper we only consider a static model. Dynamically, as the business conditions of the economy improve over time, one expects that the optimal size of the modern sector becomes larger over time.

8

3   Equilibrium Analysis of the Model

Now we solve the equilibrium of the model by backward induction, and denote the equilibrium outcome as $(\bar{w}_1, \bar{y}_1, \bar{G}_1)$.

At date 2, after observing w_1, the government chooses G_1 and $G_2 = G - G_1$ to maximize U as given by Equation (2). The first-order condition is given by

$$A_1^{\frac{1}{1-\alpha}}(\frac{\alpha}{w_1})^{\frac{\alpha}{1-\alpha}}[\frac{G_1}{1+y_1}]^{\frac{\beta}{1-\alpha}-1} = A_2^{\frac{1}{1-\alpha}}(\frac{\alpha}{w_0})^{\frac{\alpha}{1-\alpha}}[\frac{G_2}{y_2}]^{\frac{\beta}{1-\alpha}-1}. \tag{6}$$

This equation is analogous to Equation (5) and the only difference between the two equations is that the wage rate of the modern sector on the left-hand side is now w_1 instead of w_0. Clearly, given y_1, G_1 decreases in w_1 as a higher wage rate in the modern sector reduces the efficiency of this sector and hence reduces the marginal value of G_1. However, y_1 is not fixed and is in fact affected by w_1. The wage rate in the modern sector affects the number of small entrepreneurs entering this sector.

At date 2, after observing w_1, small entrepreneurs also make their sector choices. Since the number of small entrepreneurs is large, we assume that each small entrepreneur ignores the effect of his own sector choice on the government policy G_1. Small entrepreneurs will enter the modern sector if their entry costs are below a threshold level k_1. A small entrepreneur must be indifferent between the two sectors if his entry cost is exactly at k_1. Since $y_1 = mk_1$, this indifference condition is:

$$\Pi_1 - \Pi_2 = (1-\alpha)A_1^{\frac{1}{1-\alpha}}(\frac{\alpha}{w_1})^{\frac{\alpha}{1-\alpha}}(\frac{G_1}{1+y_1})^{\frac{\beta}{1-\alpha}} - (1-\alpha)A_2^{\frac{1}{1-\alpha}}(\frac{\alpha}{w_0})^{\frac{\alpha}{1-\alpha}}(\frac{G_2}{y_2})^{\frac{\beta}{1-\alpha}} = \frac{y_1}{m}. \tag{7}$$

In this equation, given G_1, y_1 decreases in w_1 because a higher wage rate reduces the profitability of the modern sector and makes it less attractive to the small entrepreneurs. Equation (7) is analogous to Equation (4). Rewriting Equation (4) yields $\Pi_1 - \Pi_2 = \frac{(1-\alpha)^2}{1-\alpha-\beta}\frac{y_1}{m}$. We can think of this first-best condition as the case that small entrepreneurs and the government simultaneously make decisions but small entrepreneurs' marginal cost of entering the modern sector is $\frac{(1-\alpha)^2}{1-\alpha-\beta}\frac{y_1}{m}$ instead of $\frac{y_1}{m}$. This difference arises from two factors: (i) small entrepreneurs maximize their profits while the social planner maximizes social outputs; and (ii) the small entrepreneurs do not take into account the externalities of their decisions on others.

To focus on the more interesting case, we make the following assumption:

Assumption (A2) $\beta < \alpha - \alpha^2$.

Assumption (A2) implies that the marginal cost to enter the modern sector is smaller in the first best solution than in the equilibrium of the model. Otherwise we might obtain the uninteresting outcome that equilibrium entry to the modern sector is more than the efficient level. Assumption (A2)

says that relative to the labor share in the production function, the share of government resources cannot be too large. Considering the fact that labor and other factors (such as capital and land) not explicitly included in our analysis are direct production inputs, this assumption is quite plausible.

If there is a solution to Equations (6) and (7), then we obtain an equilibrium of the second stage subgame. Let $(G_1(w_1), y_1(w_1))$ denote the stage equilibrium.

Lemma 1 *Under Assumptions (A1) and (A2), there is an interior equilibrium of the second stage subgame $(G_1(w_1), y_1(w_1))$. Furthermore, both G_1 and y_1 decrease in w_1.*

Because of the similarities between Equations (4) and (7) and between Equations (5) and (6), the argument of Lemma 1 parallels that of Proposition 1.

Lemma 1 can be illustrated by Figure 1 below.[9] The vertical axis is the total government resource allocated to the modern sector, the horizontal axis is the number of entrants to the modern sector. The curve $Z(y_1; w_0)$ represents the government's optimal decision rule regarding resource allocation (from Equation 5), and the curve $W(y_1; w_0)$ represents the socially optimal entry condition to the modern sector (from Equation 4), both at the wage rate w_0. The interception of $Z(y_1; w_0)$ and $W(y_1; w_0)$, labeled as "A", is the first best solution $(G_1^*(w_0), y_1^*(w_0))$. If the modern sector wage rate is kept at w_0 but small entrepreneurs make entry decisions, the entry condition (Equation 7) can be represented by $V(y_1; w_0)$, which lies in the upper left region of $W(y_1; w_0)$ in Figure 1. This is because for the same G_1, fewer small entrepreneurs will enter the modern sector than in the first best as the profit differential, instead of the revenue differential, between the two sectors needs to compensate for the entry cost. If the incumbents in the modern sector set the wage standard at $w_1 = w_0$, then the equilibrium is the interception of $V(y_1; w_0)$ and $Z(y_1; w_0)$, labeled as "B" in Figure 1.

Now consider the case when $w_1 > w_0$. Compared with the case when $w_1 = w_0$, there will be fewer entrants to the modern sector for a fixed amount of total resources allocated to the modern sector, so $V(y_1; w_1)$ lies in the upper left region of $V(y_1; w_0)$ and tilts further towards left as w_1 increases. The government will allocate less resources for a fixed size of the modern sector since firms in the modern sector are less efficient, which is illustrated by $Z(y_1; w_1)$ being below $Z(y_1; w_0)$ and becoming flatter for larger w_1. Clearly, when the wage rate in the modern sector is set at w_1, the interception of $V(y_1; w_1)$ and $Z(y_1; w_1)$, labeled as "C" at $(\bar{G}_1(w_1), \bar{y}_1(w_1))$, is the equilibrium of the second stage subgame, and both $\bar{G}_1(w_1)$ and $\bar{y}_1(w_1)$ decrease in w_1.

In the first stage of the game, incumbent firms in the modern sector choose a wage standard w_1, taking into account the effects on small entrepreneurs' sector choices and the government's resource

---

[9]All the curves in Figure 1 are not supposed to be linear. We use straight lines only for simplicity.

10

Figure 1: Characterization of Equilibrium in the Second Stage
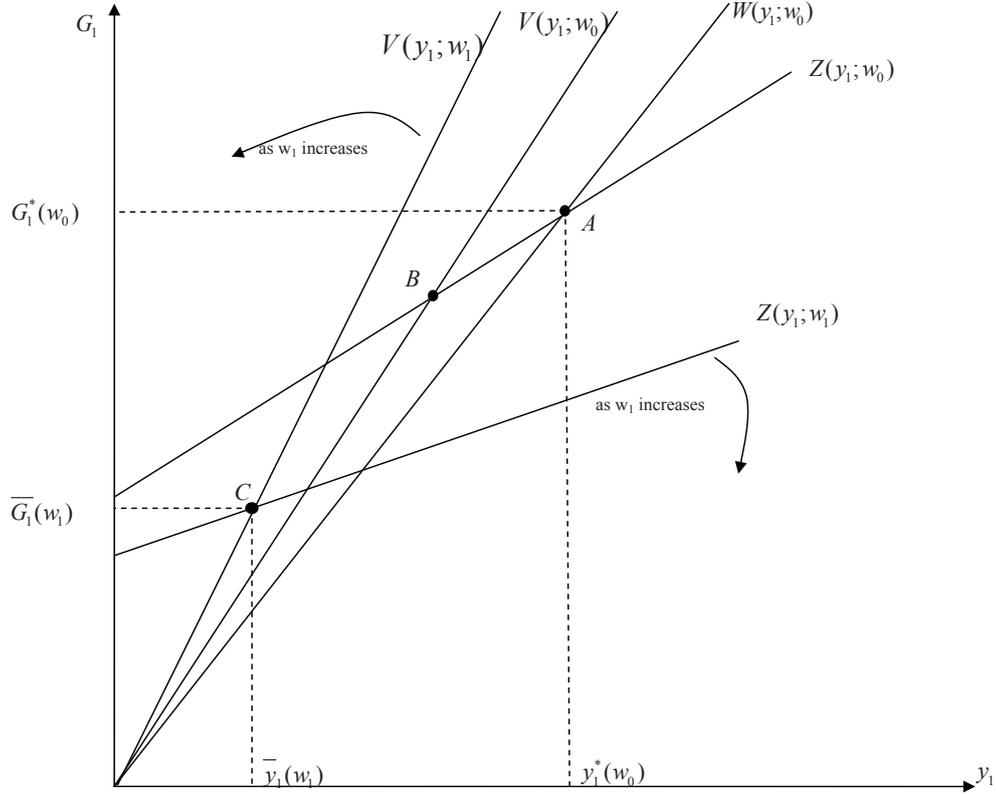
allocation. So the problem for large entrepreneurs is:

$$\max_{w_1} \quad \Pi_1 = (1-\alpha)A_1^{\frac{1}{1-\alpha}}\left(\frac{\alpha}{w_1}\right)^{\frac{\alpha}{1-\alpha}}\left[\frac{G_1(w_1)}{1+y_1(w_1)}\right]^{\frac{\beta}{1-\alpha}}.$$

Assumption (A3) (i) $\frac{A_2}{A_1} > \left(\frac{1-\alpha-\beta}{1-\alpha+\beta}\right)^{1-\alpha-\beta}$; (ii) G is relatively large (the precise condition is given in the Appendix).

Proposition 2 *Under Assumptions (A1)-(A3), in equilibrium the incumbent large entrepreneurs in the modern sector will set a wage standard higher than the market wage rate: $\bar{w}_1 > w_0$.*

Proposition 2 is our main result. It shows that under reasonable conditions, the incumbent large entrepreneurs in the modern sector will indeed use high wage standards to limit entrance to the modern sector by small entrepreneurs. As is clear from the optimization problem of the large incumbent firms, the tradeoff they face is that a higher wage standard directly reduces profit but may induce the government to allocate more valuable resources per firm to the modern sector. Assumption (A3) guarantees that at $w_1 = w_0$, the latter effect is not only possible but significant enough to dominate the former, hence the optimal wage standard is above w_0. The first part of

11

Assumption (A3) guarantees that the technological gap between two sectors are not too large. If the primitive technology is too inefficient, the government will not allocate many resources to the primitive sector. Hence the large incumbent firms in the modern sector will see no need to use high wage standards to limit entrance to their sector. The second part of Assumption (A3) ensures that government resources are sufficiently important so that the large incumbent firms are strongly motivated to limit entrance to the modern sector. The proof of Proposition 2 establishes that under these assumptions the incumbents in the modern sector obtain greater profits at point C (wage rate $w_1 > w_0$) than at point B (wage rate $w_1 = w_0$) in Figure 1.

Proposition 2 shows that one possible reason for the size wage effect in developing countries is that incumbent large entrepreneurs willingly use high wages to prevent entrance to the modern sector by small entrepreneurs. Besides providing a new explanation for the size wage effect, other implications of the equilibrium characterized in Proposition 2 are presented in the following corollaries.

Corollary 1 *The number of small entrepreneurs entering the modern sector is smaller than the efficient level ($\bar{y}_1 < y_1^*$).*

Corollary 1 shows that the primitive sector will be suboptimally too large because in equilibrium fewer small entrepreneurs will enter the modern sector than in the first best solution.[10] This result corresponds nicely with a common phenomenon of developing countries that despite the availability of modern technology and business practices, a large proportion of population is still stuck with primitive technology and are slow to "become modern." Corollary 1 can be easily seen from comparing the equilibrium point C and the first best point A in Figure 1. Intuitively, the incumbent large firms in the modern sector set a high wage standard in order to limit entrance to the modern sector. If they cannot successfully limit entrance, they will simply pay the market wage rate. When they choose to set the wage standard above the market wage rate, the size of the primitive sector has to be larger as some small entrepreneurs are deterred by the high wages in the modern sector.

Corollary 2 *The average resources allocated to each firm the modern sector are greater than the efficient level ($\bar{g}_1 > g_1^*$).*

From Figure 1, the slope of $V(y_1; w_1)$ is steeper than that of $W(y_1; w_0)$, so $\bar{G}_1/\bar{y}_1$ is greater than

---

[10]it is true both in terms of the number of firms and in terms of employment in the modern sector. With higher wage, each firm in the modern sector will hire a smaller number of workers. On the other hand, higher government resources per firm make modern firms more efficient and more willing to hire workers. It can be shown that the former effect dominates the latter and the number of workers a modern firm hires in equilibrium is less than in the first best solution. Combined with the fact that there are fewer number of modern firms in equilibrium, the total employment in the modern sector is lower than the efficient level.

G_1^*/y_1^*. By the definition of $g_1 = G_1/(1+y_1)$, the above observation does not exactly imply $\bar{g}_1 > g_1^*$. The complete proof is in the Appendix.

Corollary 2 shows that the purely benevolent government is induced to allocate resources in favor of incumbent firms in the modern sector. By setting a high wage standard, incumbent large firms can prevent some small entrepreneurs who are potential entrants from entering the modern sector. Thus, the modern sector will be smaller and the primitive sector will be larger than the first best (and than without the high wage standard, i.e., point B in Figure 1). With this change of relative sector sizes, the benevolent government will swift resources from the modern sector to the primitive sector. But such a swift is disproportional, making the average resource per firm in the modern sector greater than the first best (and than point B without wage manipulation). This benefits incumbent large firms, and is the whole purpose of manipulating the wage standard in the modern sector.

4 Extensions and Discussions

4.1 Relative Wage Equalization

In our basic model, we assume that all entrant firms in the modern sector must follow the incumbent firms to set the same wage rate by the wage equalization effect. This assumption can be easily relaxed. As long as high wage rates of the incumbent firms put enough pressure on other firms in the modern sector to raise their wage rates or cause significant increases in the effective labor costs of other firms in the modern sector (e.g., through job turnovers, searching, and shirking), our main results still hold.

Specifically, let w_1 still be the wage rate of the incumbent firms in the modern sector and consider the following general wage equalization effect:

$$w_1^{'} = w_0 + \gamma(w_1 - w_0),$$

where $w_1^{'}$ is the wage rate of entrance firms in the modern sector, and $\gamma \in [0, 1]$ represents how strong the wage equalization effect is in the modern sector. When $\gamma = 1$, $w_1^{'} = w_1$, which is the wage equalization effect assumed in our basic model. When $\gamma < 1$, $w_1^{'} < w_1$. However, by continuity, as long as γ is sufficiently close to 1, we should still have $w_1 > w_0$ in equilibrium, that is, the large incumbent firms in the modern sector will still set a wage standard greater than the market wage rate.

4.2 Non-benevolent Government

In our model, the government has been assumed to maximize the total outputs of the economy. We have shown that even under this assumption, government policies will tend to be biased in favor of incumbent big businesses. More realistically, when the government is not benevolent, one would expect the allocation outcome will be even more biased.

One way the government may not be benevolent is that it is captured by large businesses and thus favors the modern sector intrinsically in its resource allocation. To model such a bias, the government's objective function can be specified as

$$U = (1 + \eta)(1 + y_1)R_1 + y_2 R_2 - y_1^2/(2m),$$

where $\eta > 0$ represents the degree of the government's bias. Compared with the benevolent government, now the biased government will allocate more resources to the modern sector for a given y_1. As a result, more small entrepreneurs will attempt to enter the modern sector, holding the wage rate in the sector w_1 fixed. The incumbent large entrepreneurs will then raise further the wage rate in the modern sector to limit entry. Therefore, both the wage distortion and the distortion in governmental resource allocation will be exacerbated by the government's bias, but obviously our main points of the basic model are intact.

Another way the government may deviate from maximizing the total outputs of the economy is that it maximizes its own tax revenue. If the effective tax rate for all the firms in the economy is the same, there is no additional distortion, because this amounts to maximizing the total outputs.[11] However, a more realistic scenario may be that the effective tax rate for large firms is higher than that for small firms due to the government's limited tax collecting ability in the primitive sector. To capture this scenario, we specify the government's utlility function as

$$U = \tau_1(1 + y_1)\Pi_1 + \tau_2 y_2 \Pi_2,$$

where the effective tax rates for the modern sector and the primitive sectors are τ_1 and τ_2 respectively, with $\tau_1 > \tau_2$. Note that $\Pi_i = (1 - \alpha)R_i$, thus the government's objective function is effectively the same as in the preceding case when it is captured by large businesses. Then for a given y_1, the government tends to allocate more resources to the modern sector than the benevolent government. However, unlike the preceding case, it is less clear whether small entrepreneurs will find the modern sector more attractive because of the tax rate differential between the two sectors. So the incumbent large firms in the modern sector may be more or less aggressive in setting a wage standard than in

---

[11]Note that in our model, each firm's profit Π_i is $(1 - \alpha)$ of its revenue R_i, thus it does not matter whether it is sales tax or corporate income tax.

the basic model. However, unless the tax rate differential between the two sectors is too extreme, the incumbent large firms will still set a wage standard higher than the market wage rate.

In summary, when the government is not a social welfare maximizer, our main points will still hold qualitatively. Note that while government biases may explain distortions of governmental resource allocation, they cannot explain wage distortion between the modern and primitive sectors.

4.3   Regulatory Entry Barriers to the Modern Sector

Many developing economies have minimum wage regulations, which by and large are enforceable only in the modern sector. Suppose the government imposes a minimal wage $w_{min} > w_0$ in the modern sector. If $w_{min} \leq w_1$, nothing changes in the model. If $w_{min} > w_1$, then the incumbent large firms must set the wage standard at w_{min}. Based on the assumption that minimum wage regulation can only be enforced on firms beyond a certain size, Rauch (1991) shows that large firms pay a higher wage rate than small firms for homogeneous workers. However, such a theory does not explain why large firms pay wage rates higher than the legal minimum wage, and is not supported by empirical findings (e.g., Schaffner, 1998).

Governments in developing countries often impose entrance costs to the modern sector, e.g., high registration fees and cumbersome regulatory requirements. Analytically this is equivalent to shifting the entrance cost distribution from $[0, K]$ to $[\mu, K + \mu]$ in our basic model, where μ is the additional entrance cost imposed by the government. Our analysis of the basic model is largely unchanged except that the right hand side of Equation (7) needs to add μ. This additional entrance cost makes small entrepreneurs less interested in entering the modern sector. This will likely make the incumbent large firms in the modern sector less aggressive in setting the wage standard (i.e., w_1 will be lower). Thus, regulatory barriers to the modern sector cannot explain the wage premium of the modern sector, instead, they are more likely to reduce it (see also Velenchik, 1997).

5   Empirical Evidence

We have shown that higher wages can be used by large firms as a means of deterring potential entry of smaller firms. Without matched employer-employee data, we are not able to fully test this theory. However, a testable implication of our theory is that industry concentration should be positively correlated with the size-wage effect. In this section, we empirically test this correlation in order to provide some suggestive evidence for our theory.

Our data set is based on the annual surveys of manufacturing and mining firms conducted by the National Bureau of Statistics of China from 1998 to 2007. These annual surveys cover enterprises

with above five millions RMB annual sales. Number of the enterprises covered in the surveys varies from 154 thousands to 335 thousands over the sample period. A wide variety of information of these enterprises is recorded: identification, ownership, operational performances, accounting and finance, etc. The data set is usually considered of good quality and has been widely used in the studies of Chinese firms.[12] Table 1 reports the means and standard deviations of selected variables used in our analysis by survey years.

Our empirical analysis focuses on the combined effects of firm size and industrial concentration on average wage per employee. We estimate the following regression:

$$\log wage_{ijpt} = \alpha_i + \beta_1 \log asset_{ijpt} + \beta_2 HHI_{jpt}$$
$$+ \beta_3 \log asset_{ijpt} \cdot HHI_{jpt} + \beta_4 \log profit_{ijpt} + \Omega_t + \Phi_j + \Gamma_p + \varepsilon_{ijpt},$$

where the dependent variable is the natural log of average wage per employee paid by enterprise i in industry j, province p and year t; $\log asset_{ijpt}$ is the natural log of each firm's fixed asset, which is our measure of firm size; HHI_{jpt} is the Herfindahl- Hirschman Index, a common measure of industry concentration. Furthermore, we include the interaction term $\log asset_{ijpt} \cdot HHI_{jpt}$ in the regression to study the relationship between industry concentration and the size-wage effect. The interaction term is the focus of our regressional analysis. As more profitable firms are likely to pay higher wages, we include $\log profit_{ijpt}$, the average profit per employee, in the regression to control for this effect. The three variables, Ω_t, Φ_j and Γ_p are the year effect, industry effect and province effect respectively. ε_{ijpt} is the error term.

In our empirical analysis we first regress on the full pooled data, and secondly we only focus on the manufacture sector. For each sample, we estimate two different specifications. In the baseline specification, we drop the interaction term of HHI and log asset, and study the size-wage effect and concentration-wage effect by controlling year, industry and province fixed effects. In the second specification, we include the interaction term into the regression, trying to capture the effect of industry concentration on the size-wage effect.

Table 2 reports the two specifications for the two samples. Our estimations all confirm a very robust size-wage effect: one percentage increase in firm size (measured in total assets) is associated with 0.06% of increase in average wage (excluding the interaction term). Our estimations also show that industrial concentration is positively correlated with wage level (Columns 1 and 3). This is expected as concentrated industries tend to have higher average profit margins and thus may be able to pay higher wages. Notice that although the coefficients of HHI in the second and the fourth specifications are negative, the combined effects of industrial concentration on wage, including both

---

[12]For example, see Bai, Lu and Tao (2009), and Brandt, Van Biesebroeck and Zhang (2009).

HHI and the interaction term, is positive because the mean of log asset is 9.73. Also as expected, firm profitability is positively correlated with wage level, with one percentage of increase in firm average profit is associated with about 0.07 % increase in average wage.

The testable implication of our theory is confirmed by the positive coefficients of the interaction terms in the second and the fourth specifications, which reveals a positive correlation between industrial concentration and the size-wage effect. In other words, in recognition of the size-wage effect, our estimations find that industrial concentration enhances this effect. This is consistent with our theory: the higher wage levels large firms raise to deter potential entry of smaller firms, the more concentrated this industry is. Note that since we control for industrial concentration and firm profitability as well as industry, year, and location effects, this finding cannot be explained by some simple mechanic reasons such as industrial concentration leads to more profitability and thus to higher wages.

Based on annual survey data of Chinese industrial firms, our empirical findings suggest that raising the wage level can be an effective mechanism for large firms to deter potential entry of smaller firms. This implies that in empirical investigation of the size-wage effect, one should control for industrial concentration in cross-sectional analysis without industry fixed effect or in panel data analysis no matter whether industy fixed effect is used, otherwise the estimation may suffer from the problem of missing variables.

6 Conclusion

Using a simple model, we demonstrate that incumbent firms in the modern sector can strategically create entrance barriers through high wage rates and induce the benevolent government to follow policies in favor of their interests. This provides a new explanation for why there is a substantial wage premium of large firms in developing countries with abundant labor supply. One empirical implication is that the size-wage effect, the relative size of the primitive sector, and the degree of bias on government spending (which can be proxied by, for example, the ratio of higher versus elementary education expenditures), are positively correlated. If we consider the large incumbent firms' direct competition motive to limit entrance to their own industries, then we expect the size-wage effect is positively correlated with industrial concentration. This last implication is supported by our empirical analysis based on annual survey data of Chinese industrial firms.

Appendix

Proof of Proposition 1: From Equation (5), we have $g_2 = g_1(A_1/A_2)^{1/(\alpha+\beta-1)}$. Rewriting it gives

$$\frac{G_1}{G-G_1} = \frac{1+y_1}{M-y_1}\left(\frac{A_1}{A_2}\right)^{\frac{1}{1-\alpha-\beta}}. \tag{8}$$

This defines a strictly increasing function of $G_1 = Z(y_1)$ with $Z(0) > 0$ and as $y_1 \to M$, $Z \to G$.

Plugging $g_2 = g_1(A_1/A_2)^{1/(\alpha+\beta-1)}$ into Equation (4) and rearranging terms yields

$$[A_1^{\frac{1}{1-\alpha}} - A_2^{\frac{1}{1-\alpha-\beta}} A_1^{\frac{-\beta}{(1-\alpha-\beta)(1-\alpha)}}][\frac{\alpha}{w_0}]^{\frac{\alpha}{1-\alpha}}(G_1)^{\frac{\beta}{1-\alpha}} = \frac{1-\alpha}{1-\alpha-\beta}\frac{y_1}{m}(1+y_1)^{\frac{\beta}{1-\alpha}}. \tag{9}$$

This also defines a strictly increasing function $G_1 = W(y_1)$ with $W(0) = 0$.

Thus, an interior first best solution exists if Equations (8) and (9) have a solution, or, the two functions, $Z(y_1)$ and $W(y_1)$, intersect at a point $y_1 < M$. Since $Z(0) > W(0)$ and as $y_1 \to M$, $Z \to G$, the solution exists if $W(M) > G$, which is satisfied by Assumption (A1).

Given that $Z(0) > W(0)$ and $Z(M) < W(M)$, the uniqueness is guaranteed if the slope of Z is always less than that of W. From Equation (8), the slope of Z is given by

$$Z'(y_1) = \frac{(G-G_1)^2}{(M-y_1)^2}\left(\frac{A_1}{A_2}\right)^{\frac{1}{1-\alpha-\beta}}\frac{1+M}{G} = g_1 \times g_2 \times \frac{1+M}{G}. \tag{10}$$

Since $g_2 = g_1(A_1/A_2)^{1/(\alpha+\beta-1)} < g_1$, we must have $g_2 < G/(1+M)$, the average government resource (per firm) in the whole economy. Therefore, the slope of Z, given by Equation (10), is less than g_1.

Let us define: $\phi = (A_1^{\frac{1}{1-\alpha}} - A_2^{\frac{1}{1-\alpha-\beta}} A_1^{\frac{-\beta}{(1-\alpha)(1-\alpha-\beta)}})(\frac{\alpha}{w_0})^{\frac{\alpha}{1-\alpha}}$. From Equation (9), the slope of W is

$$
\begin{aligned}
W'(y_1) &= \frac{1-\alpha}{1-\alpha-\beta}\frac{y_1(1+y_1)^{\frac{\beta}{1-\alpha}}/m}{\phi G_1^{\frac{\beta}{1-\alpha}}}\frac{G_1}{1+y_1} + \frac{(1+y_1)^{\frac{\beta}{1-\alpha}}/m}{\phi G_1^{\frac{\beta}{1-\alpha}-1}}\frac{(1-\alpha)^2}{\beta(1-\alpha-\beta)} \\
&> \frac{1-\alpha}{1-\alpha-\beta}\frac{y_1(1+y_1)^{\frac{\beta}{1-\alpha}}/m}{\phi G_1^{\frac{\beta}{1-\alpha}}}\frac{G_1}{1+y_1} = g_1.
\end{aligned}
$$

Hence, $Z'(y_1) < W'(y_1)$. So the uniqueness is established.   Q.E.D.

Proof of Lemma 1: Because of the similarities between Equations (4) and (7) and between Equations (5) and (6), the proof of the lemma follows closely that of Proposition 1.

First, we prove that $w_1 \le w_0(A_1/A_2)^{1/\alpha}$. Suppose $w_1 > w_0(A_1/A_2)^{1/\alpha}$, we have $g_1 < g_2$, and $g_1 < \frac{G}{1+M}$. Therefore, the profit of a modern firm is given by $\Pi_1 = A_1^{\frac{1}{1-\alpha}}(\frac{\alpha}{w_1})^{\frac{\alpha}{1-\alpha}}(g_1)^{\frac{\beta}{1-\alpha}} \le A_2^{\frac{1}{1-\alpha}}(\frac{\alpha}{w_0})^{\frac{\alpha}{1-\alpha}}(\frac{G}{1+M})^{\frac{\beta}{1-\alpha}}$. However, if the modern firm set the wage to be w_0, its profit becomes $\Pi_1 = A_1^{\frac{1}{1-\alpha}}(\frac{\alpha}{w_0})^{\frac{\alpha}{1-\alpha}}(g_1)^{\frac{\beta}{1-\alpha}} \ge A_2^{\frac{1}{1-\alpha}}(\frac{\alpha}{w_0})^{\frac{\alpha}{1-\alpha}}(\frac{G}{1+M})^{\frac{\beta}{1-\alpha}}$. Because the profit is always higher in the latter case, the modern firm will never set the wage rate higher than $w_0(A_1/A_2)^{1/\alpha}$.

Rewriting Equation (6) gives

$$\frac{G_1}{G-G_1} = \frac{1+y_1}{M-y_1}\left(\frac{A_1}{A_2}\right)^{\frac{1}{1-\alpha-\beta}}\left(\frac{w_0}{w_1}\right)^{\frac{\alpha}{1-\alpha-\beta}}. \tag{11}$$

Similar to Equation (8), this defines a strictly increasing function of $G_1 = Z(y_1; w_1)$. For all $w_1 \geq w_0$, it can verified that $Z(0; w_1) > 0$ and as $y_1 \to M$, $Z(y_1; w_1) \to G$. Furthermore, $Z(y_1; w_1)$ is strictly decreasing in w_1.

Define $e = (\frac{A_1}{A_2})^{\frac{1}{1-\alpha-\beta}} (\frac{w_1}{w_0})^{\frac{-\alpha}{1-\alpha-\beta}}$. Since $w_1 \leq w_0 (A_1/A_2)^{1/\alpha}$, $e \geq 1$. From Equation (6), we know $g_1/g_2 = e$, thus $g_1 \geq g_2$.

Similar as before, we can rewrite Equation (7) to obtain

$$[A_1^{\frac{1}{1-\alpha}} - A_2^{\frac{1}{1-\alpha-\beta}} A_1^{\frac{-\beta}{(1-\alpha-\beta)(1-\alpha)}} (\frac{w_1}{w_0})^{\frac{\alpha}{1-\alpha-\beta}}][\frac{\alpha}{w_1}]^{\frac{\alpha}{1-\alpha}} (G_1)^{\frac{\beta}{1-\alpha}} = \frac{1}{1-\alpha} \frac{y_1}{m} (1+y_1)^{\frac{\beta}{1-\alpha}}. \tag{12}$$

Note that the left hand side is positive when $e > 1$. If $e \leq 1$, then no small entrepreneur will enter the modern sector ($y_1 = 0$). Equation (12) defines a strictly increasing function $G_1 = V(y_1; w_1)$ with $V(0; w_1) = 0$. It can be easily seen that for all $w_1 \geq w_0$, $V(y_1; w_1)$ is strictly increasing in w_1. Furthermore, by Assumption (A2), the right hand side of Equation (12) is greater than that of Equation (9). Thus, $V(y_1; w_0) \geq W(y_1)$. Therefore, $V(y_1; w_1) > W(y_1)$ for $w_1 > w_0$.

Since $W(M) > G$ under Assumption (A1), we have $V(M; w_1) > G$. Therefore, for any $w_1 \geq w_0$, there is an interior solution to Equations (7) and (6), because $Z(y_1; w_1)$ and $V(y_1; w_1)$ intersect at least once. It is also easy to show that the slope of $Z(y_1; w_1)$, $\partial Z/\partial y_1$ is strictly decreasing in w_1, and that the slope of $V(y_1; w_1)$, $\partial V/\partial y_1$ is strictly increasing in w_1 and is steeper than the slope of $W(y_1)$. Then from the proof of Proposition 1, it follows that the slope of $V(y_1; w_1)$ is steeper than the slope of $Z(y_1; w_1)$ for all w_1. Thus, the uniqueness is proven. The comparative static result with respect to w_1 follows from the fact that $Z(y_1; w_1)$ is strictly decreasing in w_1 and $V(y_1; w_1)$ is strictly increasing in w_1. Q.E.D.

Proof of Proposition 2: A conventional way of proving the proposition is to use the Implicit Function Theorem to derive $dg_1(w_1)/dw_1$ from Equations (6) and (7). However, because it is impossible to solve explicitly for $G_1(w_1)$ and $y_1(w_1)$, we still cannot directly solve the large entrepreneurs' optimization problem.

Our approach is to use Equations (6) and (7) to express Π_1 directly as a function of w_1. First, by definition,

$$\Pi_1 = (1-\alpha) A_1^{\frac{1}{1-\alpha}} (\frac{\alpha}{w_1})^{\frac{\alpha}{1-\alpha}} (g_1)^{\frac{\beta}{1-\alpha}}; \Pi_2 = (1-\alpha) A_2^{\frac{1}{1-\alpha}} (\frac{\alpha}{w_0})^{\frac{\alpha}{1-\alpha}} (g_2)^{\frac{\beta}{1-\alpha}},$$

we can get

$$g_1 = [\frac{\Pi_1}{(1-\alpha) A_1^{\frac{1}{1-\alpha}} (\frac{\alpha}{w_1})^{\frac{\alpha}{1-\alpha}}}]^{\frac{1-\alpha}{\beta}}; g_2 = [\frac{\Pi_2}{(1-\alpha) A_2^{\frac{1}{1-\alpha}} (\frac{\alpha}{w_0})^{\frac{\alpha}{1-\alpha}}}]^{\frac{1-\alpha}{\beta}}.$$

Note that Equation (6) can be rewritten as $\Pi_1/g_1 = \Pi_2/g_2$. Substituting the expressions of g_1 and g_2 in terms of profits into this equation, we have

$$\Pi_1^{1-\frac{1-\alpha}{\beta}} [(1-\alpha) A_1^{\frac{1}{1-\alpha}} (\frac{\alpha}{w_1})^{\frac{\alpha}{1-\alpha}}]^{\frac{1-\alpha}{\beta}} = \Pi_2^{1-\frac{1-\alpha}{\beta}} [(1-\alpha) A_2^{\frac{1}{1-\alpha}} (\frac{\alpha}{w_1})^{\frac{\alpha}{1-\alpha}}]^{\frac{1-\alpha}{\beta}}.$$

Thus, Π_2 can be expressed as a function of Π_1:

$$\Pi_2 = (\frac{A_1}{A_2})^{\frac{1}{\beta+\alpha-1}} (\frac{w_0}{w_1})^{\frac{\alpha}{\beta+\alpha-1}} \Pi_1. \tag{13}$$

Then both g_1 and g_2 can be expressed in terms of Π_1:

$$g_1 = \Pi_1^{\frac{1-\alpha}{\beta}}(1-a)^{-\frac{1-\alpha}{\beta}}A_1^{-\frac{1}{\beta}}(\frac{\alpha}{w_1})^{-\frac{\alpha}{\beta}}; \tag{14}$$

$$g_2 = \Pi_1^{\frac{1-\alpha}{\beta}}(\frac{A_1}{A_2})^{\frac{1-\alpha}{\beta(\beta+\alpha-1)}}(\frac{w_0}{w_1})^{\frac{\alpha(1-\alpha)}{\beta(\beta+\alpha-1)}}(1-a)^{-\frac{1-\alpha}{\beta}}A_2^{-\frac{1}{\beta}}(\frac{\alpha}{w_0})^{-\frac{\alpha}{\beta}}. \tag{15}$$

From the definitions $g_1 = G_1/(1+y_1)$ and $g_2 = (G-G_1)/(M-y_1)$, we have $y_1 = (G-g_2M-g_1)/(g_1-g_2)$.
Then Equation (7) can be rewritten as:

$$\Pi_1 - \Pi_2 = \frac{1}{m}\frac{G-g_2M-g_1}{g_1-g_2}.$$

Substituting Equation (13) into the above equation yields

$$(\Pi_1-\Pi_2)m(g_1-g_2)+g_2M+g_1 = \Pi_1 m[1-(\frac{A_1}{A_2})^{\frac{1}{\beta+\alpha-1}}(\frac{w_0}{w_1})^{\frac{\alpha}{\beta+\alpha-1}}](g_1-g_2)+g_2M+g_1 = G.$$

Using Equations (14) and (15) and by manipulation, we get

$$\Phi_1 w_1^{\frac{\alpha}{\beta}}(\Pi_1^{\frac{\beta+1-\alpha}{\beta}}m+\Pi_1^{\frac{1-\alpha}{\beta}})+\Phi_2 w_0^{\frac{\alpha}{\beta+\alpha-1}}w_1^{\frac{\alpha(1-\alpha)}{\beta(1-\alpha-\beta)}}[\Pi_1^{\frac{1-\alpha}{\beta}}M-2\Pi_1^{\frac{\beta+1-\alpha}{\beta}}m]+\Phi_3\Pi_1^{\frac{\beta+1-\alpha}{\beta}}mw_0^{\frac{2\alpha}{\beta+\alpha-1}}w_1^{\frac{\alpha(\beta+1-\alpha)}{\beta(1-\alpha-\beta)}} = G, \tag{16}$$

where

$$\Phi_1 = (1-a)^{-\frac{1-\alpha}{\beta}}\alpha^{-\frac{\alpha}{\beta}}A_1^{-\frac{1}{\beta}},$$

$$\Phi_2 = (1-a)^{-\frac{1-\alpha}{\beta}}\alpha^{-\frac{\alpha}{\beta}}A_1^{-\frac{1}{\beta}}(\frac{A_2}{A_1})^{\frac{1}{1-\alpha-\beta}} = \Phi_1(\frac{A_2}{A_1})^{\frac{1}{1-\alpha-\beta}},$$

$$\Phi_3 = (1-a)^{-\frac{1-\alpha}{\beta}}\alpha^{-\frac{\alpha}{\beta}}A_1^{-\frac{1}{\beta}}(\frac{A_2}{A_1})^{\frac{2}{1-\alpha-\beta}} = \Phi_1(\frac{A_2}{A_1})^{\frac{2}{1-\alpha-\beta}}.$$

Equation (16) expresses Π_1 implicitly as a function of w_1. Total differentiating both sides of Equation (16), we have

$$\{\frac{\alpha}{\beta}\Phi_1(\Pi_1^{\frac{\beta+1-\alpha}{\beta}}m+\Pi_1^{\frac{1-\alpha}{\beta}})w_1^{\frac{\alpha-\beta}{\beta}}$$

$$+\frac{\alpha(1-\alpha)}{\beta(1-\alpha-\beta)}\Phi_2 w_0^{\frac{\alpha}{\beta+\alpha-1}}w_1^{\frac{\alpha(1-\alpha)}{\beta(1-\alpha-\beta)}-1}(\Pi_1^{\frac{1-\alpha}{\beta}}M-2\Pi_1^{\frac{\beta+1-\alpha}{\beta}}m)$$

$$+\frac{\alpha(\beta+1-\alpha)}{\beta(1-\alpha-\beta)}\Phi_3 w_0^{\frac{2\alpha}{\beta+\alpha-1}}w_1^{\frac{\alpha(\beta+1-\alpha)}{\beta(1-\alpha-\beta)}-1}\Pi_1^{\frac{\beta+1-\alpha}{\beta}}m\}dw_1$$

$$+\{\Phi_1 w_1^{\frac{\alpha}{\beta}}(\frac{m(\beta+1-\alpha)}{\beta}\Pi_1^{\frac{1-\alpha}{\beta}}+\frac{1-\alpha}{\beta}\Pi_1^{\frac{1-\alpha-\beta}{\beta}})$$

$$+\Phi_2 w_0^{\frac{\alpha}{\beta+\alpha-1}}w_1^{\frac{\alpha(1-\alpha)}{\beta(1-\alpha-\beta)}}(\frac{(1-\alpha)M}{\beta}\Pi_1^{\frac{1-\alpha-\beta}{\beta}}-\frac{2m(\beta+1-\alpha)}{\beta}\Pi_1^{\frac{1-\alpha}{\beta}})$$

$$+\frac{m\Phi_3(\beta+1-\alpha)}{\beta}w_0^{\frac{2\alpha}{\beta+\alpha-1}}w_1^{\frac{\alpha(\beta+1-\alpha)}{\beta(1-\alpha-\beta)}}\Pi_1^{\frac{1-\alpha}{\beta}}\}d\Pi_1 = 0.$$

At $w_1 = w_0$, after some manipulation, the above equation becomes

$$\alpha F_1 dw_1 + w_0 F_2 d\Pi_1 = 0,$$

where

$$F_1 = [\Phi_1 - \frac{2(1-\alpha)}{1-\alpha-\beta}\Phi_2+\frac{\beta+1-\alpha}{1-\alpha-\beta}\Phi_3]m\Pi_1+\Phi_1+\frac{1-\alpha}{1-\alpha-\beta}\Phi_2 M,$$

$$F_2 = (\Phi_1-2\Phi_2+\Phi_3)(\beta+1-\alpha)m+[\Phi_1+\Phi_2(1-\alpha)M]\Pi_1^{-1}.$$

Our goal is to show that under the conditions specified in the proposition, $d\Pi_1/dw_1 > 0$ at $w_1 = w_0$, which implies that the optimal wage choice of the large entrepreneurs must be greater than w_0.

It is easy to see that

$$
\begin{aligned}
F_2 &= (\Phi_1 - 2\Phi_2 + \Phi_3)(\beta + 1 - \alpha)m + [\Phi_1 + \Phi_2(1 - \alpha)M]\Pi_1^{-1} \\
&= \Phi_1[1 - (\frac{A_2}{A_1})^{\frac{1}{1-\alpha-\beta}}]^2(\beta + 1 - \alpha)m + [\Phi_1 + \Phi_2(1 - \alpha)M]\Pi_1^{-1} > 0.
\end{aligned}
$$

Thus, $d\Pi_1/dw_1 > 0$ at $w_1 = w_0$ if and only if $F_1 < 0$. We can rewrite F_1 as

$$
F_1 = \frac{\Phi_1}{1 - \alpha - \beta}(\eta_1 m\Pi_1 + \eta_2),
$$

where

$$
\begin{aligned}
\eta_1 &= 1 - \alpha - \beta - 2(1 - \alpha)(\frac{A_2}{A_1})^{\frac{1}{1-\alpha-\beta}} + (\beta + 1 - \alpha)(\frac{A_2}{A_1})^{\frac{2}{1-\alpha-\beta}}, \\
\eta_2 &= 1 - \alpha - \beta + (1 - \alpha)(\frac{A_2}{A_1})^{\frac{1}{1-\alpha-\beta}}M.
\end{aligned}
$$

Clearly $\eta_2 > 0$. Letting $\mu = (\frac{A_2}{A_1})^{\frac{1}{1-\alpha-\beta}}$, we can rewrite η_1 as

$$
\eta_1 = -(1 - \mu)[(1 - \alpha + \beta)\mu - (1 - \alpha - \beta)].
$$

Thus, when $\frac{1-\alpha-\beta}{1-\alpha+\beta} < \mu$, we have $\eta_1 < 0$.

We now show that under Assumption (A3) part (i), $\eta_1 m\pi_1 + \eta_2 < 0$. This implies $F_1 < 0$, and proves the proposition. At $w_1 = w_0$, Equation (16) becomes

$$
w_0^{\frac{\alpha}{\beta}}\Pi_1^{\frac{1-\alpha}{\beta}}[\Phi_1 + \Phi_2 M + (\Phi_1 - 2\Phi_2 + \Phi_3)m\Pi_1] = G.
$$

When $\Pi_1 > 1$ (which can be trivially satisfied), we have

$$
\begin{aligned}
&w_0^{\frac{\alpha}{\beta}}\Pi_1^{\frac{1-\alpha+\beta}{\beta}}[\Phi_1 + \Phi_2 M + (\Phi_1 - 2\Phi_2 + \Phi_3)m] \\
&> w_0^{\frac{\alpha}{\beta}}\Pi_1^{\frac{1-\alpha}{\beta}}[\Phi_1 + \Phi_2 M + (\Phi_1 - 2\Phi_2 + \Phi_3)\Pi_1 m] = G.
\end{aligned}
$$

So

$$
\Pi_1 > \frac{G^{\frac{\beta}{1-\alpha+\beta}}}{w_0^{\frac{\alpha}{1-\alpha+\beta}}[\Phi_1 + \Phi_2 M + (\Phi_1 - 2\Phi_2 + \Phi_3)m]^{\frac{\beta}{1-\alpha+\beta}}}.
$$

Thus, when

$$
G > w_0^{\frac{\alpha}{\beta}}\Phi_1[1 + \mu M + (1 - \mu)^2 m](-\frac{\eta_2}{m\eta_1})^{\frac{1-\alpha+\beta}{\beta}},
$$

we have $\eta_1 m\Pi_1 + \eta_2 < 0$. The above condition is Assumption (A3) part (ii), which holds when G is sufficiently large. Q.E.D.

Proof of Corollary 2: From Figure 1, it is obvious that $\bar{y}_1 < y_1^*$. From the optimization problem of the large incumbent firms in the modern sector, we must have $\bar{g}_1 > g_1(w_0)$, where $g_1(w_0)$ is the average resource per firm in the modern sector if the wage standard in the modern sector is set at w_0. Now we show $g_1(w_0) > g_1^*$.

At $w_1 = w_0$, Equations (5) and (6) are identical and both can be rewritten as $g_2 = g_1/e$, where $e = (A_1/A_2)^{1/(1-\alpha-\beta)} > 1$. From $G_1 + G_2 = G$, we have

$$g_1(w_0)(1 + y_1(w_0)) + g_1(w_0)(M - y_1(w_0))\frac{1}{e} = g_1(w_0)(1 + M/e) + g_1(w_0)y_1(w_0)(1 - 1/e) = G,$$

and

$$g_1^*(1 + M/e) + g_1^*y_1^*(1 - 1/e) = G.$$

From Figure 1, it is clear that $y_1(w_0) < y_1^*$. Thus, we must have $g_1(w_0) > g_1^*$. Therefore, $\bar{g}_1 > g_1^*$.   Q.E.D.

References

Akerlof, George A. and Janet L. Yellen, 1988, Fairness and Unemployment, American Economic Review papers and proceedings, 78:44-49.

——1990, The Fair Wage-effort Hypothesis and Unemployment, Quarterly Journal of Economics, 105:255-284.

Bai, Chong-En, Jiangyong Lu, and Zhigang Tao, 2009, How Does Privatization Work in China? Journal of Comparative Economics, 37: 453-470.

Brandt, Loren, Johannes Van Biesebroeck, and Yifan Zhang, 2009, "Creative Accounting or Creative Destruction? Firm-Level Productivity Growth in Chinese Manufacturing," NBER Working Paper No. 15152.

Brown, Charles and James Medoff, 1989, The Employer Size-Wage Effect, Journal of Political Economy, 97: 1027-1059.

Burdett, Kenneth and Dale T. Mortensen, 1998, Wage Differentials, Employer Size, and Unemployment, International Economic Review, 39:257-73

Chu, Woosik and Robert T. Masson, 1990, Waging Campain to Retain Monopoly: Employee Compensation as Price Signals to Deter Entry, mimeo, Cornell University.

Fehr, Ernst, Georg Kirchsteiger, and Arno Riedl, 1993, Does Fairness Prevent Market Clearing? An Experimental Investigation, Quarterly Journal of Economics, 108:437-459.

Feng, Shuaizhang and Bingyong Zheng, 2009, Cherry-Picking in Labor Market with Imperfect Information, IZA Discussion Paper No. 4309.

——2010, Imperfect Information, On-the-Job Training, and the Employer Size-Wage Puzzle: Theory and Evidence, IZA Discussion Papers No. 4998.

Little, Ian M.D., Dipak Mazumdar and John M. Page Jr., 1987, Small Manufacturing Enterprises: A Comparative Study of India and Other Economies, Oxford University Press, New York.

Lucas, Robert E., 1978, On the size distribution of business firms, Bell Journal of Economics, 9: 508-523.

Oi, Walter and Todd Idson, 1999, Firm Size and Wages, in O. Ashenfelter and D. Card (Eds.), Handbook of Labor Economics, 3:2165-2215. Amsterdam, North-Holland.

Rabin, Matthew, 1998, Psychology and Economics, Journal of Eocnomic Literature, 36:11-46.

Rauch, James E., 1991, Modelling the informal sector formally, Journal of Development Economics, 35: 33-47.

Salop, Steven C. and David T. Scheffman, 1983, Raising Rival's Costs, American Economic Review, 73:267-271.

Schaffner, Julie A., 1998, Premiums to Employment in Larger Establishments: Evidence from Peru, Journal of Development Economics, 55: 81-113.

Schmidt, Christoph M. and Klaus F. Zimmermann, 1991, Work Characteristics, Firm Size and Wages, Review of Economics and Statistics, 73:705-710.

Strobl, Eric and Robert Thornton, 2002, Do Large Employers Pay More in Developing Countries? The Case of Five African Countries, IZA Discussion Paper No. 660.

Summers, Lawrence H., 1988, Relative Wages, Efficiency Wages, and Keynesian Unemployment, American Economic Review papers and proceedings, 78:383-388.

Todaro, Michael P., 1994, Economic Development, Longman, New York & London.

Troske, Kenneth R., 1999, Evidence on the Employer Size-Wage Premium from Worker-Establishment Matched Data, Review of Economics and Statistics, 81:15-26.

Velenchik, Ann D., 1997, Government Intervention, Efficiency Wages, and the Employer Size Wage Effect in Bimbabwe, Journal of Development Economics, 53:305-338.

Williamson, Oliver, 1968, Wage Rates as a Barrier to Entry: The Pennington Case, Quarterly Journal of Economics, 85:85-116.

Zabojnik, Jan and Dan Bernhardt, 2001, Corporate Tournaments, Human Capital Acquisition, and the Firm Size-Wage Relation, Review of Economic Studies, 68:693-716.

Table 1: Summary Statistics

year	Obs.	stats	Wage per person	Firm Employment	Profit per person	Firm Assets
1998	157,073	mean	7.98	383.54	4.14	67,086.88
		sd	63.33	1,888.68	224.36	589,760.60
1999	154,928	mean	8.20	371.17	4.28	74,921.86
		sd	36.25	1,733.00	76.87	656,420.90
2000	156,306	mean	9.09	351.70	6.70	78,618.92
		sd	60.77	1,536.30	70.86	692,282.10
2001	165,717	mean	10.45	324.27	8.45	78,873.79
		sd	194.91	1,416.63	160.79	693,320.50
2002	177,028	mean	10.93	309.54	11.32	80,260.46
		sd	108.32	1,363.34	231.41	694,805.10
2003	194,580	mean	11.05	294.90	13.01	82,808.57
		sd	12.10	1,272.33	166.43	760,097.20
2004	274,377	mean	12.73	241.12	15.23	70,297.90
		sd	57.88	1,093.99	588.46	1,045,584.00
2005	270,227	mean	14.58	266.77	20.46	80,051.39
		sd	105.30	1,202.28	724.94	1,174,841.00
2006	299,589	mean	15.19	245.38	21.61	84,570.64
		sd	13.63	1,193.03	543.33	1,250,180.00
2007	335,258	mean	17.26	234.63	27.28	88,044.08
		sd	49.34	1,157.09	495.71	1,318,622.00
Total	2,185,083	mean	12.56	288.30	15.30	79,447.65
		sd	81.28	1,350.73	440.28	1,002,959.00

Monetary unit: thousand RMB

Table 2: Industrial Concentration and Average Wage

	All Sectors		Manufacture Sector	
	(1)	(2)	(1)	(2)
log(asset)	0.063***	0.062***	0.063***	0.062***
	[0.001]	[0.001]	[0.001]	[0.001]
HHI	0.059***	-0.643***	0.062**	-0.371**
	[0.027]	[0.016]	[0.029]	[0.181]
HHI*log(asset)		0.071***		0.044**
		[0.020]		[0.018]
log(average profit)	0.070***	0.070***	0.069***	0.069***
	[0.001]	[0.001]	[0.001]	[0.001]
intercept	1.518***	1.527***	1.565***	1.787***
	[0.177]	[0.177]	[0.176]	[0.184]
Observations	1,645,326	1,645,326	1,527,636	1,527,636
$\hat{R2}$	0.194	0.194	0.19	0.19

Dependent variable is the natural log of average wage.

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

All regressions include year FE, industry FE and province FE.

All the values reported in this table have been adjusted to

real values. We use CPI to deflate wages, PPI to deflate profits,

and IPI (Investment Price Index to deflate fixed assets.)