Ann Reg Sci (2004) 38:607–625
DOI: 10.1007/s00168-003-0188-z

The Annals of
Regional Science

Efficiency wages, agglomeration, and a developing dual economy

Bharati Basu

Department of Economics, Central Michigan University, Mt. Pleasant MI 48859, USA
(e-mail: Basu1b@cmich.edu)

Received: February 2003/Accepted: October 2003

Abstract. In analyzing the implications of rural-urban migration in the presence of efficiency wages and external economies of scale in the urban sector, this paper focuses on structural transformation of a developing dual economy. It compares the agglomeration effects in the urban sector under exogenous wage distortion with that under exogenous wage distortion and external economies of scale and also with the agglomeration under efficiency wages and external economies of scale. It shows that because of the employment enhancing effect of rural-urban migration with efficiency wages, the agglomeration economies are bigger with efficiency wages than with minimum wage distortion in the Harris-Todaro model. In exploiting the existing external economies of scale, this agglomeration reduces the sectoral wage differential, and changes the effects of factor accumulation and commodity price changes in a way that is different from the effects under migration with exogenous wage distortion.

JEL classification: O18, R23

1. Introduction

It is a conventional wisdom that the process of development in a dual economy starts with the mobilization of labor resources from the dominant rural sector to a growing urban sector. However, at times this mobilization seems inconsistent with the existence of unemployment in the urban sector. The pioneering work by Harris and Todaro (1970) on this topic has used an exogenous wage distortion such as minimum wage legislation to explain urban unemployment that persists in a developing dual economy. This wage is fixed above the equilibrium level and it is also higher than the existing rural wage. The Harris-Todaro (henceforth referred to as the H-T model) model shows how this wage rigidity and the resulting sectoral wage difference can trigger rural-urban migration in spite of unemployment in the urban sector. A large number of studies (Corden and Findlay 1975 and references there) following H-T have discussed the issues related with this migration. Another

issue that has attracted much attention in the analysis of the development and the gradual structural transformation of a dual economy is the presence of agglomeration economies in the urban sector.

The focus of many of these studies that discuss the H-T model is the problem that rural-urban migration makes the situation worse by raising the pre-migration urban unemployment level. Migration eventually stops when expected wages (actual wage times the probability of employment) in the two sectors become equal and an equilibrium is reached with urban unemployment. In order to reduce the severity of this unemployment resulting from the wage distortion, these studies (Corden and Findlay 1975 and references there) have suggested different policies. Many of these policies are usually either politically or economically difficult to implement and thus this distortion is considered to be a problem for a developing dual economy.

However, wage distortion in a developing economy can very well take an endogenous form. With an underdeveloped information system, perfect monitoring of the workers is not feasible. Employers in the urban sector, therefore, use efficiency wages to prevent shirking in the work place. This efficiency wage is higher than the market-clearing wage. The higher wage and the possibility of being fired if caught shirking, and therefore, becoming unemployed, act as control mechanisms.[1]

The model in this paper uses an efficiency wage that is set above the market-clearing wage in the urban sector and it is also higher than the existing rural wage. It is assumed that there is no shirking in the rural sector. The level of this efficiency wage, and the resulting unemployment level are determined endogenously in the economy. This efficiency wage and the unemployment serve as discipline devices to control shirking. In the case of the exogenous wage distortion, a desired wage level is achieved but the unemployment is its disturbing consequence. In the present paper, both the level of the efficiency wage and the corresponding unemployment are desired, as they serve as substitutes for a perfect monitoring system.

In addition to this endogenous form of wage distortion, the urban sector in this paper also has another feature. Production here enjoys economies of scale. These economies of scale are external in nature. The assumption of economies of scale is very reasonable for a developing urban sector. In the initial stages of development, the urban sector is at its infant stage of production and is ready to exploit the potentials of the existing factors of production. As the growth process thrives, the newly developed infrastructure facilities and technology reduces the transport cost and makes it possible to produce with economies of scale. Goldstein and Moses (1973), in fact, have criticized the assumption of constant returns to scale in the urban sector. According to them, urban growth is intimately related to scale and agglomeration economies.

A vast majority of analyses of rural-urban migration in the presence of urban unemployment does not pay any attention to these economies of scale. Also, the studies on fast urban growth totally ignore the existence of a big rural economy and the presence of unemployment in the urban sector (with the exception of Shukla and Stark 1990). The inclusion of unemployment resulting from efficiency wage distortion, together with the agglomeration

[1] A growing literature supports the existence of efficiency considerations in the modern sector of the developing countries (Agenor 1996; Gatiga et al. 1995).

economies in the analysis of rural-urban migration in this paper thus fills the gap in the literature.

This paper analyses the implications of rural-urban migration in the presence of external economies of scale and endogenously created unemployment in the urban sector. The analysis in this paper suggests that

i) Inter-sectoral labor-mobility actually increases employment in the economy when the urban wage is endogenously determined according to the efficiency wage theory. This stands in sharp contrast with the employment effect of inter-sectoral labor mobility under exogenous wage distortion (Harris and Todaro 1970; Corden and Findlay 1975 and references there).
ii) This migration also reduces sectoral wage differential and the distortion that results from it.
iii) The interaction between endogenous wage distortion and external economies of scale produces a bigger urban growth compared to that with exogenous wage distortion or with exogenous wage distortion and external economies of scale.
iv) The models of internal economies of scale with differentiated products create a wage effect (backward linkage) that helps urban concentration. Although the urban sector in this paper already has a higher wage because of the endogenous distortion, the external economies of scale raise that wage and create an effect that is somewhat similar to the backward linkage effect of internal economies of scale.
v) Growth in the supply of capital will help the manufacturing sector grow, but by less than what it would have been in the absence of efficiency wages. This growth also increases sectoral wage differential.
vi) An increase in the price of manufacturing good will increase urban sector's wage. This is in contrast to the case of exogenous wage distortion and external economies of scale. This increase in the price of manufacturing good also raises the return to capital and lowers rural sector's wage as it does in the case of exogenous distortion and external economies of scale (Panagariya and Succar 1986). But the magnitudes of these changes are expected to be less than what they are in the case of exogenous wage distortion and external economies of scale. It is because a part of the effects of external economies of scale in the urban sector is used by the endogenous rise in urban sector's wage.

In Sect. 2, I present the model. The migration equilibrium is described in Sect. 3. Section 4 compares urban growth or agglomeration under exogenous distortion with that under endogenous distortion. In Sect. 5, I compare urban agglomeration under external economies of scale with that under internal economies of scale. The comparative static analyses with respect to price and factor supply changes are presented in Sect. 6. Section 7 concludes with a brief summary of the results.

2. Model

In this dual economy, the urban sector produces the manufacturing good and the rural sector produces the agricultural good. Two factors of production, capital, K, and labor, L, are completely mobile between the sectors, and both of them enter into the production of each of the commodities. The manu-

facturing sector is assumed to be relatively capital intensive. The total output of the manufacturing sector is given by

$$X_m = h(X_m)F_m(K_m, L_m) \tag{1}$$

where X_m = output, $K_m(L_m)$ = total capital (labor) employed in the manufacturing sector, F_m is assumed to be linear and homogenous in its arguments and has the usual properties of a standard production function. The function, h, captures the scale economies.

The industry output, X_m, is the sum of the outputs of all the firms in the manufacturing industry. The production function of a typical firm, k, in this industry is written as $X_m^k = h(X_m)\, F_m^k\left(K_m^k, L_m^k\right)$ where X_m^k, K_m^k and L_m^k represent the total output, capital and labor associated with firm k.

Let's define

$$\varepsilon \equiv (h/X_m)\left(\frac{dX_m}{dh}\right) \tag{2}$$

as the scale elasticity, where $X_m = \sum X_m$. We assume that $0 < \varepsilon < 1$, i.e., although there are economies of scale to produce more output, firms still need more inputs (see Jones 1968 and Mayer 1974). When economies of scale are external in nature, the perfectly competitive structure of the market is maintained. Thus, as price takers, firms equate the private value of marginal product of each factor to the factor's price. All firms are identical and thus

$$\begin{aligned} P_m h(X_m)\left(\partial F_m/\partial L_m^k\right) &= P_m h(X_m)(\partial F_m/\partial L_m) = W_m \\ P_m h(X_m)\left(\partial F_m/\partial K_m^k\right) &= P_m h(X_m)(\partial F_m/\partial K_m) = r_m \end{aligned} \tag{3}$$

where P_m represents the price of manufacturing good and W_m and r_m denote the wage rate and rental rate in the manufacturing sector.

In addition, the labor market in the urban sector follows the non-shirking constraint (NSC) since workers tend to shirk and cannot be perfectly monitored. Employers therefore offer wages higher than the equilibrium wage to deter shirking and fire workers caught shirking. This will result in non-shirking equilibrium with unemployment and higher wages. The fear of losing high wages and being unemployed will keep these workers from shirking. This is the efficiency wage model as posited by Shapiro and Stiglitz (1984). Let $b > 0$ denote a normal exogenous turnover rate, i is the discount rate, and q is the probability of being caught if shirking. Define V_m^U, V_m^S and V_m^N respectively as the expected lifetime utilities of the unemployed, shirking employees and non-shirking employees in the urban sector. Furthermore, define a as the probability of getting out of the pool of unemployed workers.[2] Assuming risk neutrality the asset value of equations, applicable to these three types of workers are

$$iV_m^U = a\left(V_m - V_m^U\right) \tag{4a}$$

$$iV_m^S = (W_m/P) + (b+q)\left(V_m^U - V_m^S\right) \tag{4b}$$

$$iV_m^N = (W_m/P) - e_o + b\left(V_m^U - V_m^N\right) \tag{4c}$$

[2] Note that in the analysis of non-shirking equilibrium, b,q,i and a are taken to be exogenous. Thus they can be manipulated to create a particular level of employment.

where e_o is the disutility of effort. V_m is the equilibrium wage and is equal to V_m^N. Workers would not shirk if $V_m^N \geq V_m^S$. Using the above three equations, the non-shirking condition (NSC) can be written as

$$W_m/P = e_o + e_0/q(bE_m/(E_m - L_m)+\text{i}) \tag{5}$$

where E_m is the total labor supply in the urban sector and L_m is the number of employed workers in the urban sector (see appendix for derivation of NSC).

According to the efficiency wage hypothesis, there will always be unemployment if non-shirking condition is satisfied. In Fig. 1, W_m* is an above equilibrium wage and the size of the unemployment is $L_m E_m$.

In the rural sector, we will have constant returns to scale in agricultural production. The production function is given by

$$X_a = F_a(K_a, L_a) \tag{6}$$

where X_a = Total output of the agricultural good.

K_a = Total amount of capital in the rural sector. L_a = Total amount of labor in the rural sector.

$$\begin{aligned} P_a(\partial F_a/\partial L_a) &= W_a \\ P_a(\partial F_a/\partial K_a) &= r_a \end{aligned} \tag{7}$$

where P_a, W_a, and r_a, respectively denote the commodity price, wage rate and rental rate in the rural sector. Inter-sectoral mobility of capital ensures that in equilibrium we will have

$$r_a = r_m \tag{8}$$

As regards the labor market, free inter-sectoral mobility allows laborer from the rural sector to move to the urban sector because manufacturing sector's wage is higher than the rural sector's wage. Eventually equilibrium is established when expected wages of the sectors are equalized (i.e., W_m^e, the expected wage of the urban sector, is equal to W_a, the rural sector's expected wage). This expected wage is defined to be actual wage times the probability of being employed, i.e.,

$$W_m \cdot \frac{L_m}{E_m} = W_m^e = W_a \tag{9}$$

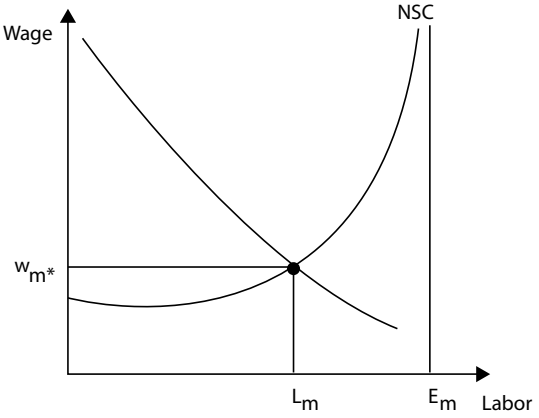

Fig. 1.

Finally, the total endowment of the economy is divided between the two sectors.

$$\begin{aligned} K_m + K_a &= K \\ L_m + L_a + U_m &= L = E_m + L_a \end{aligned} \tag{10}$$

Remember that $L_m < E_m$ and $E_m - L_m = U_m$ where E_m is the total number of workers available in the urban sector. After defining $\lambda = U_m/L_m$, Eq. (9) can be written as $W_m = W_a(1 + \lambda)$ Given $K, L, P_a, P_m, b, i, q, a$ and e_o the system of equations will solve for $X_m, X_a, L_m, E_m, L_a, U_m, K_m, K_a, W_a, W_m, r_a$, and r_m.

3. Migration equilibrium

To describe how the migration equilibrium works we first start with Corden and Findlay's geometric presentation of the H-T model. Total labor supply in the economy is presented by the horizontal distance O_mO_a in Fig. 2. L_m is measured to the right from O_m and L_a is measured to the left from O_a. V_m curve represents the value of marginal product in the manufacturing sector. In the absence of any distortion, the equilibrium is obtained at C.

But due to the rigid minimum wage (exogenously given), W_m, the urban employment is now O_mL_m. Since W_m is higher than W_a there will be migration until expected wages of both the sectors are the same, i.e. $W_m^e = W_a$.

The expected wage equality also suggests that $W_mL_m = W_aE_m$. The rectangular hyperbola through T such as qq would represent all the points whose value is W_mL_m. If V_a curve cuts this qq curve then we will have $W_mL_m = W_aE_m$. Migration equilibrium will then be reached with L_mE_m being the size of the unemployment. The migration equilibrium is shown in Corden and Findlay's analysis, when there is Constant Returns to Scale (CRS) in both the sectors.

In this paper, we have external economies of scale in the urban sector. Accordingly, we will need to look at Eq. (3) representing the value of marginal product of labor.

$$P_m h(X_m)(\partial F_m/\partial L_m) = W_m.$$

Differentiating this with respect to L_m we can obtain

$$e_{Lm} = \left(\hat{V}_m/\hat{L}_m\right) = -(\theta_{Km}/\delta_m) + \{\varepsilon/(1-\varepsilon)\}\theta_{Lm} \tag{11}$$

e_{Lm} is the elasticity of the value of marginal product curve with respect to a change in labor employed in the manufacturing sector. As L_m rises L_m/K_m rises and that will reduce the marginal product of labor. This is the factor endowment effect depicted by the first term in the above equation.[3] But as output rises due to the increase in labor supply, we see scale effect presented

[3]As it is customary in calculating the elasticity of the value of marginal product of labor, we are taking the capital supply of the manufacturing sector to be fixed. However, in the functioning of the model capital is free to move between the sectors. In other words, we are measuring e_{Lm} at a given amount of capital in the manufacturing sector. As L_m increases θ_{km}, the distributive share of capital in the manufacturing sector, goes down. The elasticity of substitutions between capital and labor, δ_m, could also change following the change in marginal product of labor. Even if we assume a constant, δ_m the ratio θ_{km}/δ_m goes down, i.e., the first term of Eq. (11) will have a negative effect on e_{Lm}.

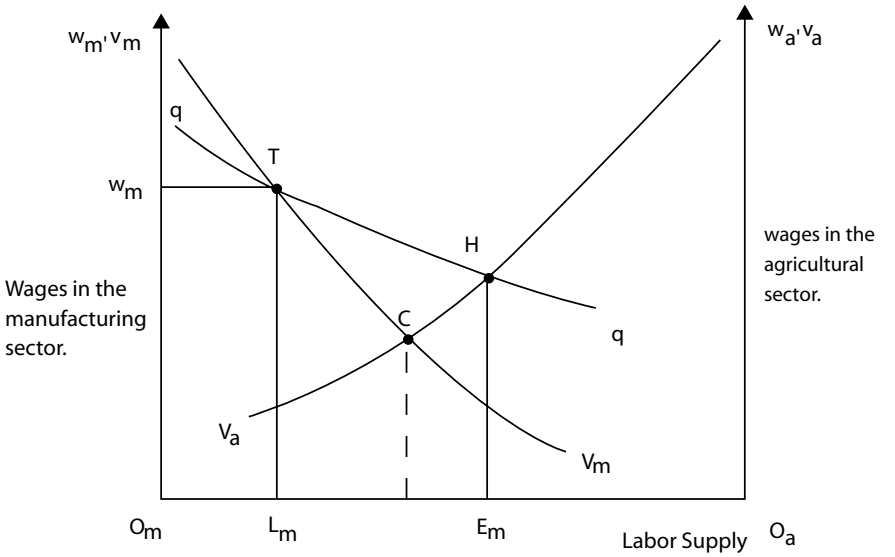

Fig. 2.

by the second term in the above equation. The scale effect would make the curve look flatter and if it is strong enough, we would have a positively sloped V_m curve. Thus, we might not have the existence of a stable equilibrium (see Neary 1981; Panagariya and Succar 1986; Wong 1995). To avoid that problem we are assuming that in spite of the presence of external economies of scale the V_m curve is negatively sloped to give us a stable equilibrium. In Fig. 2, we have W_m exogenously given. In this paper W_m is fixed endogenously by the intersection of NSC curve and the V_m curve as is shown in Fig. 3.

To start with, the urban sector has W_m and rural sector has W_a. The employment level in the urban sector is O_mL_m and in the rural sector is O_aE_m. If $W_m \cdot \frac{L_m}{E_m} > W_a$ labor starts moving from the rural sector. According to the non-shirking condition (5), *NSC* curve moves downward to *NSC*′. W_m consequently moves to W'_m, L_m moves to L'_m and W_a moves to W'_a. The employment in the urban sector rises. But the rise in L_m is less than the increase in E_m. In this case $W_m \cdot \frac{L_m}{E_m}$ or W^e_m goes down while W_a goes up. Eventually, they will be equal at H, when migration equilibrium with unemployment will be reached.

The difference between original HT model with full factor mobility and the model in this paper is that rural-urban migration in this model helps to increase employment in the urban sector, even though the unemployment rate might not go down. So either the decline in urban wage or the decline in both the urban wage and the probability of getting a job in the urban sector helps to bring the equilibrium following migration. In fact, urban wage in this model is not fully rigid. By forcing the wage to go down migration, in fact, reduces the distortion brought in by the wage differential.

Proposition. An increase in E_m increases equilibrium L_m, i.e., rural-urban migration increases urban employment.

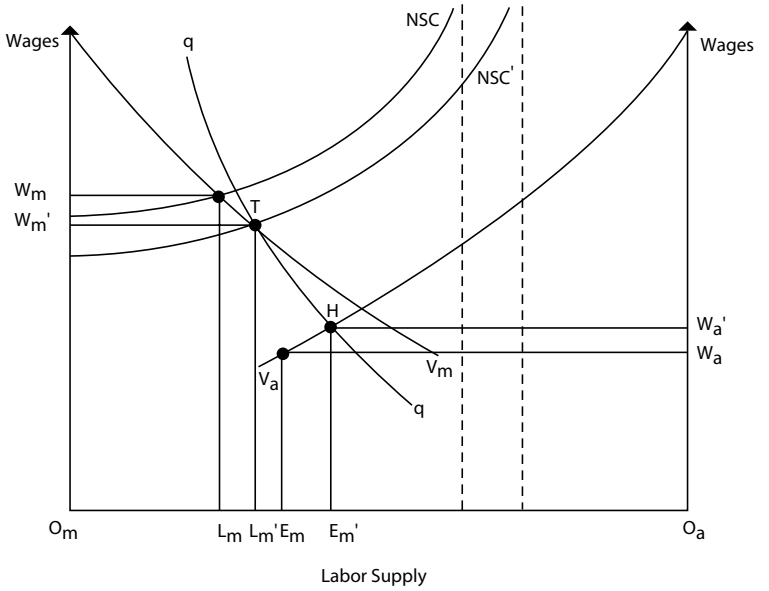

Fig. 3.

Proof. The equilibrium condition in urban labor market is given by

$$e_a + \frac{e_0 i}{q} + \frac{e_0 b E_m}{q(E_m - L_m)} = V_m \tag{12}$$

$$\text{or } A + B\frac{E_m}{E_m - L_m} = V_m \tag{13}$$

where $A = e_0 + e_0 i/q$ and $B = \frac{e_0 b}{q}$

$$L_m = \frac{E_m(A + B - V_m)}{A - V_m} \tag{14}$$

or

$$\frac{\partial L_m}{\partial E_m} = \frac{A + B - V_m}{A - V_m} + E_m\left\{\frac{\partial V_m}{\partial E_m} \cdot \frac{B}{(A - V_m)^2}\right\}$$

$$= \frac{(A + B - V_m)}{B}\left(\frac{A - V_m}{V_m}\right) + \frac{\hat{V}}{\hat{E}_m} \tag{15}$$

The second term is very similar to the elasticity of marginal product of labor that includes both scale effect and the factor intensity effect. We know from Eq. (3) that

$$P_m h(X_m)\left(\partial F_m / \partial L_m^k\right) = P_m h(X_m)(\partial F_m / \partial L_m) = V_m$$

where V_m is the value of marginal product of labor. Differentiating this term with respect to L_m we can get

$$-\frac{\theta_{km}}{\delta_m} + \left(\frac{\varepsilon}{1 - \varepsilon}\right)\theta_{Lm}. \text{ Thus}$$

$$\partial L_m/\partial E_m = \left(\frac{A+B-V_m}{B}\right)(A-V_m)\frac{1}{V_m} - \frac{\theta_{km}}{\delta_m} + \left(\frac{\varepsilon}{1-\varepsilon}\right)\theta_{Lm} > 0 \qquad (16)$$

From (13), we know that the first term on the right hand side will be positive and for large values of δ_m and ϵ the expression in (16) will be positive. That may also be true even when ε is not large enough so that the factor intensity effect, $\frac{\theta_{km}}{\delta_m}$, will dominate the scale effect, $\left(\frac{\varepsilon}{1-\varepsilon}\right)\theta_{Lm}$, to keep the V_m curve negatively sloped. Q.E.D.

Even though as a result of immigration employment in the urban sector increases, $\partial L_m/\partial E_m < 1$ because of the NSC condition which leaves some workers unemployed or $\frac{\partial(E_m - L_m)}{\partial E_m} > 0$.[4]

From Eq. (5), it follows that

$$\frac{\partial(W_m/P)}{\partial(E_m - L_m)} < 0^5$$

Thus in this model following rural-urban migration, urban employment increases and the sectoral wage difference goes down because of the increase in rural wage and the decline in the urban wage at the same time. This reduces the severity of distortion which is in sharp contrast to the effects of rural-urban migration with exogenous wage distortions.

It is important to note that the urban region in this paper has only one sector producing manufacturing good. Sometimes the urban region of a developing dual economy may have a formal sector and an informal sector.[5] Under the scenario, rural migrants may move to the urban informal sector and become self-employed by engaging in activities like selling street food or may enter into service industry (e.g., maid services to private household). Obviously, the informal sector providing these types of jobs is not expected to have either the efficiency wages or the external economies of scale. In that case we will have a formal sector with the efficiency wage distortion and external economies of scale and an informal sector with perfect wage flexibility. The conclusions of this paper will still hold even when we consider the existence of an informal sector. This is due to the fact that some of the rural migrants will opt to join the unemployment pool in the formal sector because they believe that it will provide a better chance of being employed in that sector compared to the chance they will get if they join the informal sector. Others may decide to find employment in the informal sector. The rural-urban migration will still be triggered by expected wage difference. However, for those who wait in the formal sector's unemployment pool, the expected wage W_m^e still equals $W_m \frac{L_m}{E_m}$ as in this paper's model, and for those who join the informal sector $W_m^e = \Pi W_m + (1-\Pi)W_I$ where $\Pi\left(= W_m \frac{L_m}{E_m}\right)$ is the probability of employment

[4] Even though following migration the number of unemployed increases, unemployment rate may go down if the scale effect dominates the factor intensity effect and V_m curve becomes more flat. In other words, the employment effect of migration depends on the elasticity of wage with respect to unemployment which is dependent on the shape of the V_m curve and the NSC curve.

[5] Agenor (1996) and Gatiga et al. (1995) have shown the large inter-industry wage differential within the modern urban sector which might be due to the co-existence of both a formal and an informal sectors in the urban regions.

in the formal sector and $(1 - \Pi)$ is the probability of employment in the informal sector and W_I is the wage of the informal sector.

4. Agglomeration under exogenous and endogenous distortion

The model in this paper's analysis differs from the original H-T model in two ways: a) it introduces external economies of scale, and b) it deals with efficiency wage distortion. In this section, I will show how the growth of the urban sector or the magnitude of structural transformation of a developing dual economy varies depending on whether we have only exogenous wage rigidity in the urban sector as in the H-T model, or we have exogenous wage rigidity together with external economies of scale in the urban sector, or we have endogenous wage rigidity and external economies of scale.

In ranking the policy alternatives involving employment subsidy and wage subsidy, Shukla and Stark (1990) have compared the first two cases. In the original H-T model, $X_m = F_m(K_m, L_m)$ and the exogenously given wage $\overline{W} = (\partial F_m/\partial L_m)P_m$. In terms of Fig. 4, the urban population according to H-T model is $O_m L_u$ and urban unemployment is equal to $L_m L_u$. If we take a specific production function of Cobb-Douglas (henceforth referred as C-D) nature then $X_m = AL_m^{\alpha}$ and $\bar{W} = \alpha A L_m^{\alpha-1}$.

If the manufacturing industry enjoys external economies of scale then following (1) the output of manufacturing sector is given by $X_m = h(X_m)F_m(K_m, L_m)$. In that case, although firms employ labor according to the value of private marginal product of labor, i.e.,

$$\overline{W} = h_m(X_m)\frac{\partial F_m}{\partial L_m}P_m,$$

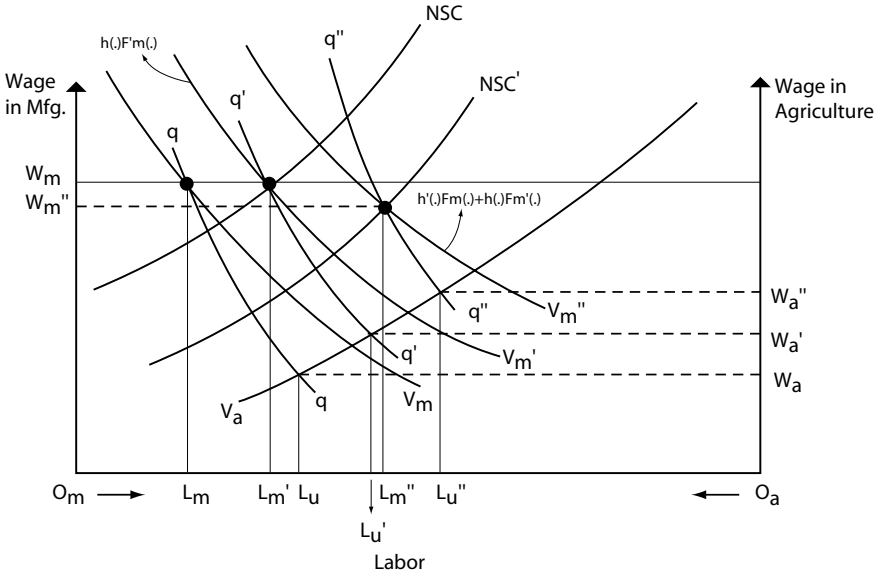

Fig. 4.

this is not the efficient employment for the industrial sector. To be efficient, labor should be hired according to the value of social marginal product of labor which says

$$\overline{W} = \left\{ h'(X_m)F_m(\cdot) + h(X_m)\frac{\partial X_m}{\partial L_m} \right\} P_m$$

Assuming that elasticity of economies of scale is such that stability condition is satisfied (see Sect. 3), the size of the urban sector and urban unemployment will be bigger than what it is under constant returns to scale production technology. In Fig. 4, employment under the scenario increases to $O_mL'_m$ and the size of the urban sector is now given by $O_mL'_u$. Again, using C-D production function

$$X_m = AL_m^{\gamma+\alpha}$$

$$\overline{W} = A(\gamma+\alpha)L_m^{\gamma+\alpha-1}$$

where γ is the scale parameter.

Now I want to introduce NSC curve in Fig. 4 following the efficiency wage distortion. Suppose the initial non-shirking wage is set at W_m.The employment-enhancing effect of inter-sectoral migration under the endogenous wage distortion together with external economies of scale, pushes the urban employment level further and agglomeration in the urban sector produces a very big effect as the size of the urban sector moves to $O_mL'_u$ and the urban employment is $O_mL'_m$. It is obvious that agglomeration effect will produce the best result when it works with efficiency wage distortion.

However, the size of the structural transformation and/or the growth of urban employment depend on the shape of the value of marginal product curve in the urban sector, and the shape of the NSC curve. In Fig. 5, the urban manufacturing sector is originally in equilibrium with W_m wage and O_mA employment. The rural sector's employment is at O_aC with wage W_a. If $A'C'(= CD)$ number of workers now moves to the urban sector, NSC moves to NSC', urban employment increases to O_mE, and the size of the urban sector increases to O_mD. The production in the rural sector decreases by $BCDF$ and production in the urban sector increases by $AEGA'$. If the growth of the urban sector's output falls short of the decline in rural sector's output, national income goes down. This is known as "urban bias" in the literature and as mentioned before, it depends on the elasticity of urban wage with respect to unemployment, i.e., the shape of the value of marginal product curve of labor and the NSC curve.[6] In their analysis of urban bias, Krichel and Levine (1999) have developed a measure of elasticity of wage with respect to urban unemployment using urban wage as a declining function of unemployment. In this section, I will develop a measure of wage elasticity using the non-shirking constraint and show that this wage elasticity depends not only on unemployment, but also on parametric values of turnover rate, b, the rate of discount, i, and the probability of getting caught while shirking, q. Thus it is

[6] This paper does not focus on the analysis of urban bias. However, urban bias depends on the growth of the urban sector relative to the reduction of the size (negative growth) of the rural sector. Thus, both the analysis of urban bias and the analysis of agglomeration in the urban sector depend on the elasticity of wage with respect to unemployment.

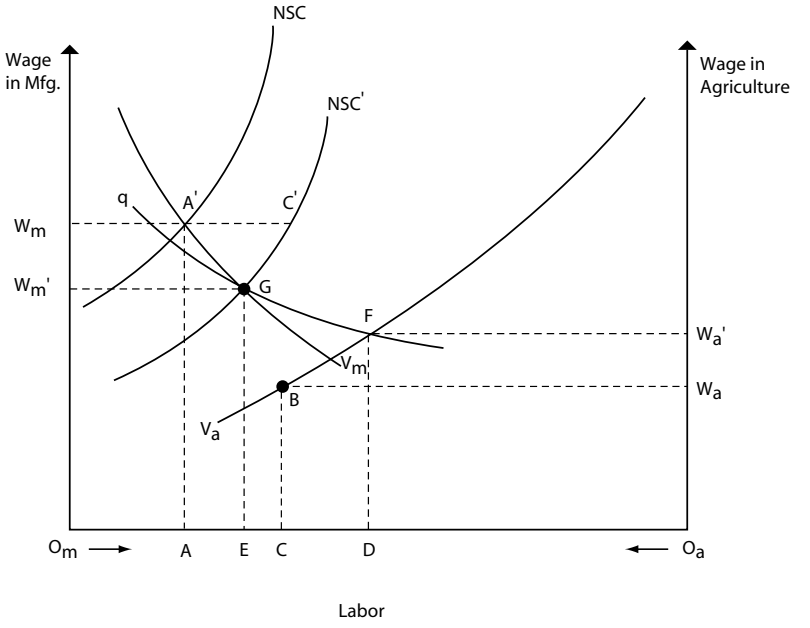

Fig. 5.

possible to achieve desirable level of urban growth by manipulating these parameters.

Using the C-D production function in the urban sector, the demand for labor is presented by

$$A(\alpha+\gamma)L_m^{\alpha+\gamma-1}$$

where $X_m = AL_m^{\alpha+\gamma}$.

From Eq. (5) we get the NSC as

$$w_m = e_0 + \frac{e_0 i}{q} + \frac{e_0 b}{q}\left(\frac{E_m}{E_m - L_m}\right) \text{ where } w_m = W_m/P$$

From implicit differentiation of this equilibrium condition we get

$$\frac{dL_m}{dE_m} = \frac{F_{E_m}}{F_{L_m}}$$

$$\frac{dL_m}{dE_m}\frac{E_m}{L_m} = \frac{\frac{e_0 b}{q}\frac{E_m}{(E_m-L_m)}\frac{L_m}{(E_m-L_m)}\cdot\frac{(E_m-L_m)}{L_m}}{A(\alpha+\gamma)(\alpha+\gamma-1)\frac{L_m^{\alpha+\lambda-1}}{L_m}\cdot L_m\frac{E_m-L_m}{L_m} - \frac{e_0 b}{q}\frac{L_m}{E_m-L_m}\cdot\frac{E_m}{E_m-L_m}\cdot\frac{E_m-L_m}{L_m}} \tag{17}$$

Let's define

$$\pi = \frac{\frac{e_0 b}{q}\cdot\frac{1}{U}}{w_m(\alpha+\gamma-1)\frac{E_m-L_m}{L_m} - \frac{e_0 b}{q}\cdot\frac{1}{U}} \tag{18}$$

π is the elasticity of employment with respect to migration and U is the unemployment rate. Let's write the NSC

$$w_m = e_0 + e_0 i/q + \frac{e_0 b}{q}\left(\frac{E_m}{E_m - L_m}\right)$$

as $w_m = e_0 + e_0 i/q + \frac{e_0 b}{q}\left(\frac{1}{U}\right)$

$$\frac{dw_m}{dU} = \frac{-e_0 b}{qU^2}$$

$$\text{or } \eta = \frac{dw_m}{dU}\cdot\frac{U}{w_m} = \frac{-e_0 b}{qU^2}\cdot\frac{U}{w_m} = \frac{-e_0 b}{qU\left(e_0 + \frac{e_0 i}{q} + \frac{e_0 b}{q}(1/U)\right)}$$

Dividing both numerator and denominator by qU we get

$$\eta = \frac{-e_0 b/qU}{e_0 + \frac{e_0 i}{q} + \frac{e_0 b}{qU}}$$

$$\eta\left(e_0 + \frac{e_0 i}{q} + \frac{e_0 b}{qU}\right) = e_0 b/qU$$

Substituting the value of $-e_0 b/qU$ in (18) we get

$$\pi = \frac{-\eta\beta}{w_m(\alpha + \gamma - 1)\frac{E_m - L_m}{L_m} + \eta\beta}$$

where $\beta = e_0 + \frac{e_0 i}{q} + \frac{e_0 b}{qU}$

$$\text{or } \pi = \frac{-\eta\beta}{w_m(\alpha + \gamma - 1)\frac{U}{1-U} + \eta\beta}$$

Define total output of the economy.

$$Z = X_m + PX_a - CM$$

where $X_m =$ output in the urban sector,
X_a = output in the rural sector,
CM = cost of migration,
P = Price level

$$\frac{dZ}{dE_m} = \frac{dX_m}{dL_m}\cdot\frac{dL_m}{dE_m} + P\frac{dX_a}{dL_a}\cdot\frac{dL_a}{dE_m} - \frac{dM}{dE_m}C$$

The optimization condition is

$$w_m\frac{dL_m}{dE_m} - W_a - C = 0$$

since $dL_a/dE_m = -1$, and $\frac{dM}{dE_m} = 1$

$$\text{or } w_m\pi\frac{L_m}{E_m} = W_a + C$$

$$\text{or } w_m\pi\frac{L_m}{E_m} = \frac{L_m}{E_m}w_m$$

$$\text{or } \frac{-\eta\beta}{w_m(\alpha+\gamma-1)\frac{U}{1-U}+\eta\beta}\frac{L_m}{E_m}=\frac{L_m}{E_m}$$

$$\text{or } \eta=\frac{E_mL_mw_m(\alpha+\gamma-1)U/1-U}{2\beta L_mE_m} \text{ Dividing by } E_mL_m$$

$$\eta=\frac{w_m(\alpha+\gamma-1)U/1-U}{-2\beta}$$

$$\eta=w_mD \text{ where } D=\frac{((\alpha+\gamma-1)U/1-U)^7}{-2\beta}$$

Since D and thus η can be controlled by manipulating b, i, q and a, it is possible to influence the growth of the urban sector or the structural transformation by manipulating these parameters.

5. Agglomeration under external and internal economies of scale

Rural-urban migration in this paper helps in structural transformation of a dual economy by promoting urban growth. This happens in the presence of endogenous wage distortion and external economies of scale. The equilibrium is still an unemployment equilibrium. The literature on economic geography using internal economies of scale explains agglomeration without any form of wage distortion (Fujita et al. 1999; Krugman 1993). In the studies of internal economies of scale, resources are fully employed all the time. In that sense, a comparison between external and internal economies of scale might not have any direct bearing on the results of this paper. However, it will help to gain further insight into the analysis of rural-urban migration and urban growth.

Rural-urban migration in the presence of unemployment in the urban sector increases employment in this paper. Thus, it reduces the severity of unemployment compared to that under exogenous wage distortion. Furthermore, as shown in the preceding section, the volume of migration is larger in this paper's model than it is under exogenous distortion with a rigid wage (original H-T model) or under exogenous wage distortion with external economies of scale. In the original H-T model the dynamics of the adjustment process works through the changes in the probability of employment in the urban sector. In this paper's model, the dynamics works through the declining urban wage (because of the endogenous distortion) and the changes in the probability of employment. Although employment goes up, the probability of employment (measured by the ratio of urban employed to total urban labor force) may move in either direction depending on the elasticity of wage with respect to unemployment. The adjustment process in this model may, therefore, be faster than that under the exogenous distortion.

[7] Detailed derivations can be made available upon request.

External economies in this model work through the scale of industry output. In the case of internal economies, scale effect works through the variety of the manufacturing goods produced. The utility increases as variety increases. With internal economies of scale we see the forward linkage effect when increased variety reduces the price index of the region of concentration (Fujita et al. 1999). The backward linkage works through higher wage in the region of concentration. No such price index effect is possible with external economies of scale. However, since rural-urban migration is triggered by higher expected wage, we can see some form of the backward linkage type effect under external economies of scale. The models of economic geography in the internal economies of scale relies on transportation costs (for trading goods) to explain labor mobility. It is important to note that the models with external economies of scale do not use transportation costs at all.

Is it at all practical to see the urban sector enjoying or exploiting agglomeration economies with endogenous wage distortion? We can say that at the very early stage of development, the developing countries have used minimum wage legislation (India, Hong Kong, China, etc.). As development proceeds and these countries have emerged as newly industrialized nations, more private enterprising is encouraged and firms accordingly are eager to provide incentive for work effort. These incentives may come in the form of efficiency wage (see also Footnote 1), and thus it is quite feasible for urban sectors of these countries to reduce the severity of unemployment resulting from wage distortion and experience the structural transformation by exploiting the external economies of scale that are present at the early stages of development

6. Comparative Static Analysis[8]

In a regional economy with labor market distortion (such as H-T economy), Panagariya and Succar (1986) have shown that the Rybczynski effect will be valid with external economies of scale.[9] In this paper I examine the effects of factor endowment changes in the presence of external economies of scale, when wage distortion takes an endogenous form. A growth in capital supply increases manufacturing output, as it does in the case of exogenous wage distortion, i.e.,.

$$\hat{X}_m/\hat{K} = (\Pi_{22}\Pi_{33} - \Pi_{23}\Pi_{32})/|\Pi| > 0 \text{ where}$$

$$\Pi_{22} = \lambda_{La}$$

$$\Pi_{33} = ((\theta_{Lm} - \theta_{La} + \theta_{Km}\delta_m\theta_{La})/(\theta_{Lm} - \theta_{La}))$$

$$\Pi_{23} = (a\lambda_{Lm}\theta_{Km}\delta_m\theta_{La}/(\theta_{Lm} - \theta_{La})) + (\lambda_{La}\theta_{Ka}\delta_a\theta_{Lm}/(\theta_{Lm} - \theta_{La})) + \lambda_{Lm}\alpha$$

$$\Pi_{32} = 0$$

In my analysis $|\Pi| > 0$ and all $\Pi_{i's}$ except Π_{32} are positive.

[8] Detailed derivation of all comparative static analyses can be obtained from the author.

[9] According to the Rybczynski theorem, an increase in the endowment of one factor of production must cause the industry using that factor intensively to expand given the factor and commodity prices.

The changes in factor supply will affect the factor prices in a paradoxical way, e.g., an increase in K will raise r. However, it will be less paradoxical compared to the case where production enjoys economies of scale and the labor market has an exogenously fixed wage. In the analysis of this paper, a part of the scale effect is outweighed by the endogenous nature of the distortion in the labor market. As K increases, X_m expands. For an exogenously fixed W_m, r_m would increase by a large amount to take care of external economies of scale, and W_a would fall when r_a increases to match with r_m. In this paper, as K_m rises, the V_m curve shifts to the right and NSC condition allows to have a higher W_m. So r_m and r_a will not rise as much as they would have in the absence of this endogenous labor market distortion. As W_m and L_m rises, the expected wage in the urban sector goes up and more labor start flowing to the urban sector. Since $\partial L_m/\partial E_m < 1, W_m^e$ in the long run equilibrium may move in the either direction but the new equilibrium wage would be higher than the initial one. In other words, there will be a positive effect on actual wage, but the effect on the expected wage is ambiguous. Thus we see that

$$\frac{\hat{W}_m}{\hat{K}} = ((A_{22}A_{33} - A_{23}A_{32})/|A|) > 0$$

$$\frac{\hat{r}}{\hat{K}} = -(A_{21}A_{33} - A_{23}A_{31}/|A|) > 0$$

$$\frac{\hat{W}_a}{\hat{K}} = ((A_{21}A_{32} - A_{22}A_{31})/|A|) < 0$$

where

$$A_{21} = \alpha\lambda_{Lm}\left(\frac{1-\varepsilon}{\varepsilon}\right)\theta_{Lm} + \frac{\lambda_{La}T'X_m}{X_a}\frac{\theta_{Lm}}{\varepsilon} - \alpha\lambda_{Lm}\theta_{Km}\delta_m + \lambda_{Lm}\alpha$$

$$A_{22} = \alpha\lambda_{Lm}\left(\frac{1-\varepsilon}{\varepsilon}\right)\theta_{Km} + \frac{\lambda_{La}T'X_m}{X_a}\frac{\theta_{Km}}{\varepsilon} + (\alpha\lambda_{Lm}\theta_{Km}\delta_m + \lambda_{La}\theta_{Ka}\delta_a)\frac{\theta_{Km}}{\theta_{Lm} - \theta_{La}}$$

$$A_{23} = \left\{-\frac{1}{\theta_{Lm} - \theta_{La}}(\alpha\lambda_{Lm}\theta_{Km}\delta_m\theta_{La} + \lambda_{La}\theta_{Ka}\delta_a\theta_{lm}) - \lambda_{Lm}\alpha\right\}$$

$$A_{31} = \big(\theta_{Lm} - \theta_{La} + \theta_{Km}\delta_m\theta_{La}\big)/\big(\theta_{Lm} - \theta_{La}\big) + \big(((1-e)(\theta_{Lm} - \theta_{La}) - \theta_{Km}\delta_m\varepsilon)/(\theta_{Lm} - \theta_{La})\big)\frac{\theta_{Lm}}{\varepsilon}$$

$$A_{32} = \frac{\theta_{Km}}{\varepsilon}\{((1-\varepsilon)(\theta_{Lm} - \theta_{La}) - \theta_{Km}\delta_m\varepsilon)/(\theta_{Lm} - \theta_{La})\}$$

$$A_{33} = ((\theta_{Lm} - \theta_{La} + \theta_{Km}\delta_m\theta_{La})/\theta_{Lm} - \theta_{La})$$

As far as changes in the commodity prices are concerned, the results (Stolper-Samuleson effect) deviate from what we have seen with exogenously fixed wages in the urban sector. When P_m goes up and there are external economies of scale in the urban sector with W_m fixed exogenously, the rise in urban production is larger than what it could be under constant returns to scale. If W_m is fixed, the entire benefit goes to r, i.e., capital enjoys a much

higher return compared to that under CRS.[10] Thus W_a falls to compensate for an increase in r_a. The fall in L_a and K_a reduces X_a.

In this paper as P_m goes up, W_m changes. This happens because the value of marginal product goes up, i.e., V_m curve moves upward and the intersection of this new V_m curve and NSC curve produces a higher W_m. As W_{m} goes up, the increase in r_{m} will be less than what it would have been with external economies of scale and exogenously fixed urban wage. Still the increase in r_m causes capital to flow from the rural sector. This results in a fall in W_a. Thus both capital and labor flows out of the rural sector. In the long run equilibrium r will rise and W_a will fall. Although the direction of change in r and W_a is the same as it is in the models of exogenously fixed wages and external economies of scale, the magnitude of these changes will be much less because a part of the effect of economies of scale is used by the endogenous change (rise) in W_m. As in the case of exogenously fixed wage, L_a, K_a and X_a will go down. In the long run L_m will rise, but since the size of the urban labor market rises more, unemployment rate may increase too unless the scale effect dominates the factor intensity effect and/or elasticity of wage with respect to unemployment is favorable.

7. Conclusion

The original HT model presents a migration equilibrium in a dual economy with exogenously given wage in one of the sectors. This equilibrium is an unemployment equilibrium because of the labor market distortion created by the rigid wage in the urban sector. Similar unemployment equilibrium is also observed when external economies of scale are added to the labor market distortion in the urban sector (Panagariya and Succer 1986). My paper shows that a wage distortion that is tied to the work efficiency or shirking by workers in the urban sector also generates a migration equilibrium with unemployment even when it works along with external economies of scale in that sector. Did we gain anything by looking into this new type of wage distortion? We have shown that the equilibrium resulting from free rural-urban migration in this model raises the level of employment in the urban sector as well as in the entire economy. Unfortunately, it still may have to settle with some amount of unemployment.

Unlike with the exogenous wage distortion, the severity of unemployment can be reduced in this model by reducing turnover rate or increasing the probabilities of getting caught while shirking. These will not have as much budgetary pressure as can be created by tax and subsidy policies that are recommended to reduce exogenous wage distortion.

Furthermore, we have seen that scale of structural transformation of a dual economy is bigger under endogenous distortion than under exogenous distortion. The economies of scale are better exploited under endogenous wage distortion because of the employment enhancing effect of migration with efficiency wages.

[10] Before we allow any capital mobility, the immobile factor capital captures all the gains from an increase in P_m.

It has also been shown that a change in the factor endowment changes factor prices in a paradoxical way. For example, an increase in capital supply raises return to capital as is it does in the case of H-T model with external economies of scale (Panagariya and Succar 1986), but the paradox is less severe in my paper because of the endogenous nature of the labor market distortion. This is due to the fact that a part of the gains from external economies of scale is used by the labor market. It also intensifies the sectoral wage differential.

The analysis in this paper suggests that the authorities should evaluate their options before imposing an exogenous distortion in the labor market (like minimum wage legislation) and/or before trying to remedy the effects of such distortions, which may have severe budgetary consequences. If employers are already using efficiency wages or are eager to use them, they should be encouraged so that they actively try to increase the probability of getting caught if shirking, or to reduce the turnover rate. This will reduce the distortion and at the same time help to enjoy the employment enhancing effect of inter-sectoral labor mobility (until unemployment equilibrium is reached). Some of these can be achieved by improving the information systems or making the use of existing information systems more efficient. This might be done at a cost that is less than the cost of tax and subsidy policies recommended for reducing exogenous wage distortions.

References

Agenor RP (1996) The labor market and economic adjustment. IMF Staff Papers 43: 261–335

Corden WM, Findley R (1975) Urban unemployment, intersectoral capital mobility and development policy. Economica: 59–78

Fujita M, Krugman P, Venables AJ (1999) The spatial economy. The MIT Press, Cambridge, MA

Gatiga J, Mizala A, Romaguera P (1995) Interindustry wage differential in Brazil. Economic Development and Cultural Change 43: 315–331

Goldstein GS, Moses LN (1973) A survey of urban economics. Journal of Economic Literature:471–515

Harris JR, Todaro MP (1970) Migration, unemployment and development: A two sector analysis. American Economies Review: 126–142

Jones RW (1968) Variable returns to scale in general equilibrium theory. International Economic Review 3: 261–235

Krichel T, Levine P (1999) The welfare economies of rural-to-urban migration: The Harris-Todaro model revisited. Journal of Regional Science 39(3): 429–447

Krugman P (1993) First nature, second nature and metropolitan location. Journal of Regional Science 33: 129–144

Mayer W (1974) Variable returns to scale in general equilibrium Theory: A comment. International Economic Review: 225–235

Neary JP (1981) On the Harris-Todaro model with intersectoral capital mobility. Economica: 219–234

Panagariya A, Succar P (1986) The Harris-Todaro model and economies of scale. Southern Economics Journal: 986–998

Shapiro C, Stiglitz JE (1984) Involuntary unemployment as a worker discipline device. American Economic Review 74: 433–444

Shukla V, Stark O (1990), Policy comparisons with an agglomeration effects augmented dual economy model. Journal of Urban Economies 27: 1–27

Wong K (1995) International trade in goods and factor mobility. The MIT press, Cambridge MA

Appendix

Derivation of the non-shirking condition and a subsequent comparative static analysis

According to the Efficiency Wage theory posited by Shapiro and Stiglitz (1984), workers' asset utility varies depending on whether the worker is employed and shirking, or employed but not shirking or unemployed. Given an exogenous turnover rate, b , a probability of being caught if shirking, q , a discount rate , i and a job acquisition rate, a , the asset value equation of a worker who shirks is given by

$$iV^s = (b+q)(V^u - V^s) + w_m \tag{A1}$$

Here iV_s (interest rate times asset value) equals flow benefits ($(w_m = W_m/P)$ plus expected loss (which includes unemployment resulting either from normal turn over or from being fired after caught shirking). The similar equation for a worker who is employed but not shirking is given by

$$iV^n = w_m - e_0 + b(V^u - V^n). \tag{A2}$$

Note that disutility from work, e_0, is taken into consideration. For an unemployed worker the asset value equation is given by

$$iV^u = a(V^n - V^u) \tag{A3}$$

Workers stop shirking if and only if $V^n \geq V^s$ (utility from non-shirking $\geq$ utility from shirking). Using (A1) and (A2), the no shirking condition (NSC) is written as

$$w_m \geq iV^u + (i+b+q)e_0/q \tag{A4}$$

Using (A2) and (A3) we can write

$$iV^n = (w_m - e_0)(a+i)/(a+b+i) \tag{A5}$$

$$iV^u = (w_m - e_0)a/(a+b+i) \tag{A6}$$

Substituting the expression for V^u in Eq. (A4), we get the NSC

$$w_m > e_0 + e_0(a+b+i)/q \tag{A7}$$

Since in steady state equilibrium $a = bL_m/E_m - L_m$, Eq. (A7) can be written as

$$w_m > e_0 + e_0 i/q + (e_0 b E_m)/(qE_m - L_m) \tag{A8}$$

From (A8) we can write $L_m = E_m(A - e_0 b)/A$

where $A = w_m q - e_0 q - e_0 i$

E_m = Total labor supply in the urban sector

L_m = Employed labor in the urban sector.

Thus, $0 < (\partial L_m/\partial E_m) < 1$ or $(\partial U_m/\partial E_m) = 1 - (\partial L_m/\partial E_m) > 0$.

This implies that any increase in the size of the urban labor market will increase the number of both employed and unemployed workers. Thus, when the rural migrants join the pool of existing unemployed workers they raise both employment and unemployment levels in the economy. The unemployment rate, however, may go down since the urban sector enjoys economies of scale (see p. 615 and Footnote 1).