

Volume 68

2008

**Rural-to-urban migration, human capital,
and agglomeration**

By C. SIMON FAN and ODED STARK

Reprinted from

**JOURNAL OF ECONOMIC
BEHAVIOR & ORGANIZATION**

© 2008 Elsevier B.V.

Rural-to-urban migration, human capital, and agglomeration

C. Simon Fan^a, Oded Stark^{b,c,d,e,f,*}

^a Lingnan University, Hong Kong, China

^b University of Klagenfurt, Austria

^c University of Bonn, Germany

^d University of Vienna, Austria

^e Warsaw University, Poland

^f Warsaw School of Economics, Poland

Received 30 November 2006; received in revised form 18 April 2008; accepted 18 April 2008

Available online 26 April 2008

Abstract

A new general-equilibrium model that links together rural-to-urban migration, the externality effect of the average level of human capital, and agglomeration economies shows that in developing countries, unrestricted rural-to-urban migration reduces the average income of both rural and urban dwellers in equilibrium. Various measures aimed at curtailing rural-to-urban migration by unskilled workers can lead to a Pareto improvement for both the urban and rural dwellers. In addition, the government can raise social welfare by reducing the migration of skilled workers to the city. Moreover, without a restriction on rural-to-urban migration, a government's efforts to increase educational expenditure and thereby the number of skilled workers may not increase wage rates in the rural or urban areas.

© 2008 Elsevier B.V. All rights reserved.

JEL classification: B12; H21; O15; O18; R13; R23

Keywords: Rural-to-urban migration; The externality effect of the average level of human capital; Agglomeration economies; Public policies

1. Introduction

In this paper, a new model of rural-to-urban migration is developed, with an emphasis on the role of human capital in both urban and rural economic activities. Notable exceptions notwithstanding, a substantial literature on rural-to-urban migration has taken its cue from the dual economy model of Harris and Todaro (1970). Their model assumes that the urban sector produces manufactured goods using (homogeneous) labor and physical capital as factors of production, and that the rural sector produces agricultural goods using (homogeneous) labor and land as factors of production. The model has been widely used as a basic analytical framework for studying rural-to-urban migration in developing countries and as a platform for policy formation.

However, since the Harris and Todaro model ignores human capital as a factor of production, it appears to have become increasingly less applicable to many developing countries in modern times. For example, due to continuing structural changes in recent decades, cities in the developing world have become more oriented toward service and

* Corresponding author at: ZEF, University of Bonn, Walter-Flex-Strasse 3, D-53113 Bonn, Germany.
E-mail address: ostark@uni-bonn.de (O. Stark).

(relatively) high-tech industries. Also, with a growing share of manufactured goods becoming standardized, these goods can be produced anywhere that offers basic infrastructure such as piped water and reliable electricity. Consequently, the production of manufactured goods increasingly takes place in the rural areas of developing countries.¹ This realignment enhances the importance of human capital in the rural areas. For example, Taylor and Yunez-Naude (2000) find that in rural Mexico, the returns from schooling are high both in crop and noncrop activities; as schooling levels increase, the returns from schooling arise from activities other than crop production. In China, the township and village enterprises (TVEs) have played a significant role in the country's economic growth since the early eighties. In 2000, for example, TVEs accounted for 47 percent of the total industrial output in China (Fu and Balasubramanyam, 2003), and the output value of TVEs has been far greater than the output value of agriculture. Yang and An (2002), and Yang (2004) show that education not only increases productivity in the nonagricultural sector in rural China, but that it also facilitates and encourages the relocation of productive inputs from agricultural to nonagricultural pursuits. Jonasson (2007) finds that in rural Peru, nonagricultural rural employment is a prerequisite for positive returns to education, and that education is rewarded by rural-based nonagricultural work. A perception that the rural areas in developing countries are an exclusive domain of uneducated peasants who apply physical labor to eke out a living had better be discarded.

In this paper we develop a new policy-yielding model of migration in which human capital is important in both the urban and rural areas. In line with considerable research in urban economics and economic growth, agglomeration economies in the cities are built into the model.² In a simple general-equilibrium framework, our model interlinks three key factors: the process of migration from the rural area to the urban area, the externality effect of the average level of human capital, and agglomeration economies.³

We postulate that a city's productivity is determined by its average level of human capital and by the size of its labor force. The productivity of the rural area is determined by its average level of human capital. Right at the outset, the analysis yields a rather surprising implication: in developing countries, unrestricted rural-to-urban migration reduces the average income of both rural and urban dwellers in equilibrium. This result implies that although a city attracts all the skilled individuals and enjoys the benefit of agglomeration economies (which the rural areas do not), with free labor mobility, the city's productivity is still very low in equilibrium. The intuition underlying this result is quite simple: since the returns to skills are higher in the urban areas than in the rural areas, as is typically the case in developing countries, skilled workers are likely to concentrate in the cities. Consequently, the wage rate of the unskilled workers in the rural areas will be low, which in turn will induce a large number of unskilled rural workers to leave for the cities. With free labor mobility, the rural-to-urban migration process will come to a halt when the urban and rural wages for unskilled workers are equalized: the urban wage will decline continuously with the in-migration of unskilled workers which, in turn, will reduce the average level of human capital in the city. In other words, unrestricted rural-to-urban migration results in a lower wage for unskilled workers in both the urban and rural areas. Furthermore, since the wages of skilled and unskilled workers in the urban area are affected by common productivity factors, the wage for the skilled workers will be driven to a low level by the unrestricted rural-to-urban migration. Thus, our model explains the *negative* consequences of rural-to-urban migration in the developing world, as is amply highlighted, for example, in nearly every leading development economics textbook (cf. Gillis et al., 1996; Ray, 1998; Todaro, 2000).⁴

In essence, our results arise from the difference between the private human capital and the social returns to human capital, a difference that implies that free labor mobility leads to an equilibrium that is not socially optimal. In all

¹ For surveys on the importance in recent decades of the rural non-farm sector for the economies of developing countries see Lanjouw and Lanjouw (2001) and Reardon et al. (2001).

² Several interesting studies that incorporate agglomeration economies into models of rural-to-urban migration (for example, the survey by Abdel-Rahman and Anas, 2004) tend to abstract from the unique characteristics of developing countries.

³ The agglomeration economies of a city have been studied extensively in the urban economics literature and, lately, in the economic growth literature (Black and Henderson, 1999a,b; Henderson, 2003). There is also a growing awareness of late of the importance of the (positive) externality effect of the average level of human capital for a city's productivity (Black and Henderson, 1999b; Lucas, 2001; Glaeser and Saiz, 2004; Moretti, 2004). However, to the best of our knowledge, no attempt has been made thus far to examine systematically the combined repercussions of rural-to-urban migration, human capital spillovers, and agglomeration effects. Shukla and Stark (1990) analyze several policy implications of agglomeration economies in the city for rural-to-urban migration. They assume homogeneous workers and do not attend to human capital considerations. Bertinelli and Black (2004) investigate a model of rural-to-urban migration with congestion costs in the city. They abstract from the consideration of agglomeration economies and the contemporaneous externality effect of the average level of human capital.

⁴ Moreover, our model applies not only to migration from rural areas to urban areas, but also to migration from towns and counties to cities, or even from small cities to large cities.

countries in general, and in developing countries in particular, there is an urban–rural wage gap for educated workers.⁵ Consequently, with free movement of labor, all skilled workers will cluster in the urban area, leaving the rural area with an average level of human capital that is below the social optimum.

Our analysis yields several interesting policy insights. First, our model shows that various measures aimed at curtailing rural-to-urban migration by unskilled workers, which include subsidizing the rural sector and restricting rural-to-urban migration,⁶ can lead to a Pareto improvement for both urban and rural dwellers. Second, and somewhat surprisingly, our analysis shows that the government can raise social welfare by reducing the migration of skilled workers into the city and by subsidizing some skilled workers to move from the city to the rural areas or, for that matter, to stay in the rural areas. Third, our model shows that in developing countries, when there is nothing to deter rural-to-urban migration, a government’s effort to increase educational expenditures and thereby the number of skilled workers may not increase the wage rates in rural or urban areas as long as there are still a large number of unskilled workers in the rural area in equilibrium. In fact, a rather perplexing result is that as the number of skilled workers (in the city) increases, the average level of human capital in the city may very well *decrease*.

2. The basic analytical framework

In what follows we develop a quite simple model in order to highlight our essential ideas. A detailed discussion of the robustness of the model is provided in Section 5.

Consider a small open economy that consists of an urban area (the city) and a rural area (the countryside). There are two types of workers: skilled and unskilled. Efficiency labor is the only factor of production in both the city and the countryside. The labor markets in the city and in the countryside are both perfectly competitive. The urban sector produces a single good, and the rural sector produces a different single good. The aggregate production in the city is according to a constant returns to scale production technology

$$Y^c = Fn, \quad (1)$$

where Y^c and n are the total output and the sum total of the efficiency units of labor employed in the city, respectively. Akin to Black and Henderson (1999b), we assume that F , the productivity factor of the city, is determined by n , the agglomeration effect of the city, and by the city’s fraction of skilled labor (the number of skilled workers in the city divided by the total number of workers in the city), which is denoted by h^c . Namely, we define

$$F \equiv f(n, h^c).$$

We assume that

$$\frac{\partial f(n, h^c)}{\partial n} = f_1(n, h^c) > 0, \quad \frac{\partial f(n, h^c)}{\partial h^c} = f_2(n, h^c) > 0. \quad (2)$$

The assumption $f_1(n, h^c) > 0$ means that there is an agglomeration effect in the production in the city. This is a standard assumption in urban and regional economics (see, for example, Black and Henderson, 1999a; Henderson, 2003). The assumption $f_2(n, h^c) > 0$ means that the production efficiency of the city depends positively on the average level of human capital of the city; that is, there is a positive externality effect of the average level of human capital (see, for example, Acemoglu, 1996; Black and Henderson, 1999b; Lucas, 2001; Glaeser and Saiz, 2004; Moretti, 2004; Liu, 2007).

Since the economy is small and open, the prices of its outputs are determined in the world market and are thus independent of the economy’s outputs. We normalize the price of the good produced in the urban sector of the economy to one. Since the labor market in the city is perfectly competitive, the wage rate per efficiency unit of labor in the city is equal to the marginal product of efficiency labor in the city, namely to

$$f(n, h^c).$$

⁵ For example, see McCormick and Wahba (2005) and Naughton (2007).

⁶ The policy of restricting rural-to-urban migration has been implemented by a number of developing countries and, as noted in Section 3, is exemplified by the “Hukou” system in China that strictly restricts rural-to-urban migration (Au and Henderson, 2006).

We assume that an unskilled worker is endowed with one unit of efficiency labor. Skilled workers are assumed to be homogeneous. We also assume that the private returns to human capital are higher (or at least marginally higher) in the city than in the countryside.⁷ Accordingly, we assume that a skilled worker is endowed with α units of efficiency labor if he works in the city, and with β units of efficiency labor if he works in the countryside, and that

$$\alpha > \beta > 1. \quad (3)$$

Thus, in the urban area, the earnings of an unskilled worker and the earnings of a skilled worker are $f(n, h^c)$ and $\alpha f(n, h^c)$, respectively.

The aggregate production in the rural sector is according to a constant returns to scale production technology

$$Y^r = HL^r, \quad (4)$$

where Y^r and L^r are the total output and the sum total of the efficiency units of labor employed in the countryside, respectively. In line with the discussion in the Introduction, we assume that H , the productivity factor of the rural sector, is determined by the fraction of skilled workers in the rural area, which is denoted by h^r .⁸ Namely, we define

$$H = \frac{1}{P}g(h^r),$$

where P is a positive coefficient, and where we assume that $g'(h^r) > 0$. Thus, the average level of human capital is a production factor in both the rural area and the urban area, although the private returns to human capital may well be greater in the city than in the countryside. However, in line with the received literature, we assume that the agglomeration effect exists only in the urban area.

We assume that the price of the good produced in the rural sector is P . Then, the value of the marginal product of an efficiency unit of labor in the rural area is $g(h^r)$. Since the labor market in the rural area is perfectly competitive, the wage rate per efficiency unit of labor in the countryside is equal to $g(h^r)$. Thus, in the rural area, the earnings of an unskilled worker (that is, the wage for his efficiency unit of labor) and the earnings of a skilled labor are $g(h^r)$ and $\beta g(h^r)$, respectively.

Finally, we assume that all individuals supply their efficiency units of labor inelastically. Then, since prices are constant, the indirect utility function of every individual is his income. Thus, in this paper we assume that every individual seeks to maximize his income.

Consider now unhindered rural-to-urban migration. To concentrate on essentials, for most of this paper we do not consider the process of human capital formation, taking the numbers of skilled and unskilled workers as exogenously given. (This assumption is relaxed, however, in Section 4.2.) We introduce the following notations:

s^c : the number of skilled workers in the city before migration occurs.

l^c : the number of unskilled workers in the city before migration occurs.

s^r : the number of skilled workers in the rural area before migration occurs.

l^r : the number of unskilled workers in the rural area before migration occurs.

m^s : the number of skilled migrants.

m^l : the number of unskilled migrants.

We first examine the equilibrium conditions when there are no restrictions on rural-to-urban migration. The rural-to-urban migration of the unskilled workers ceases if and only if the rural wage and the urban wage for unskilled workers are equalized, as long as there are unskilled workers in the rural area in equilibrium. That is, if the solution is interior, then the number of unskilled migrants, m^l , will be determined by the following equation:

$$f(n, h^c) = g(h^r). \quad (5)$$

⁷ This assumption is in line with the evidence (see, for example, Schultz, 2004; Naughton, 2007) and is consistent with the presumption that the returns to an individual's skills are positively affected by the average level of human capital. In fact, in the related theoretical literature such as Bertinelli and Black (2004) and Lucas (2004), an extreme assumption is made that the returns to human capital in the rural area are zero.

⁸ That the wage of an unskilled worker increases with the proportion of skilled workers in the rural area can be attributed to the complementarity between skilled and unskilled labor in the rural area, as well as to the externality effect of the average level of human capital.

From the assumptions in (3), we know that the wages for a skilled labor in the city and in the countryside are $\alpha f(n, h^c)$ and $\beta g(h^r)$, respectively. From (5) and the assumptions in (3), we know that

$$\alpha f(n, h^c) = \alpha g(h^r) > \beta g(h^r), \quad (6)$$

which implies that if there is rural-to-urban migration of both unskilled and skilled workers, then all skilled workers will end up in the city, which in turn implies that $h^r = 0$. Thus, we can rewrite (5) as

$$f \left[l^c + \alpha(s^c + s^r) + m^1, \frac{s^c + s^r}{l^c + s^c + s^r + m^1} \right] = g(0). \quad (7)$$

As an added comment, note that we assume that the following condition is satisfied so that not all unskilled workers migrate to the city in equilibrium:

$$f \left[l^c + \alpha(s^c + s^r) + l^r, \frac{s^c + s^r}{l^c + s^c + s^r + l^r} \right] < g(0). \quad (8)$$

Since the solution is interior, we must have that $m^1 < l^r$. Then, clearly, we have the following proposition.

Proposition 1. *Without restrictions on rural-to-urban migration, the wage rates for unskilled workers and for skilled workers are, respectively, $g(0)$ and $\alpha g(0)$ in equilibrium.*

Note that $g(0)$ is the lowest possible wage for unskilled workers, and $\alpha g(0)$ is the lowest possible wage for skilled workers in the urban area. Thus, Proposition 1 implies that no matter how strong the agglomeration effect in the urban area, if the city cannot absorb all the unskilled workers from the rural area, which is typically the case in the developing world, then everyone will end up being poor. In other words, without any restrictions on, or barriers to, rural-to-urban migration, the agglomeration effect can result in high production efficiency only in a partial equilibrium framework. However, our general-equilibrium framework, as simple as it is, shows that the efficiency effect of urban agglomeration can be completely diluted by the falling level of average human capital in the city resulting from the rural-to-urban migration of unskilled workers. As noted in the introduction, this inefficient outcome stems from the misallocation of the externality effect of human capital between the rural and urban areas and from the inability of a free market regime to yield efficient outcomes when the externality effect of the average level of human capital is important in production. This result provides an explanation for the empirical observation that urbanization per se may not lead to efficiency gains despite the benefit of agglomeration economies in the cities of the developing countries (Gillis et al., 1996; Todaro, 2000; Henderson, 2003).

3. Restricting rural-to-urban migration can improve social welfare

The market failure identified in Section 2 suggests that the government should assume a role in managing the rural-to-urban movement of labor. In this section, we will demonstrate that by restricting rural-to-urban migration in various ways, the government can improve social welfare.

3.1. Restrictions placed on the migration of the unskilled

In this section, we consider only a restriction on the migration of unskilled workers. In such a case, all the skilled workers will be in the city. We assume that the government chooses m^1 to maximize the output of the entire economy; that is,

$$[l^c + \alpha(s^c + s^r) + m^1] f \left[l^c + \alpha(s^c + s^r) + m^1, \frac{s^c + s^r}{l^c + s^c + s^r + m^1} \right] + (l^r - m^1)g(0). \quad (9)$$

Since the objective function in (9) is continuous with respect to m^1 , which in turn belongs to the compact set $[0, l^r]$, an optimal solution of m^1 must exist. Let the optimal solution to (9) be denoted by m^* . From the assumption in (8), we know that $m^* < l^r$. Then, if the government can restrict migration by allowing only m^* unskilled rural workers into the

city, then the urban wage will increase.⁹ As to the rural unskilled workers, their wage rate will still be at the level of $g(0)$, which is their equilibrium wage rate without any restrictions on rural-to-urban migration. Thus, a restriction of rural-to-urban migration can lead to a Pareto improvement.

In summary, we thus have the following proposition.

Proposition 2. *The restriction of rural-to-urban migration of unskilled workers can result in a Pareto improvement.*

Proposition 2 implies that from the perspective of production efficiency in the economy at large, absent restrictions on rural-to-urban migration there are too many unskilled migrants. The Proposition provides a rationale for the policies of restricting rural-to-urban migration of unskilled workers that have been implemented by a number of developing countries including Indonesia, the Philippines, Tanzania, and in particular, China.¹⁰

In China, a policy called the Hukou system is in effect. Essentially, this is a household registration system. In a way, Hukou is akin to a citizenship in the context of international migration: it entitles an individual to free or subsidized housing, medical care, children's education, and other social benefits, but only in a certain specific location. Hukou is an effective tool for restricting rural-to-urban migration, particularly of unskilled workers (see, for example, Chan and Zhang, 1999; Wu and Treiman, 2004; Au and Henderson, 2006). In particular, it is costly for rural individuals to migrate to the cities. For example, Au and Henderson (2006, pp. 352–353) write, "Permanently leaving a village means abandoning ownership claims without compensation to agricultural land that one's family may have farmed for decades and to the profits of local rural industries which are distributed in-kind, as for example with township housing. . . . Migrants may still have to pay taxes to their rural home village for services they don't consume and on land left fallow. . . . There is a license fee to work outside the home township paid to the township that can be equivalent to several months' wages. At the destination there can be fees for city management, for being a "foreign" worker, for city construction, for crime fighting, for temporary residence, and even for family planning if the migrant is female."

Proposition 2 suggests an economic rationale for the Hukou system in China. Massive rural-to-urban migration will significantly drive down the average human capital in the cities, although it enhances the agglomeration economies of the cities. Without any restrictions on rural-to-urban migration, the net effect is negative and urban productivity will be driven to a low level. The Hukou system may not though necessarily lead to an optimal production efficiency in the cities, yet it enhances urban production efficiency by more than free rural-to-urban migration would. Thus, although the Hukou system was largely an outgrowth of political considerations, our model shows that it may have significant and surprisingly positive economic benefits.

3.2. Restrictions placed on the migration of the skilled

This subsection shows that if some skilled workers can be held back in the rural area (albeit with some financial compensation from the government), then, obviously, the proportion of skilled workers in the rural area will be greater than zero, which implies that there will be an overall increase in the wage rate and, hence, a possible enhancement of social welfare.

To establish this possibility rigorously, we denote the proportion of the skilled workers that are kept in the rural area by x . Then, with free migration of rural unskilled workers, the wage rate for unskilled and skilled workers in the countryside will be, respectively, $g(x)$ and $\beta g(x)$ in equilibrium; the wage rate for unskilled and skilled workers in the city will be, respectively, $g(x)$ and $\alpha g(x)$ in equilibrium. Thus, total output (in the economy at large) will be

$$[l^c + l^r + \alpha(s^c + m^s) + \beta(s^r - m^s)]g(x). \quad (10)$$

⁹ Intuitively, we may consider the following scenario: suppose that prior to the incidence of rural-to-urban migration, the urban wage (for unskilled workers) is higher than the rural wage. If the government can impose a restriction that only skilled workers can migrate from the countryside to the city, then the urban wage will *increase* because the migration of the skilled workers into the city will enhance the agglomeration effect *and* raise the average level of human capital in the city.

¹⁰ However, a policy of restricting rural-to-urban migration of only unskilled individuals will result in a large urban-rural gap in earnings. Apparently, efficiency and equality do not move in tandem.

It is subject to

$$x = \frac{s^r - m^s}{s^r - m^s + l^r - m^l} \quad (11)$$

and

$$f \left[l^c + \alpha(s^c + m^s) + m^l, \frac{s^c + m^s}{l^c + s^c + m^s + m^l} \right] = g(x). \quad (12)$$

The first order condition of (10) with respect to x is

$$(\alpha - \beta) \frac{d(m^s)}{dx} g(x) + [l^c + l^r + \alpha(s^c + m^s) + \beta(s^r - m^s)] g'(x) \leq 0 \quad (13)$$

with strict equality holding if $x > 0$. In (13), the expression of $d(m^s)/dx$ can be obtained by totally differentiating (11) and (12) with respect to m^s , m^l and x . It is easy to verify that if $g'(0)$ is sufficiently large, then $x=0$ cannot satisfy (13), which means that $x=0$ is not the optimal solution.

The preceding discussion leads then to the following proposition.

Proposition 3. *If $g'(0)$ is sufficiently large, then total output will be larger if some skilled workers work in the rural area.*

Proposition 3 implies that if the cost of tax and transfer that the government incurs is sufficiently small relative to $g'(0)$ which is the marginal benefit of increasing the average level of human capital in the rural area when the level is at zero, then the (benevolent) government will find it beneficial to induce some skilled individuals to work in the rural area by means of financial compensation. Since human capital is increasingly important in the rural areas of many developing countries, when the initial level of average human capital is low, the increase in output resulting from some skilled individuals working in the rural area will be large, which implies that this condition is likely to be satisfied in these countries. Thus, an interesting policy implication is that a Pareto improvement can be had if the government were to subsidize some skilled labor in the rural areas so that they will not migrate to the city or that they will move from the city to the rural areas.¹¹ The presence of a certain number of skilled workers in the rural areas may significantly reduce the rural-to-urban migration of *unskilled* workers.

Proposition 3 further implies that without restrictions on migration, there may be too many *skilled* migrants.¹² The intuition is that the migration of skilled workers from the rural areas is likely to bring in its wake the migration of many more unskilled workers, which will reduce the average human capital of the city, and hence the city's productivity.¹³ The migration of *skilled* workers from the rural areas will reduce rural productivity, which will make the rural areas even less attractive to rural *unskilled* workers. Consequently, the migration of skilled workers can lead to a large number of unskilled workers migrating from the rural areas. The common wisdom in the received literature is that in developing countries, while unskilled rural-to-urban migration may harm the urban economy, skilled rural-to-urban migration will have a positive effect on the urban economy (for example, see McCormick and Wahba). The common wisdom is deficient because it is based on a partial-equilibrium analysis, concentrating on the beneficial impact of skilled rural-to-urban migration for the cities, while ignoring both the harmful impact of this migration for the rural

¹¹ For the United States, an analysis based on data taken from the National Longitudinal Survey of the Labor Market Experience 1979 until 1998 (Gould, 2007) suggests that for "white-collar" workers, human capital gains acquired from working in the city are transferable to the rural areas. There is no evidence that moving to the rural areas wipes out the usefulness of human capital even if acquired in the city. Of course the United States' evidence may be of limited relevance to the setting that we have studied, but it serves to hint that a "downward" movement from the city to the rural area is not a harbinger of inevitable human capital erosion.

¹² In this model, we abstract from the consideration that individuals may have idiosyncratic tastes for living in the countryside or in the city. With this additional consideration, whether the implication holds will depend on whether there is a sufficiently large number of skilled individuals who prefer to live in the countryside due to their idiosyncratic tastes despite the urban-rural skilled wage gap.

¹³ It might be argued that if there are few skilled individuals in the countryside, then the (private) returns to the skilled in the countryside will be high, but this argument is not supported by evidence. In most developing countries, or even developed countries, there is a much larger fraction of skilled workers in the city. However, the private returns (the earnings) of the skilled are much higher in the cities. In fact, it appears that the less skilled there are in the countryside, the larger the urban-rural gap and the less the earnings of both the skilled and the unskilled in the countryside.

areas and the harmful impact for cities of the subsequent unskilled rural-to-urban migration. Our general-equilibrium analysis implies that excessive skilled rural-to-urban migration harms the urban economy as well.

Again, the Hukou system in China provides a good case study of the benefits of restricting rural-to-urban migration in a developing country.¹⁴ A fairly unique feature of the Chinese economy is its township and village enterprises (TVEs), which, as already noted in the Introduction, have played a significant role in China's growth since the early eighties. An important contributing factor for the rapid development of TVEs is the plentiful presence of talented and skilled workers in China's rural areas (Fu and Balasubramanyam). Conversely, the development of rural industries dampens the incentive of skilled workers to migrate from the rural to the urban areas.

4. Additional policy implications

In this section, we analyze the implications of two policies that have often been implemented in many developing countries: (1) subsidy to the rural sector, and (2) increasing educational expenditures. We assume that labor is freely mobile between the rural and the urban areas, that is, that no restrictions are imposed on rural-to-urban migration.

4.1. Subsidy to the rural sector

We denote the subsidy to every worker in the rural sector by λ . Then, from Proposition 1, the net wage rate of every unit of efficiency labor in the rural area will rise from $g(0)$ to $g(0) + \lambda$. We postulate that the subsidy is financed by taxing the workers in the city, and we assume that the tax rate is flat. Then, in equilibrium, the following two conditions must be satisfied.

(1) Urban–rural wage equalization

With free labor mobility and perfectly competitive labor markets, the urban wage rate for the unskilled workers must be equal to the rural wage rate for the unskilled workers so that there is no further rural-to-urban migration. By logic similar to the reasoning of Section 2, all skilled individuals work in the city in equilibrium. Thus, with government intervention of tax and subsidy, the equilibrium condition for the equalization of the (net) wage rate in the rural and urban areas is

$$(1 - \tau)f \left[l^c + \alpha(s^c + s^r) + m^l, \frac{s^c + s^r}{l^c + s^c + s^r + m^l} \right] = g(0) + \lambda, \quad (14)$$

where τ denotes the tax rate (per an efficiency of unit labor in the city).

(2) The government's budget constraint

Since we study a static model, we assume that the government's budget must be balanced: the total subsidy payments to the rural sector are equal to the total tax revenue in the urban sector; that is,

$$\tau[l^c + \alpha(s^c + s^r) + m^l]f \left[l^c + \alpha(s^c + s^r) + m^l, \frac{s^c + s^r}{l^c + s^c + s^r + m^l} \right] = (l^r - m^l)\lambda. \quad (15)$$

We now show that the subsidy to the rural sector can lead to a Pareto improvement for both urban and rural dwellers. The reasoning is straightforward: with the subsidy in place, the net wage rate (that is, the wage rate after the tax and subsidy) for the unskilled workers in both the urban sector and the rural sector in equilibrium will increase from $g(0)$ to $g(0) + \lambda$. Meanwhile, in equilibrium, the net wage rate for the skilled workers (in the urban sector) will increase from $\alpha g(0)$ to $\alpha[g(0) + \lambda]$. Thus, the subsidy to the rural sector can lead to a strict Pareto improvement for both urban and rural dwellers, which unambiguously improves social welfare.

Next, we examine how the subsidy to the rural sector can be *optimally* chosen. We assume that the government chooses m^l to maximize the output of the entire economy after the imposition of taxes and the disbursement of subsidies. With free labor mobility, all the skilled workers will work in the city, by the assumption that the returns to their skills are higher in the city. Thus, the total income of the whole economy (before the government's

¹⁴ The Hukou system imposes much less stringent migration restrictions on skilled workers than on unskilled workers (Wu and Treiman). Still, there are significant barriers even for a skilled worker in rural China to obtain an urban Hukou. Consequently, many skilled workers remain in the rural areas of China.

imposition of taxes and disbursement of subsidies) is given by (9). Note that the total income after the government's redistribution is the same as the total income before the government's intervention. Therefore, the government will choose m^1 to maximize the total output of the whole economy (before the government's tax and subsidy), that is, it will choose $m^1 = m^*$ (recall the analysis in Section 3.1). Then, by inserting $m^1 = m^*$ into (14) and (15) and solving these two simultaneous equations, we obtain the optimal subsidy (λ^*) and the optimal tax (τ^*) as follows:

$$\tau^* = \left[1 - \frac{g(0)}{f} \right] \frac{l^r - m^*}{l^c + \alpha(s^c + s^r) + l^r} \quad (16)$$

and

$$\lambda^* = [f - g(0)] \frac{l^c + \alpha(s^c + s^r) + m^*}{l^c + \alpha(s^c + s^r) + l^r}. \quad (17)$$

It is easy to ascertain that the tax rate in the city, τ^* , is between zero and one, and that the amount of subsidy per rural unskilled worker, λ^* , is between zero and $f - g(0)$. In summary, we have the following proposition.

Proposition 4.

- (1) *The subsidy to the rural sector can lead to a Pareto improvement for both the urban and rural dwellers, which unambiguously improves social welfare.*
- (2) *The optimal tax and the optimal subsidy given, respectively, by (16) and (17), yield an equilibrium number of unskilled migrants that maximizes the output of the whole economy.*

The underlying logic of Proposition 4 is similar to that of Section 3. The subsidy to the rural sector effectively restricts rural-to-urban migration, which in turn mitigates the inefficient allocation of the average level of human capital between the rural and urban areas.

4.2. *Governmental expenditures on education*

In this subsection, we denote the total *number* of skilled, which is now a variable, by s . In this new setting we denote by l^u the initial number of unskilled workers in the city (before migration occurs), and by m the number of the unskilled rural-to-urban migrants.

We consider the realistic, developing world situation that there will still be unskilled in the rural area, no matter how much the government spends on education (up to a realistic amount). Then, with free labor mobility in equilibrium, the following condition for the urban and rural wage equalization must obtain:

$$f \left(l^u + \alpha s + m, \frac{s}{l^u + s + m} \right) = g(0). \quad (18)$$

Thus, from (18), we can see that with free labor mobility, an increase in educational expenditures will not increase the wage rates for the skilled workers and the unskilled workers in both the rural and the urban areas. Namely, no matter how large s is (up to a certain level so that there are still unskilled workers in the rural area), the wage rate for the unskilled workers and the skilled workers will remain constant at the levels of $g(0)$ and $\alpha g(0)$, respectively.

Now, we denote the fraction of skilled workers in the city by R , namely

$$R \equiv \frac{s}{l^u + s + m}. \quad (19)$$

It implies that

$$l^u + \alpha s + m = \frac{s}{R} + (\alpha - 1)s. \quad (20)$$

Inserting (19) and (20) into (18), we get

$$f \left[\frac{s}{R} + (\alpha - 1)s, R \right] = g(0). \quad (21)$$

Totally differentiating (21) with respect to s and R , we get

$$\frac{dR}{ds} = \frac{[R + (\alpha - 1)R^2]f_1}{sf_1 - R^2f_2}. \quad (22)$$

Recall that $\alpha > 1$. Then, from (22), we can see that $dR/ds < 0$; that is, an increase in the total number of skilled workers will result in a lower fraction of skilled workers in the city if and only if

$$sf_1 < R^2f_2 \quad (23)$$

namely, if and only if

$$(l^u + s + m)^2 f_1 < sf_2. \quad (24)$$

In summary, we have the following proposition.

Proposition 5.

- (1) *With free labor mobility, an increase in governmental expenditures on education will not increase the wage rates for skilled or unskilled workers in both the rural and urban areas.*
- (2) *An increase in the total number of skilled workers will result in a lower fraction of skilled workers in the city if and only if $(l^u + s + m)^2 f_1 < sf_2$.*

The model shows that without restrictions of (deterrence to) rural-to-urban migration, a government's effort to increase educational expenditures and thereby the number of skilled workers will not increase the wage rates in the rural and urban areas as long as a large number of unskilled workers remain in the rural area in equilibrium. Thus, as the number of skilled workers (in the city) increases, the average human capital in the city may *decrease*. Note that the intuition of part (2) of Proposition 5 is that an increase in s may result in a much greater increase in m , which reduces R . If the economic prosperity of a city depends on its average human capital, as emphasized, for example, by Lucas (2001), then a rather surprising outcome is that larger expenditures on education will result in a lower level of average human capital in the city and the same level of average human capital in the countryside, which will thus reduce the economic prosperity of the entire economy. Therefore, the proposition implies that increasing educational expenditures alone in a developing country may yield an inefficient outcome. In other words, in a developing country where unskilled workers far outnumber skilled workers, increasing the human capital stock will not in and by itself reverse the inefficient allocation of the average human capital between the rural and urban areas that is caused, in turn, by unrestricted rural-to-urban migration. Consequently, increasing the human capital stock alone may not result in the country's workers experiencing an increase in their wage rates.

An additional comment is called for. Suppose that the government can freely decide on its educational expenditures in the rural and urban areas. Then, with free rural-to-urban migration, the government will incur educational expenditures only in the city if the educational system in the city is more efficient than in the countryside. This is so because with free labor mobility, all the skilled workers will end up in the city. With a restriction on rural-to-urban migration, however, the government will have a stronger incentive to spend on rural education since it can increase rural wages and reduce the migration of the unskilled workers to the city, which, as analyzed in Section 3, will, in turn increase everyone's wage rate.

5. Robustness of the model

The analysis in the preceding sections yields several new results, some of which are in sharp contrast to the received literature. Since our analysis is based on a simple foundation, one may wonder whether our results depend critically on our underlying assumptions. In response to this possible concern we discuss in this section the robustness of our results. We show that the fundamental assumption of our model is that there is an abundant supply of unskilled labor in the rural areas, an assumption which is in line with the prevailing reality in most developing countries.¹⁵ We argue

¹⁵ In fact, this assumption is similar to Ricardo's (1817) argument that at the time of early industrialization in many of the now developed countries, there was always a "labor surplus" in the rural areas.

that the results obtained thus far will qualitatively hold even if our model is extended. We consider six extensions: (1) capital is yet another factor of production, (2) there are several cities, (3) individuals may have idiosyncratic tastes for living in the countryside, (4) there is a congestion cost in the city, (5) the marginal productivity of unskilled labor in the rural sector is diminishing and (6) housing prices and rents in the urban areas are higher than in the rural areas.

In our model, we have assumed that labor is the only factor of production and that skilled labor and unskilled labor are substitutes in production. If, instead, we were to assume that capital too is a factor of production and that skilled and unskilled labor are complements in production, then, *ceteris paribus*, the returns to capital and skilled labor in the city will increase with increased rural-to-urban migration of unskilled labor. However, the returns to both capital and skilled labor also depend on the average level of human capital in the city, which declines with the increase in the inflow of unskilled labor. The net returns to capital and skilled labor in response to an increase in the rural-to-urban migration of unskilled labor will depend then on the magnitude of the negative impact on productivity of a declining average level of human capital. In other words, the greater the impact of the average level of human capital on productivity, the more likely the net returns to capital and to skilled labor will decline with an increase in the rural-to-urban migration of unskilled labor.

In historical times, such as the period of the Industrial Revolution, production was not intensive in sophisticated knowledge, which suggests that the agglomeration effect in itself resulted in high productivity, with urbanization positioning the economy on a growth path (Goodfriend and McDermott, 1995).¹⁶ In that setting, rural-to-urban migration of unskilled labor benefited both capitalists and skilled labor (such as engineers and entrepreneurs). However, in the current era of knowledge-based production, the average level of human capital is a vital determinant of a city's productivity. In this environment, unrestricted rural-to-urban migration is likely to lead to a significant reduction in the returns to all factors of production.

Consider next a setting of multiple cities. In such a setting, since the production functions in different cities are all likely to be non-linear, the general-equilibrium framework of this new setting is likely to yield multiple equilibria. However, as long as there is a sufficiently abundant supply of unskilled labor in the rural areas, the result will continue to hold that in a developing country, the agglomeration effect does not bring about high production efficiency in a general-equilibrium framework if labor mobility is free. The reason is that if the wage of the unskilled workers in any city is higher than the wage of the rural unskilled workers, then there will be an inflow of unskilled migrants from the rural area into that city that will result in a steadily declining level of average human capital in that city. In equilibrium there will be a complete equalization of the wage rate of the unskilled workers in the rural area and in that city, and the efficiency effect of urban agglomeration in that city will be completely diluted by the unrestricted rural-to-urban migration of unskilled workers. By a similar logic, the policy implications of the model will also continue to hold.

To concentrate on essentials, our model does not incorporate the possibility that some individuals may have idiosyncratic tastes for living in the countryside, which could imply that some skilled individuals may choose to accept a lower rural wage in order to live in the countryside. In this case, not all the skilled workers will end up in the cities. However, it is reasonable to argue that particularly in a poor economy, for most people this kind of idiosyncratic taste is not strong in comparison with the lure of a higher income. Thus, the essential implication of our model, that without government intervention too many of both the skilled workers and the unskilled workers are in the cities, will materially continue to hold.

Also, in the received literature it is nearly always postulated that the limiting factor on migration to the cities is congestion (or commuting costs and pollution). This perspective can be conveniently incorporated into our model upon a slight modification of an assumption in (2): instead of $f_1(n, h) > 0$ for all n , we could assume that $f_1(n, h) > 0$ if n is below a certain level, and that $f_1(n, h) < 0$ if n is above a certain level. Clearly, such a modification will not qualitatively change any of the model's results. As a matter of fact, this consideration implies that there is a negative externality effect of rural-to-urban migration, which in turn implies that restricting rural-to-urban migration is even more desirable. In other words, the modification would only reinforce the ramifications of our model.

Next, it might be argued that the marginal productivity of unskilled labor in the rural sector diminishes and, hence, that additional rural-to-urban migration may increase rural welfare. However, since there are at least two factors of production in the rural sector viz. skilled labor and unskilled labor, these two factors are likely to be complementary in production, and thus, an increase in skilled labor in the rural sector will increase the marginal productivity of unskilled

¹⁶ In those times, manufacturing existed only in cities, and the rural areas produced only agricultural goods.

labor in that sector. Consequently, restricting the rural-to-urban migration of skilled labor will increase the marginal productivity of unskilled labor in the rural sector and thereby could well increase welfare in that sector. In other words, a more efficient way of allocating labor might be to induce some skilled labor to work in the rural sector, rather than encouraging additional rural-to-urban migration of unskilled labor.

Finally, it could be argued that agglomeration in the cities will bring in its wake high housing prices that could drive unskilled labor out of the cities. However, this argument is inconsistent with the observed reality in most developing countries. Large segments of the cities of the developing world consist of shanty towns and poor neighborhoods in which the price of housing is very low, even though the price of housing in other parts of the cities can be as high as that in the cities of the developed world (cf. Todaro). At least two explanations account for this. First, the presence of a large number of unskilled workers in a certain part of a city can well result in a low housing price in *that* part of the city, for example, due to concerns about crime and pollution. Second, a great many of the unskilled workers who migrate to the cities in the developing world leave their families behind. These migrants typically send much (often most) of their (meager) earnings to their families in the rural areas, and their demand for housing in the cities is usually quite modest. The shanty towns and poor neighborhoods in the cities evolve so as to offer low-quality, cheap, and small housing units to cater for this demand.

In conclusion, the results obtained in our simple model appear to be robust to an array of possible extensions, as long as the assumption that there is a sufficiently abundant supply of unskilled labor in the rural areas holds.¹⁷ Since this assumption is in line with the prevailing reality of most developing countries but may not be reflective of the conditions that obtain in the developed world, we re-emphasize that our model applies only to developing countries.

6. Conclusions

Somewhat surprisingly, the received theoretical literature hardly attends to the subject of rural-to-urban migration in developing countries *in modern times*. As noted in the Introduction, this neglect appears to stem from two interrelated misconceptions: that rural areas produce only agricultural goods, and that the efficient production of manufactured goods can take place only in (large) cities.¹⁸ Presumably, it is because of these misperceptions that the role of human capital in the production of the rural sector is largely ignored in the received theoretical literature.

We have sought to set up a new framework that bridges the gap. We postulate that a city's productivity is determined by its average level of human capital and by the size of its labor force. The productivity of the rural area is determined by the average level of human capital there. Our analysis yields a rather surprising implication: unrestricted rural-to-urban migration reduces the average income of both rural and urban dwellers in equilibrium. With free inter-area labor mobility, rural-to-urban migration will come to a halt when the urban and rural wages for unskilled workers are equalized: the urban wage will fall continuously with the in-migration of unskilled workers, which reduces the average level of human capital in the city. Furthermore, since the wages of skilled and unskilled workers in the urban area are affected by common productivity factors, the wage of the skilled workers will be driven to a low level by the unrestricted rural-to-urban migration. Thus, our analysis explains the *negative* consequences of rural-to-urban migration in the developing world. Moreover, our analysis yields several interesting policy insights: the analysis reveals that measures aimed at curtailing rural-to-urban migration by both unskilled and skilled workers can potentially lead to a Pareto improvement for both the urban and rural dwellers, and it shows that without restrictions on rural-to-urban migration, increasing educational expenditures alone may not increase the wage rates in the rural or urban areas.

Models of rural-to-urban migration are at the heart of theories of economic development and growth. Ricardo argued that the urban industrial sector can draw away surplus rural labor without causing a rise in wages in either the rural or urban areas. Ricardo's insight was expanded in numerous subsequent writings, including the Nobel-Prize winning treatise of Lewis (1955) and the seminal contribution of Fei and Ranis (1964). This body of work explains nicely the historical demographic transition across regions as well as the economic development of the currently developed countries. Recently, Lucas (2004) has revitalized Ricardo's original insight and developed it further, noting that rural-

¹⁷ Recall that this assumption is represented by inequality (8) when there are no restrictions on labor mobility, and when there are no government interventions.

¹⁸ In fact, manufacturing activities have become more decentralized even in developed countries such as the USA (Desmet and Fafchamps, 2006).

to-urban migration is essentially a process of the transfer of labor from a traditional, land-intensive technology to a human capital-intensive technology.

Despite its prominence in the development economics literature, the current body of research runs into great difficulties in explaining the consequences of the massive rural-to-urban migration flows that have occurred in a great many developing countries in the past few decades. Our theory explains the different impacts of rural-to-urban migration in the past versus nowadays. We posit that the production efficiency of a city depends not only on the size of its labor force (the agglomeration effect), but also on the average level of human capital of its labor force. In a general-equilibrium framework, we demonstrate that unrestricted rural-to-urban migration leads to inefficiency when the average level of human capital plays a significant role in productivity, which might not have been the case in historical times. In the past, production was not knowledge intensive, which suggests that it was the agglomeration effect in and by itself that resulted in high productivity, with urbanization placing the economy on a solid growth path. In the current era of knowledge-intensive production, the average level of human capital is a vital factor in the productivity of both the urban and rural areas. In this setting, unrestricted rural-to-urban migration in developing countries leads to significant negative outcomes for all individuals, in cities and countryside alike.

Acknowledgements

We are indebted to two anonymous referees for gratifying evaluations, helpful comments, and insightful suggestions.

References

- Abdel-Rahman, H.M., Anas, A., 2004. Theories of systems of cities. In: Henderson, J.V., Thisse, J.F. (Eds.), *Handbook of Urban and Regional Economics*. Elsevier, Amsterdam, pp. 2294–2339.
- Acemoglu, D., 1996. A microfoundation for social increasing returns in human capital accumulation. *Quarterly Journal of Economics* 111, 779–804.
- Au, C.-C., Henderson, J.V., 2006. How migration restrictions limit agglomeration and productivity in China. *Journal of Development Economics* 80, 350–388.
- Bertinelli, L., Black, D., 2004. Urbanization and growth. *Journal of Urban Economics* 56, 80–96.
- Black, D., Henderson, J.V., 1999a. Spatial evolution of population and industry in the United States. *American Economic Review* 89, 321–327.
- Black, D., Henderson, J.V., 1999b. A theory of urban growth. *Journal of Political Economy* 107, 252–284.
- Chan, K.W., Zhang, L., 1999. The Hukou system and rural–urban migration in China: processes and changes. *China Quarterly* 160, 818–855.
- Desmet, K., Fafchamps, M., 2006. Employment concentration across US counties. *Regional Science and Urban Economics* 36, 482–509.
- Fei, J.C.H., Ranis, G., 1964. *Development of the Labor Surplus Economy: Theory and Policy*. Irwin, Homewood, IL.
- Fu, X., Balasubramanyam, V.N., 2003. Township and village enterprises in China. *Journal of Development Studies* 39, 27–46.
- Gillis, M., Perkins, D.H., Roemer, M., Snodgrass, D.R., 1996. *Economics of Development*. W.W. Norton, New York.
- Glaeser E.L., Saiz, A., 2004. The rise of the skilled city. *Brookings-Wharton Papers on Urban Affairs*, pp. 47–105.
- Gould, E.D., 2007. Cities, workers, and wages: a structural analysis of the urban wage premium. *Review of Economic Studies* 74, 477–506.
- Goodfriend, M., McDermott, J., 1995. Early development. *American Economic Review* 85, 116–133.
- Harris, J.R., Todaro, M.P., 1970. Migration, unemployment and development: a two-sector analysis. *American Economic Review* 60, 126–142.
- Henderson, J.V., 2003. The urbanization process and economic growth: the so-what question. *Journal of Economic Growth* 8, 47–71.
- Jonasson, E., 2007. Non-agricultural employment as rural poverty exit path? Evidence from Peru. Mimeo, Lund University.
- Lanjouw, J.O., Lanjouw, P., 2001. The rural non-farm sector: issues and evidence from developing countries. *Agricultural Economics* 26, 1–23.
- Lewis, W.A., 1955. *The Theory of Economic Growth*. Allen & Unwin, London.
- Liu, Z., 2007. The external returns to education: evidence from Chinese cities. *Journal of Urban Economics* 61, 542–564.
- Lucas Jr., R.E., 2001. *Lectures on Economic Growth*. Harvard University Press, Cambridge, MA.
- Lucas Jr., R.E., 2004. Life earnings and rural–urban migration. *Journal of Political Economy* 112, S29–S59.
- McCormick, B., Wahba, J., 2005. Why do the young and educated in LDCs concentrate in large cities? Evidence from migration data. *Economica* 72, 39–67.
- Moretti, E., 2004. Workers' education, spillovers, and productivity: evidence from plant-level production functions. *American Economic Review* 94, 656–690.
- Naughton, B., 2007. *The Chinese Economy: Transitions and Growth*. MIT Press, Cambridge, MA.
- Ray, D., 1998. *Development Economics*. Princeton University Press, Princeton, NJ.
- Reardon, T., Berdegue, J., Escobar, G., 2001. Rural nonfarm employment and incomes in Latin America: overview and policy implications. *World Development* 29, 395–409.
- Ricardo, D., 1817. *The Principles of Political Economy and Taxation*. Dent, London [1911 Reprints].
- Schultz, T.P., 2004. Human resources in China: the birth quota, returns to schooling, and migration. *Pacific Economic Review* 9, 245–267.
- Shukla, V., Stark, O., 1990. Policy comparisons with an agglomeration effects-augmented dual economy model. *Journal of Urban Economics* 27, 1–15.

- Taylor, J.E., Yunez-Naude, A., 2000. The returns from schooling in a diversified rural economy. *American Journal of Agricultural Economics* 82, 287–297.
- Todaro, M.P., 2000. *Economic Development*. Addison Wesley, Reading, MA.
- Wu, X., Treiman, D.J., 2004. The household registration system and social stratification in China: 1955–1996. *Demography* 41, 363–384.
- Yang, D.T., 2004. Education and allocative efficiency: household income growth during rural reforms in China. *Journal of Development Economics* 74, 137–162.
- Yang, D.T., An, M.Y., 2002. Human capital, entrepreneurship, and farm household earnings. *Journal of Development Economics* 68, 65–88.