

**Cities, Towns, and Poverty:
Migration Equilibrium and Income Distribution in a
Todaro-type Model with Multiple Destinations***

Luc Christiaensen, The World Bank

Joachim De Weerd, University of Antwerp and KU Leuven

Ravi Kanbur, Cornell University

12 April, 2017

Contents

1. Introduction
 2. Model and Equilibrium
 3. Poverty Gradients
 4. Job Creation and Poverty Reduction
 5. An Empirical Illustration
 6. Heterogeneous Migrants and Jobs
 7. Conclusion
- References

Abstract

Should public investment be targeted to big cities or to small towns, if the objective is to minimize national poverty? To answer this policy question we extend the basic Todaro-type model of rural-urban migration to the case of migration from rural areas to two potential destinations, secondary town and big city. We first derive conditions under which a poverty gradient from rural to town to city will exist as an equilibrium phenomenon. We then address the policy question and show how the answer depends on the migration response, where the poverty line lies relative to incomes in the three locations, and at times also the poverty index itself. In particular, we develop sufficient statistics for the policy decisions based on these income parameters and illustrate the empirical remit of the model with long running panel data from Kagera, Tanzania. Further, we show that the structure of the sufficient statistics is maintained in the case where the model is generalized to introduce heterogeneous workers and jobs. Overall, the findings confirm that, given migration responses, national poverty outcomes are not immune to whether urban employment generation takes place in the towns or the city.

Key words: Secondary Towns versus Big Cities, Poverty Reduction, Poverty Gradient, Todaro Model, Migration Equilibrium, Equilibrium Income Distribution.

JEL Codes: O18, O41, I3, J61

* The authors acknowledge support from the World Bank. The findings, interpretations, and conclusions expressed in this paper are entirely those of the authors. They do not necessarily represent the views of the International Bank for Reconstruction and Development/World Bank and its affiliated organizations, or those of the Executive Directors of the World Bank or the governments they represent.

1. Introduction

Should public investment be targeted to big cities or to small towns, if the objective is to minimize national poverty? This perennial policy question has grown in importance as the debate on “megacities versus secondary towns” has intensified (Christiaensen and Kanbur, 2016). A first cut answer to this question might look at which location had greater poverty. If there is a declining poverty gradient from towns to cities, an initial answer might be to invest in secondary towns for poverty reduction.

But it should be clear that a more considered answer to the question depends on at least three factors. First, what will be the first round impact, on income distribution and poverty, of public investment in either location? Here the poverty gradient may give us an initial clue. But, second, how will this first round impact change the attractiveness of migration to cities and towns from rural areas? And third, how will this migration in turn affect income distribution in cities and towns, and thence migration and distribution in subsequent rounds? The key point is that any public investment will induce a reallocation of population, and any assessment of the impact of public investment on national income distribution and thus national poverty will have to model the new migration equilibrium. Indeed, the poverty gradient will itself be an equilibrium phenomenon, taking into account migration responses to income differentials.

There are thus many moving parts in attempting to provide an answer to the basic question. Our focus in this paper will be on migration equilibrium, population reallocation and the national income distribution, for given efficiency of investment in big cities versus small towns. Migration equilibrium has of course been a staple of development theory since the seminal papers of Todaro (1969) and Harris and Todaro (1970). In these papers there are two sectors (“rural”, naturally thought of as the location of origin in migration, and “urban”, thought of as the destination location). Location choice is determined by economic prospects in the two sectors and equilibrium is attained when there are no private incentives to migrate from one location to another.

As is now well known, in the standard formulation of the Todaro/Harris-Todaro model the rural prospect is a certain income while the urban prospect consists of a higher than rural income (“modern sector employment”) with a probability, and a lower than rural income (“unemployment” or “informality”) otherwise. Equilibrium in a model with identical agents is described by indifference to location, and is attained by endogenous adjustment of the urban prospect (changes in the unemployment rate) as population relocates. With fixed incomes in the rural sector, in modern employment and in unemployment, and a fixed number of jobs in the urban modern sector, the equilibrium outcomes are determined by these incomes and the number of jobs, and by the degree of risk aversion (the base assumption is usually one of risk neutrality).

The basic model has endured as a platform for analysis and debate despite its simplified nature. From this base, one strand of the literature has progressively complicated the urban prospects, for example by introducing different features of an informal sector including endogeneity of income. Further complications like skill acquisition while waiting for a modern sector job, which brings into play dynamic considerations, have also been introduced. A recent

example of this development, which also has references to previous work in this tradition, is Basu, Chau, Fields and Kanbur (2016).

How about income distribution? With the basic set up, income distribution is determined by the migration equilibrium, which gives the share of population at each of the three income levels—rural, modern employment and unemployment/informal. Comparative statics on this income distribution has been a staple in the literature. For example, in an early contribution, Anand and Kanbur (1985) trace out the implications of increasing the urban modern wage or the number of urban modern jobs. More recently, papers such as those by Temple (2005) and Temple and Ying (2014) continue this tradition of tracing out the evolution of income distribution in a two-sector model of population reallocation between agriculture and non-agriculture under different economic environments.

There is thus a strong tradition of analyzing migration equilibrium and income distribution in the setting of two sector Todaro/Harris-Todaro type models. However, the key is that these are two sector models—typically a rural sector and an urban sector. The question of big cities versus small towns simply cannot be posed in such models because there is only a single urban sector encompassing the two. In order to answer the question posed at the start of this paper, on whether poverty reduction is best served by public investment in cities or in towns, we clearly need a model in which the urban sector is so disaggregated. The migration equilibrium will then have to encompass population allocation in three locations—one rural and two urban. And the national income distribution will now be more complex, taking into account income and respective population shares at each of the income levels in the three locations.

Multiple destinations are also a staple of the rural-urban migration literature. For example, Fafchamps and Shilpi (2013) is only one of the recent additions to the literature, looking at the choice of migration destination in Nepal using two rounds of the Living Standards Survey. But, like that paper, nearly all of this strand of the literature is empirical and econometric in nature, and there is no focus on the distributional implications of a multiple destinations model. The missing piece of the puzzle, then, is indeed a model of migration equilibrium and income distribution with multiple destinations.

The object of this paper is, first of all, to develop a Todaro-type model of migration equilibrium and income distribution where rural migrants face two urban destinations—a big city and a secondary town, each with its own stylized prospects. This is done in Section 2. Section 3 examines the properties of such an equilibrium focusing on the nature of the distributional differences between the three locations—specifically on whether there is a “poverty gradient” from rural to towns to cities, as has been found in the empirical literature (Ferre, Ferreira and Lanjouw, 2012). Section 4 turns to the motivating question of this paper—should public investment be targeted to towns or cities, if the objective is poverty reduction? This is analyzed in terms of the impact of job creation in the town or in the city. Conditions under which one or the other policy stance dominates are then derived. Section 5 then presents an empirical illustration of our arguments with long running panel data from Kagera, Tanzania. Section 6 extends the simple, identical agents, model of the previous sections to the case where agents are heterogeneous. It is shown that heterogeneity complicates the picture but the basic economic forces identified in the previous analyses still hold sway. Section 7 concludes the paper.

2. Model and Equilibrium

We consider an extension of the basic Todaro (1969) model by introducing two potential destinations for out-migration from the rural area – small town and city. Rural, small town and city are identified by subscripts r , s and c , respectively. In each destination the migrant faces a prospect of getting a high paying modern sector job with wage w (w_s , w_c), or ending up unemployed with a low paying informal sector income w_o (w_{os} , w_{oc}). The probability of securing modern sector job is e (e_s , e_c). There is a single rural sector income which is certain, and denoted w_r . Incomes in the model are exogenously given.

Total population is \bar{N} , and the populations of the three sectors are denoted N_r , N_s and N_c . Modern sector employment is denoted by E (E_s , E_c) and informal sector employment (or unemployment) is denoted U (U_s , U_c). Throughout we treat the words “unemployment” and “informal employment” as synonymous. Thus

$$\bar{N} = N_r + N_s + N_c \quad (1)$$

$$N_s = E_s + U_s \quad (2)$$

$$N_c = E_c + U_c \quad (3)$$

Further, as in the classic Todaro (1969) formulation, the probability of getting a modern sector job in a destination is identified with the modern sector employment rate in that destination:

$$e_s = E_s / N_s \quad (4)$$

$$e_c = E_c / N_c \quad (5)$$

The “unemployment rate” is then the fraction of the population in the informal sector.

$$u_s = 1 - e_s \quad (6)$$

$$u_c = 1 - e_c \quad (7)$$

We come now to the migration decisions and migration equilibrium in this model with a single origin and two destinations. The agent in the rural sector faces the choice of staying in the rural sector with income w_r , migrating to the small town for a prospect of w_s with probability e_s and w_{os} with probability $1 - e_s$, or migrating to the city for a prospect of w_c with probability e_c , and w_{oc} with probability $1 - e_c$. In addition, we suppose that there is a “migration cost” for each destination, denoted t_s^* and t_c^* . These can be transportation costs to, and settlement and job search costs in, each of the destinations.¹ We assume that $t_c^* > t_s^*$.

¹ In practice, transport costs are often only a fraction of the annual income gain that could be obtained from moving. Among migrants for Kagera, in northern Tanzania, Christiaensen, De Weerd and Kanbur (2017) show for example that transport costs to towns were only about 22 percent of the average annual income gain from informal employment in towns (w_{os}), and only 14 percent from the average income gain from informal employment in the city (w_{oc}). Schwartz (1973) alerted us to this already early on. In her paper, “Interpreting the effect of distance on migration”, she shows that the adverse effect of distance on migration follows especially from diminishing-information (i.e. rising jobs search costs), and less likely from actual transport (or psychic costs). Migration costs, t_i^* , thus reflect both transport costs to, as well as settlement and jobs search costs in the destinations. This makes it also more likely that the interior equilibrium migration conditions hold empirically (see equations (18) and (19)). As

In a two period framework we assume that the costs of migration are paid in the first period and the benefits come in the second period. Let the discount factor be δ . We make the Todaro assumption that the agent is risk neutral so the migration decision depends on the comparison:

$$w_r + \delta w_r \gtrless w_r - t_s^* + \delta [e_s w_s + (1 - e_s) w_{os}]$$

for the secondary town and the corresponding comparison for the city. Rearranging terms, the comparison becomes:

$$w_r + t_s^*/\delta \gtrless [e_s w_s + (1 - e_s) w_{os}]$$

Let

$$t_s = t_s^*/\delta$$

be the discounted migration cost and let the corresponding term for city be

$$t_c = t_c^*/\delta$$

Then all of the relevant comparisons can be written as follows:

$$w_r \gtrless e_s w_s + (1 - e_s) w_{os} - t_s \quad (8)$$

$$w_r \gtrless e_c w_c + (1 - e_c) w_{oc} - t_c \quad (9)$$

$$e_s w_s + (1 - e_s) w_{os} - t_s \gtrless e_c w_c + (1 - e_c) w_{oc} - t_c \quad (10)$$

It should be clear that we will not get an interior solution to population allocation if any of the inequalities hold strictly – one or two of the three locations will have zero population in that case. For an interior equilibrium, we thus require the core conditions:

$$w_r = e_s w_s + (1 - e_s) w_{os} - t_s = e_c w_c + (1 - e_c) w_{oc} - t_c \quad (11)$$

or

$$w_r + t_c = e_c w_c + (1 - e_c) w_{oc} \quad (12)$$

$$w_r + t_s = e_s w_s + (1 - e_s) w_{os} \quad (13)$$

As in the basic Todaro model, we take $w_r, w_s, w_{os}, t_s, E_s, w_c, w_{oc}, t_c, E_c$ to be fixed. The variables e_s and e_c adjust to satisfy the equilibrium conditions (12) and (13). These together with

the migration costs tend to increase with distance to the destination, they could also be seen as a proxy for distance/proximity. The critical role of distance in understanding migration patterns is also longstanding in the literature (see for example Sjaastad (1962) for an early discussion).

(1) – (7) determine population allocation across the different locations and income levels. In other words, the full equilibrium income distribution in the current period can now be specified.

Solving (12) and (13), in conjunction with (1) – (7), gives us the following equilibrium values:

$$e_s = \frac{w_r + t_s - w_{os}}{w_s - w_{os}} ; \quad u_s = \frac{w_s - w_r - t_s}{w_s - w_{os}} \quad (14)$$

$$N_s = \left[\frac{w_s - w_{os}}{w_r + t_s - w_{os}} \right] E_s ; \quad U_s = \left[\frac{w_s - w_r - t_s}{w_r + t_s - w_{os}} \right] E_s \quad (15)$$

$$e_c = \frac{w_r + t_c - w_{oc}}{w_c - w_{oc}} ; \quad u_c = \frac{w_c - w_r - t_c}{w_c - w_{oc}} \quad (16)$$

$$N_c = \left[\frac{w_c - w_{oc}}{w_r + t_c - w_{oc}} \right] E_c ; \quad U_c = \left[\frac{w_c - w_r - t_c}{w_r + t_c - w_{oc}} \right] E_c \quad (17)$$

The equilibrium rural population is then given by using (15) and (17) in (1).

The equilibrium values in (14) – (17) are thus determined by the exogenous parameters. However, the two employment rates clearly need to lie between zero and one, and this restricts the range of parameter values.

$$1 > e_s > 0 \quad \Leftrightarrow \quad w_s > w_r + t_s > w_{os} \quad (18)$$

$$1 > e_c > 0 \quad \Leftrightarrow \quad w_c > w_r + t_c > w_{oc} \quad (19)$$

The full equilibrium and its income distribution are now pinned down. What are its properties?

3. Poverty Gradients

The equilibrium income distribution in this model consists of five incomes and their associated populations:

$$(w_r, N_r); (w_s, E_s); (w_{os}, U_s); (w_c, E_c); (w_{oc}, U_c) \quad (20)$$

N_r , U_s and U_c are endogenous equilibrium variables derived from equations (1), (15) and (17). The population shares are simply (N_c / \bar{N}) , (E_s / \bar{N}) , (U_s / \bar{N}) , (E_c / \bar{N}) and (U_c / \bar{N}) . We can now investigate the properties of the equilibrium income distribution.

Consider first the observed mean income in the secondary town and city, μ_s and μ_r respectively:

$$\begin{aligned}
\mu_c &= e_c w_c + (1 - e_c) w_{oc} = w_r + t_c \\
&\stackrel{\geq}{\leq} w_r + t_s = e_c w_c + (1 - e_c) w_{oc} = \mu_s \\
&\Leftrightarrow \\
t_c &\stackrel{\geq}{\leq} t_s
\end{aligned} \tag{21}$$

Thus the basic stylized fact, observed almost universally, that mean income in the city exceeds mean income in town, requires that the cost of migration to the city is greater than the cost of migration to the town. From now on we will assume that this is the case.

What about the poverty comparison between rural, town and city? Is there a poverty gradient? This depends on where exactly the poverty line lies relative to the five incomes and how the five incomes lie relative to each other. We assume that the ranking of the five incomes is as follows:

$$w_r < w_{os} < w_{oc} < w_s < w_c \tag{22}$$

In other words, the lowest income is in the rural area and the highest is in the city modern sector. The modern sector income in the secondary town is less than modern sector income in the city, but it is greater than informal income in the city. Informal income in the city is greater than informal income in the secondary town, which is in turn greater than rural income. We can now consider different cases as the poverty line is increased progressively to cover a larger and larger share of the national population.

We focus on poverty as measured by the FGT class of poverty indices P_α . Denote the poverty line by z . Then if incomes are w_i for $i = 1, 2, 3 \dots, n$ individuals in the society,

$$P_\alpha = \frac{1}{n} \sum_{i \text{ s.t. } z > w_i} \left[\frac{z - w_i}{z} \right]^\alpha$$

As is well known, α is the degree of poverty aversion. When $\alpha = 0$, we recover the standard poverty head count ratio. When $\alpha = 1$, we have the poverty gap measure. As α increases beyond 1, the index gives more and more weight to the poorest of the poor.

We now consider poverty across the different groups for progressively higher poverty lines. The first case is where

$$w_r < z < w_{os} < w_{oc} < w_s < w_c \tag{24}$$

In this case there is poverty in the rural sector but not in either of the two urban locations. There is thus a rural-urban poverty gradient but no such gradient between town and city. So consider the second case, where

$$w_r < w_{os} < z < w_{oc} < w_s < w_c \tag{25}$$

Poverty in the three locations – r , s and c , – is given by

$$P_{\alpha,r} = \left[\frac{z - w_r}{z} \right]^\alpha \quad (26)$$

$$\begin{aligned} P_{\alpha,s} &= u_s \left[\frac{z - w_{0s}}{z} \right]^\alpha \\ &= \left[\frac{w_s - w_r - t_s}{w_s - w_{0s}} \right] \left[\frac{z - w_{0s}}{z} \right]^\alpha \end{aligned} \quad (27)$$

$$P_{\alpha,c} = 0 \quad (28)$$

Let us compare $P_{\alpha,r}$ and $P_{\alpha,s}$. When $\alpha = 0$, the head count ratio is 100% in the rural sector but less than 100% in the secondary town, and it is zero in the city. Thus there is a clear declining poverty gradient in the equilibrium, as has been found empirically by Ferrer, Ferreira and Lanjouw (2012). The same holds true for $\alpha = 1$ because not only are a small proportion of the secondary town's population in poverty, the depth of their poverty is smaller. For this same reason, $P_{\alpha,s}$ is lower than $P_{\alpha,r}$ for all α . There is thus a clear poverty gradient from r , to s , to c .

Now consider a third case where both the urban informal sectors are in poverty:

$$w_r < w_{0s} < w_{0c} < z < w_s < w_c \quad (29)$$

In this case:

$$P_{\alpha,r} = \left[\frac{z - w_r}{z} \right]^\alpha \quad (30)$$

$$P_{\alpha,s} = \left[\frac{w_s - w_r - t_s}{w_s - w_{0s}} \right] \left[\frac{z - w_{0s}}{z} \right]^\alpha \quad (31)$$

$$\begin{aligned} P_{\alpha,c} &= u_c \left[\frac{z - w_{0c}}{z} \right]^\alpha \\ &= \left[\frac{w_c - w_r - t_c}{w_c - w_{0c}} \right] \left[\frac{z - w_{0c}}{z} \right]^\alpha \end{aligned} \quad (32)$$

When $\alpha = 0$, the highest poverty is once again in the rural sector. But the comparison between $P_{\alpha,s}$ and $P_{\alpha,c}$ essentially comes down to which location has the higher unemployment rate (or fraction of population in informality) since given the poverty line this is in fact the poverty rate.

The comparison of unemployment rates in this case is ambiguous. An examination of (14) and (16) shows that there are three opposing forces that determine unemployment: (i) higher migration costs (transport, search and settlement) lower unemployment, (ii) a higher gap between the formal wage and the rural wage increases unemployment and (iii) a higher gap between the formal and informal sector lowers unemployment. While higher migration costs and lower informal wages unambiguously lower unemployment in the model, the role of formal income appears ambiguous. On the one hand a high formal wage raises the gap with the rural wage (increasing unemployment) and on the other hand it raises the gap with the informal wage (reducing unemployment). Taking the derivative,

$$\frac{\partial u_c}{\partial w_c} = \frac{\partial \left[\frac{w_c - w_r - t_c}{w_c - w_{oc}} \right]}{\partial w_c} = \frac{(w_r + t_c - w_{oc})}{(w_c - w_{oc})^2},$$

which is positive given (19), it can however be seen that, in the model, higher formal wages come along with larger unemployment rates (and lower formal wages with lower unemployment rates).

Thus a falling P_0 poverty gradient for rural to town to city then becomes more likely when cities have higher migration costs, lower formal sector wage gaps with the rural wage and higher formal-informal wage gaps (or when cities have higher migration costs and lower informal and formal wages, which all discourage city migration, in turn leading to lower city unemployment and poverty). And if the gradient holds for P_0 it will also hold for P_1 and for all α greater than 1, given that $(z - w_{os}) > (z - w_{oc})$. Hence, under certain conditions there will again be a general poverty gradient from rural, through secondary town, to big city.

4. Job Creation and Poverty Reduction

Consider now the motivating policy question of whether the government should invest in small towns or big cities. We specify this in the context of our model as a policy of increasing formal sector jobs in either of the two locations. If the objective is to reduce national poverty, will the impact be greater with an increase of E_s or an increase of E_c ? Of course the final answer to the policy question will depend on the costs of creating a job in the town versus the city. That is not the focus of this paper, but the answers we provide will can be used to benchmark the cost differential which will flip the targeting of public investment in one direction or the other.

Let us start again, with scenario (24), where the poverty line is so low that the only poverty is in rural areas. Then national poverty is simply

$$\begin{aligned} P_\alpha &= \frac{N_r}{\bar{N}} \left[\frac{z - w_r}{z} \right]^\alpha \\ &= \frac{1}{\bar{N}} [\bar{N} - N_s - N_c] \left[\frac{z - w_r}{z} \right]^\alpha \end{aligned}$$

$$= \frac{1}{\bar{N}} \left[\bar{N} - \frac{E_s}{e_s} - \frac{E_c}{e_c} \right] \left[\frac{z - w_r}{z} \right]^\alpha \quad (33)$$

Differentiating (33), we get

$$\begin{aligned} \frac{dP_\alpha}{dE_s} &\leq \frac{dP_\alpha}{dE_c} \\ &\Leftrightarrow \\ 1/e_s &\geq 1/e_c \\ &\Leftrightarrow \\ \frac{w_s - w_{os}}{w_r + t_s - w_{os}} &\geq \frac{w_c - w_{oc}}{w_r + t_c - w_{oc}} \end{aligned} \quad (34)$$

Thus the poverty reduction is greatest ($\frac{dP_\alpha}{dE_s}$ is more negative) where the employment rate is smallest or the informality rate is largest ($1/e_s$ is largest). The intuition behind this is as follows. Each new job of course attracts a migrant. But it also attracts, in equilibrium, additional migrants drawn by the higher probability of getting a job. At the margin, the number of additional migrants attracted is inversely proportional to the employment rate in the location—the lower the employment rate, the more additional migrants are attracted (from (15) and (17): $dN_i/dE_i = 1/e_i$). Even if the additional migrants end up in the informal sector, in the scenario depicted in (24), they are still above the poverty line.

With migration costs to towns typically substantially lower than to cities ($t_s \ll t_c$), favoring migration to towns, the formal employment rate is likely to be lower in towns (unless offset by much higher wage differentials in the cities compared to the towns), as can be seen from comparing (14) and (16). Under such a scenario, poverty reduction from employment generation in towns would be larger than from employment generation in the city.

Scenario (25), where only the rural income and small town informal income are below the poverty line, is a relatively straightforward case where the inherent advantage of the city drives the results. In this case there is no poverty in the city. Any migration induced to the city informal sector from job creation in the city modern sector will reduce P_0 . But the same is not true for the small town, since its informal sector income is below the poverty line. Thus job creation in the city will always dominate for poverty reduction as measured by change in P_0 . The same holds true for P_I since w_{os} is less than z .

Let us now turn to scenario (29) where rural income and both of the urban informal incomes are below the poverty line. In this case:

$$\begin{aligned}
P_\alpha &= \frac{N_r}{N} \left[\frac{z - w_r}{z} \right]^\alpha \\
&+ \frac{U_s}{N} \left[\frac{z - w_{0s}}{z} \right]^\alpha \\
&+ \frac{U_c}{N} \left[\frac{z - w_{0c}}{z} \right]^\alpha \\
&= \left[\frac{z - w_r}{z} \right]^\alpha \\
&- \left(\frac{E_s}{N} \right) \left[\frac{z - w_r}{z} \right]^\alpha - \left(\frac{E_c}{N} \right) \left[\frac{z - w_r}{z} \right]^\alpha \\
&- \left(\frac{U_s}{N} \right) \left\{ \left[\frac{z - w_r}{z} \right]^\alpha - \left[\frac{z - w_{0s}}{z} \right]^\alpha \right\} \\
&- \left(\frac{U_c}{N} \right) \left\{ \left[\frac{z - w_r}{z} \right]^\alpha - \left[\frac{z - w_{0c}}{z} \right]^\alpha \right\} \tag{35}
\end{aligned}$$

This seemingly complicated expression simplifies when $\alpha = 0$, to

$$P_0 = 1 - \left[\frac{E_s}{N} \right] - \left[\frac{E_c}{N} \right] \tag{36}$$

Thus

$$\frac{dP_0}{dE_s} = \frac{dP_0}{dE_c} = - \left[\frac{1}{N} \right] \tag{37}$$

This should be clear intuitively as well. This is a scenario where the only income above the poverty line are modern sector incomes in the two urban locations. Job creation either in the city or in town reduces poverty one for one. There are migration consequences, of course, from rural areas to the informal sectors of the two destination locations. But all of these moves are below the poverty line. Since the index of poverty is the incidence or head count of poverty, these moves below the poverty line do not affect poverty at all. What is left is the effect of a new job in the modern sectors of the two locations, whose consequence for the numbers in poverty is identical. The choice thus depends only on the relative cost of creating jobs in either of the two locations, not on its poverty consequences.

When $\alpha = 1$ the complicated expression in (35) also simplifies, giving us:

$$\begin{aligned}
P_I &= \left[\frac{1}{z\bar{N}} \right] \{ (z - w_r)\bar{N} - X_s E_s - X_c E_c \} \\
X_s &= \frac{(w_{0s} - w_r)(w_s - z) + t_s(z - w_{0s})}{w_r + t_s - w_{0s}} \\
X_c &= \frac{(w_{0c} - w_r)(w_c - z) + t_c(z - w_{0c})}{w_r + t_c - w_{0c}} \tag{38}
\end{aligned}$$

Job creation in either sector reduces poverty, but the impact depends on relative magnitudes of the various income levels and the poverty line.

$$\frac{dP_i}{dE_i} = - \left[\frac{1}{z\bar{N}} \right] X_i \quad ; \quad i = s, c \tag{39}$$

Thus, whether to invest in city or town depends on the relative magnitudes of X_s and X_c –these are the sufficient statistics for whether to create the marginal job in city or town.

The intuition behind expressions (38) and (39) is as follows. Creating a modern sector job in the city, say, reduces the poverty gap of a rural person by $(z - w_r)$. But this job creation also moves a number of people from the rural area to the urban informal sector. For each person so moved, the reduction in poverty gap is given by $(w_{0c} - w_r)$. The total number of persons moved to the city informal sector by creating one job in the modern city sector is given from (17) by $(w_c - w_r - t_c)/(w_r + t_c - w_{0c})$. The reduction in poverty gap from moves to the informal sector is thus given by the product of these two expressions. The total reduction in poverty gap, from moves to the modern and the informal sector in the city as the result of creating one modern sector job is then $(z - w_r) + [(w_c - w_r - t_c)/(w_r + t_c - w_{0c})] (w_{0c} - w_r)$, which is the expression X_c in (38). An analogous argument establishes the intuition for X_s .

The decline in poverty will be greater for modern job generation in the town if $X_s > X_c$, i.e. if

$$\frac{w_s - (w_r + t_s)}{(w_r + t_s) - w_{0s}} (w_{0s} - w_r) > \frac{w_c - (w_r + t_c)}{(w_r + t_c) - w_{0c}} (w_{0c} - w_r) \tag{40}$$

which is more likely if secondary town job generation induces more migration and the larger is the reduction in their poverty gap, i.e. the larger are w_{0s} and w_s ² (and the smaller is t_s ³).

In conclusion, from this extended H-T model with two destinations and explicitly accounting for migration costs, we learn that, in general, to assess whether to invest in cities or towns for poverty reduction, we need to take into account many interactions. Especially, the migration response to job generation as a result of the investment is key. The poverty outcome

² Or the smaller the difference with w_{0c} and w_c respectively (recall from (22): $w_r < w_{0s} < w_{0c} < w_s < w_c$)

³ Compared to t_c –note that $dX_s/dt_s < 0$.

depends further intimately on where the poverty line falls in relation to the income levels in cities and towns. Finally, the poverty index chosen to evaluate the poverty effects can also matter.

For an empirically plausible income gradient scenario ($w_r < w_{os} < w_{oc} < w_s < w_c$) and three different locations of the poverty line within this gradient we have also derived exact conditions under which one or the other policy will dominate. For convenience, Table 1 summarizes the findings, together with the associated poverty gradient. Given the importance of the migration response in determining the empirical outcome, migration costs and thus the proximity of the poor to towns versus cities, which is closely related to the migration cost they face (Schwarz, 1973), will play an important role (as will the difference in the income gaps between formal and informal employment across towns and cities).

To add perspective, globally, a larger share of the population lives closer to secondary towns than to cities. While 15.9 percent live within 1-3 hours from a secondary town (defined as $\leq 500,000$ inhabitants), 9.9 percent live within 1-3 hours from a city ($> 500,000$ inhabitants); another 15.3 percent live in the rural hinterlands (at more than 3 hours from an urban center) (FAO, 2017). In Sub-Saharan Africa, which houses about half of the world's extreme poor (World Bank, 2016), 22 percent of the population live within 1-3 hours from a secondary town, compared with 9 percent living within 1-3 hours from a city. About 1/3 of the rural population lives in the rural hinterland at more than 3 hours from an urban center (FAO, 2017). With 82 percent of Africa's poor living in rural areas (Beegle et al., 2016) and many of these closer to towns than to Africa's cities, it becomes important for global poverty reduction to understand how distance and migration costs interact in practice with the difference in income gaps between formal and informal jobs in those towns and cities.

Table 1: Poverty gradient and poverty reduction from modern job generation under different positions of the poverty line in the income gradient

Scenario	Poverty gradient	Poverty reduction from modern job generation
(24): $w_r < z < w_{os} < w_{oc} < w_s < w_c$	<i>Rural-Urban</i> : $P_{\alpha,r} > P_{\alpha,s} = P_{\alpha,c}$	$\frac{dP_\alpha}{dE_s} < \frac{dP_\alpha}{dE_c}$: greater in town if $\frac{1}{e_s} > \frac{1}{e_c} \Leftrightarrow \frac{w_s - w_{os}}{w_r + t_s - w_{os}} > \frac{w_c - w_{oc}}{w_r + t_c - w_{oc}}$
(25): $w_r < w_{os} < z < w_{oc} < w_s < w_c$	<i>Rural-Town-City</i> : $P_{\alpha,r} > P_{\alpha,s} > P_{\alpha,c}$	$\frac{dP_\alpha}{dE_s} > \frac{dP_\alpha}{dE_c}$: always greater in the city
(29): $w_r < w_{os} < w_{oc} < z < w_s < w_c$	<i>Rural-Town-City</i> : $P_{\alpha,r} > [P_{\alpha,s} > P_{\alpha,c}]$ if $u_s < u_c \Leftrightarrow \frac{w_s - w_r - t_s}{w_s - w_{os}} < \frac{w_c - w_r - t_c}{w_c - w_{oc}}$	$\alpha = 0$: no difference $\alpha = 1$: $\frac{dP_\alpha}{dE_s} < \frac{dP_\alpha}{dE_c}$, greater in town, if $\frac{w_s - (w_r + t_s)}{(w_r + t_s) - w_{os}} (w_{os} - w_r) > \frac{w_c - (w_r + t_c)}{(w_r + t_c) - w_{oc}} (w_{oc} - w_r)$

5. An empirical illustration

The Kagera Health and Development Survey (KHDS) provides an ideal opportunity to bring the model to the data and to begin to explore the empirical relevance of its insights. Under the KHDS, 915 households, sampled to be representative of (rural and urban) Kagera, a region in the north-western part of Tanzania, were first interviewed up to four times during 1991-1994. In 2010 the survey attempted to trace and reinterview all 6,353 original household members, even if they had migrated or split off from the original household. By 2010 the original sample respondents had spread across Tanzania's rural areas, towns and cities. This unique dataset is one of the longest-running panel data sets of individuals in Africa. By 2010, 68 percent (4,336) of the original household members were interviewed, 20 percent had died (1,275) and 12 percent were untraced (739). Out of the 4,336 respondents, comparable consumption data is available for 4,323. With at least one person contacted in 92 percent of the initial households and only 12 percent of individuals not found, survey attrition rates were extremely low, especially considering the long time period of 20 years.

From the sample of 4,323 individuals we drop 313 who were already living in an urban area in 1991/94, leaving 4,010 individuals who originate from rural Kagera. Out of those 1,899 migrated out of their original villages. Twenty years later, in 2010, we find 315 of these individuals had migrated to a city, 557 individuals to a town and 1,027 individuals to another rural area.⁴ Of the 2,111 individuals who did not move, 331 live in an area that got reclassified from rural to urban and we drop them from our analysis to avoid confusion. Virtually everyone in the rural areas is in the informal sector (only 3.5 percent rely on a formal wage job), so this group corresponds closely to the rural population we have in mind in the model. In the towns, 18.9 percent lives in a household that relies on formal wage employment. This rises to 22.2 percent in the cities.⁵

The 2010 income distribution of these individuals, could essentially be thought of as the migration equilibrium outcome of the rural population of Kagera in 1991 across the 3 destinations, the rural area (i.e. no migration or another rural area), the secondary towns and the cities and across the informal and formal sectors within the two latter destinations. It provides an ideal backdrop to give empirical content to the findings of the model under the different scenarios considered (as summarized in Table 1) and explore the model's empirical relevance.

We begin by examining the income gradient observed in the data across the five different income groups considered in the model (informal rural employment, informal town employment,

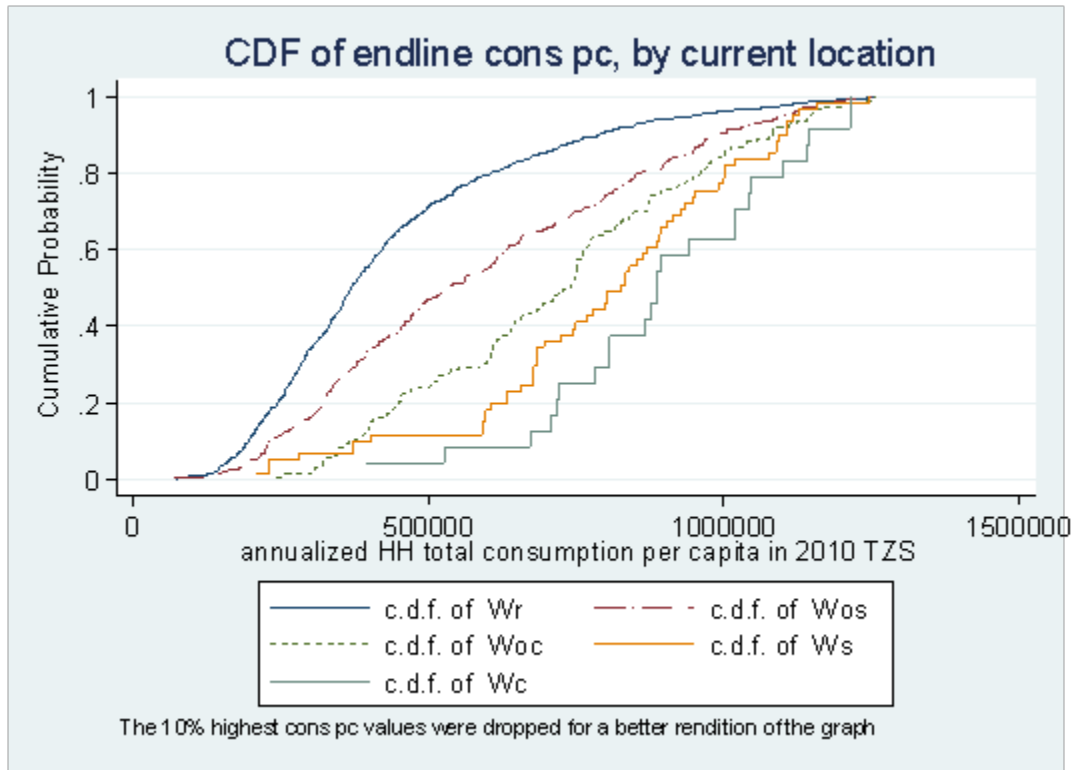
⁴ We distinguish between urban and rural centers using the 2002 census classification. Urban centers with more than 500,000 inhabitants are considered cities, the others are towns. Using this definition the cities are Dar es Salaam (the economic capital of Tanzania with an estimated 4.36 million people according to the 2012 census) and Mwanza (with an estimated 0.7 million inhabitants)

⁵ Of the formally wage employed, 70 and 80 percent work in the private sector in towns and cities, respectively. In towns, 21 percent also work in the public sector, while that share is 11 percent in the cities. The remaining 7 percent work for NGOs or religious organizations, in both the towns and the city.

informal city employment, formal town employment, formal city employment). Given the difficulties in obtaining reliable income data in African settings, consumption per capita⁶, spatially deflated and expressed in 2010 Tanzanian Shilling (TSH), is taken as proxy for income.⁷

Taking median incomes (in ‘000 TSH) of each group yields: $w_r = 372 < w_{os} = 613 < w_{oc} = 871 < w_s = 1,088 < w_c = 1,475$. This conforms with the income gradient assumed in the model. A similar gradient is observed when taking average incomes, which are more sensitive to outliers: $w_r = 495 < w_{os} = 765 < w_{oc} = 1,063 < w_s = 1,325 < w_c = 1,605$. Figure 1 overlays the cumulative density functions (cdf) for each of the 5 incomes of interest and clearly shows a similar dominance pattern across the different distributions. For a largest part the income distributions first order dominate each other in the expected manner, with perhaps the only exception W_s crossing W_{oc} in the left and right tail of the distribution.⁸

Figure 1: Cumulative density function of income (2010) by income groups



While the poverty line can be set at different levels (Ravallion, 2016), one natural anchor point, adopted as benchmark by the international community in the Sustainable Development

⁶ Similar results are obtained when using consumption per adult equivalent.

⁷ For a detailed description of the data and the consumption variable construction, we refer to De Weerd et al. (2012).

⁸ Formal secondary town employment dominates informal city employment over most of the income range, except at the very bottom and top of the income distribution, which may also be linked to measurement errors.

Goals, is the international \$1.9-day poverty line (in 2011 prices), which basically corresponds to the national poverty line of the poorest nations, corrected for 2011 differences in purchasing power (Ferreira et al. 2016). Converted in '000TSH, this yields: $z=406$ for our 2010 Kagera population. It puts us in scenario (24) in Table 1 ($w_r < z < w_{os} < w_{oc} < w_s < w_c$).

The observed poverty headcount gradient across our population is: $P_{0_r}= 0.56 > P_{0_s}= 0.25 > P_{0_c}= 0.08$. For the poverty gap, we find: $P_{1_r}= 0.17$, and $P_{1_s}= 0.07$, and $P_{1_c}=0.01$. This suggests a clear poverty gradient declining from rural to town to city.⁹ Finally, according to (34), the creation of an additional formal wage job in town would be more poverty reducing than the creation of an additional formal job in the city if $1/e_s > 1/e_c$ or if $e_c > e_s$. With $e_c=0.222$ and $e_s=0.189$, the model would favor formal wage job creation in the towns. While overall 5.3 more poor people would be lifted out of poverty with one additional formal wage job in town, according to the model, 4.5 more poor people would be lifted out of poverty for an additional formal wage job in the city. This is because formal job generation in town induces a larger migrant flow out of the rural areas than job generation in the city, which the model links to 1) differences in migration costs; and 2) differences in the informal-formal wage gap in towns and cities.

While there is no direct information on migration costs, reported one-way transport costs are on average twice as high to the city as to the towns (47,140 TSH vs 23,346 TSH respectively) and migrants to towns are 8 percentage point more likely to have found work before moving than migrants to city, illustrating differences in search costs. Reliance on friends and family for the first residence was also much more important for city migrants (56 percent) than for town migrants (30 percent) who were more likely to own a dwelling (32 percent versus 16 percent in cities) or have housing provided by the employer (8 percent, versus 2% in cities). In each case about 30 percent rented a dwelling. Overall, this suggests higher settlement barriers for city migrants, and a greater role of family and friends to overcome them. As distance reduces the likelihood of a first mover, and thus the likelihood of having family and friends far away, migration costs to cities are bound to increase, or poorer households may simply be locked out (McKenzie et al., 2010; Beegle et al., 2011; Bryan, Chowdhury, Mobarak, 2014).

While these calculations abstract from other consideration such as differences in the cost of formal job creation and further spillover effects, they illustrate the empirical relevance of the model and the different feature affecting the poverty reducing effects of formal job creation in different urban locations. Different poverty lines and urban definitions could further be taken. Considering only Dar as a city (and Mwanza as a secondary town or secondary city, as is typically done) would for example reinforce the notion that job generation in towns could be more poverty reducing in Tanzania's current context. Based on the model's findings (and thus abstracting from differences in the costs of employment generation and spillover effects), one additional formal wage job in Dar would lift 3.36 poor people out of poverty ($e_c=0.296$), though 5.74 ($e_s=0.174$) when created in the town (including Mwanza in this case).

⁹ The slight deviation from the model's prediction, whereby poverty levels in town and city are equal ($P_{\alpha_r} > P_{\alpha_s} = P_{\alpha_c}$) follows from the fact that the income distribution in the respective destinations contains in practice more than two income profiles (formal and informal), contrary to what is assumed in the model.

6. Heterogeneous Migrants and Jobs

In the spirit of the basic Todaro model, the previous sections assume that the migrants are identical. But the empirical example already alludes to the fact that there is bound to be heterogeneity and a major strand in the migration literature has relaxed this assumption and considered the consequences, for example, when agents in the rural area differ by wealth or by education. How would such a complication affect the analysis of poverty gradients and job creation and poverty reduction? We show below that although the results are further complicated, they maintain their basic structure and economic intuition.

Let us consider the situation where migrants and jobs are segmented by education or skill level. The distinction is indicated by superscript “*a*” for advanced education and superscript “*b*” for basic education. There are separate labor markets and migration streams for each type of labor. Total population is now

$$\begin{aligned}\bar{N} &= \bar{N}^a + \bar{N}^b \\ &= [N_r^a + N_s^a + N_c^a] + [N_r^b + N_s^b + N_c^b]\end{aligned}\tag{41}$$

But the migration streams of *a* and *b* types are segmented and follow their own logic, comparing the costs and benefits of migration to small towns and cities.

The returns and costs to migration are now superscripted by *a* and *b* to distinguish between the two types of labor. We assume that the pattern of wages set out in (18), (19) and (22) holds for each type of labor – in other words with superscripts *a* and *b* respectively for advanced and basic education.¹⁰ Further, we assume that each wage for advanced exceeds the corresponding wage for basic and that the transport costs of *a* are less than the transport costs of *b*. With this specification, the equilibrium conditions set out in (12) and (13) are simply replicated with superscripts *a* and *b*. Thus the equilibrium values (14) – (17) are also replicated with superscripts *a* and *b*.

Since we assume that each wage for type *a* exceeds each corresponding wage for type *b*, poverty within group *a* will be lower than poverty within group *b*:

$$P_\alpha^a < P_\alpha^b\tag{42}$$

Overall poverty P_α is simply a population weighted sum of poverty among *a* and *b* types:

¹⁰ This assumption finds empirical support in Figure 1, which shows that the income gradient (informal rural to informal town, informal city, formal town and formal city) is retained across most of the income distribution. To the extent that heterogeneity in ability coincides with the ranking position in the income distribution (say the 20th percentile for those with basic and the 80th percentile for those with advanced education), it can then be seen that the income gradient is replicated for both low and high performers.

$$P_\alpha = \left[\frac{\bar{N}^a}{\bar{N}} \right] P_\alpha^a + \left[\frac{\bar{N}^b}{\bar{N}} \right] P_\alpha^b \quad (43)$$

The poverty gradients for $(P_{\alpha,r}^a, P_{\alpha,s}^a, P_{\alpha,c}^a)$ and $(P_{\alpha,r}^b, P_{\alpha,s}^b, P_{\alpha,c}^b)$ are determined by the same arguments as used for comparing (26), (27) and (28) when wages for each type a and b lie in the relation (24). In this case there is a poverty gradient for each type, with poverty decreasing from rural to secondary town to city. When wages in relation to the poverty line are given by (29) for each type, the gradient depends on the relative values of (30), (31) and (32) for each type. If the unemployment (or informality) rate for each type is lower in the city than in the town, then we will again have a declining poverty gradient from rural to town to city. Of course, if each type a and b has a declining poverty gradient, then overall poverty will also show this gradient.

The poverty impact of job creation in the secondary town or the city, by type a or b , is now given analogously to (34):

$$\begin{aligned} \frac{dP_\alpha}{dE_i^i} &\begin{matrix} \geq \\ \leq \end{matrix} \frac{dP_\alpha}{dE_i^j} \\ &\Leftrightarrow \\ e_i^j &\begin{matrix} \geq \\ \leq \end{matrix} e_i^i \quad ; \quad \begin{matrix} j = a, b \\ i = s, c \end{matrix} \end{aligned} \quad (44)$$

As in the discussion of (34), the relative poverty impact of modern job creation in one of the four sub-sectors specified by $i = s, c$ and $j = a, b$ is given by the relative modern employment rate in that sector. For scenario (29), the analogy is with (35) – (39). In this case the equivalent expression to (39) can be written as

$$\begin{aligned} \frac{dP_1}{dE_i^j} &\begin{matrix} \geq \\ \leq \end{matrix} \frac{dP_1}{dE_i^i} \\ &\Leftrightarrow \\ X_i^j &\begin{matrix} \geq \\ \leq \end{matrix} X_i^i \quad ; \quad \begin{matrix} j = a, b \\ i = c, s \end{matrix} \end{aligned}$$

Thus the efficacy of job creation by location and education level is captured by the four sufficient statistics X_i^j , analogously to the two sufficient statistics in (39). The basic structure of the policy rule is maintained in this more complicated setting with heterogeneous migrants and jobs.

7. Conclusion

In this paper we first extended the basic Todaro-type model of rural-urban migration to the case of migration from rural areas to two potential destinations, secondary town and big city, and derived conditions under which a poverty gradient from rural to town to city will exist as an equilibrium phenomenon. We then addressed the motivating question: Should public investment be targeted to big cities or to small towns, if the objective is to minimize national poverty?

We have shown that the answer to this question is not self-evident and depends on the migration response (which in turn depends on the migration cost and the difference in the informal-formal wage gaps between the town and the city), where the poverty line lies relative to incomes in the three locations, and at times also the poverty index. We developed sufficient statistics for the policy decisions based on these income parameters and illustrated the empirical relevance of this model with an example of a long running panel from rural Kagera, Tanzania, which represents a natural migration equilibrium, and thus an ideal platform to give empirical content to the model and explore its relevance in practice. Further, we showed that the structure of the sufficient statistics is maintained in the case where the model is generalized to introduce heterogeneous workers and jobs.

Together these results show that the common consideration of urban areas in the aggregate foregoes important dimensions of these urban areas, which can have an important bearing on the poverty outcomes of urban investment. More analysis of the distributional effects of urban investment, which explicitly incorporates such disaggregation is called for. This paper has taken one step in that direction.

References

- Anand, Sudhir and Ravi Kanbur. 1985. "Poverty under the Kuznets Process." *The Economic Journal*, Vol. 95, Issue 380a, pp. 42-49, 1985.
- Basu, Arnab, Nancy Chau, Gary Fields and Ravi Kanbur. 2016. "Job Creation in a Multi-Sector Labor Market Model for Developing Economies." CEPR Discussion Paper, No. 11386.
- Beegle, Kathleen, Joachim De Weerd, and Stefan Dercon. 2011. "Migration and Economic Mobility in Tanzania: Evidence from a Tracking Survey." *Review of Economics and Statistics* no. 93 (3):1010-1033.
- Beegle, Kathleen, Luc, Christiaensen, Andrew, Dabalen, and Isis, Gaddis, 2016. "Poverty in a Rising Africa." World Bank: Washington D.C.
- Bryan, Gharad, Shyamal, Chowdhury, and Ahmed Mushfiq, Mobarak, 2014. "Underinvestment in a Profitable Technology: The Case of Seasonal Migration in Bangladesh." *Econometrica*, 82-5: 1671-1748.
- Christiaensen, Luc and Ravi Kanbur. 2016. "Secondary Towns and Poverty Reduction: Refocusing the Urbanization Agenda." World Bank Policy Research Paper, No. 7895.
- Christiaensen, Luc, Joachim De Weerd and Ravi Kanbur, 2017. "What Contributes More to Poverty Reduction: Migration to Cities, Migration to Towns, or Migration to other Rural Areas?" mimeographed.
- De Weerd, Joachim, Kathleen Beegle, Helene Bie Lilleør, Stefan Dercon, Kalle Hirvonen, Martina Kirchberger, and Sofya Krutikova. 2012. "Kagera Health and Development Survey 2010: Basic Information Document." Rockwool Foundation Working Paper Series no. 46.
- Ferre, Celine, Francisco H.G. Ferreira, and Peter Lanjouw. 2012. "Is There a Metropolitan Bias? The relationship between poverty and city size in a selection of developing countries." *The World Bank Economic Review*, Vol. 26, pp. 351-382.
- Food and Agriculture Organization, 2017. "State of Food and Agriculture 2017." Food and Agriculture Organization: Rome, forthcoming.
- Harris, John and Michael Todaro. 1970. "Migration, Unemployment, and Development: A Two Sector Analysis," *American Economic Review* 40: 126-142.
- Fafchamps Marcel, and Forhad Shilpi. 2013. "Determinants of the Choice of Migration Destination", *Oxford Bulletin of Economics and Statistics*, 75(3): 388-409.
- McKenzie, David, John Gibson, and Steven Stillman. 2010. "How Important is Selection? Experimental vs Non-experimental Measures of the Income Gains from Migration." *Journal of the European Economic Association* 8(4): 913-45.
- Ravallion, Martin. 2016. *The Economics of Poverty: History, Measurement and Policy*. Oxford University Press, New York and Oxford.
- Schwartz, Aba. 1973. "Interpreting the Effect of Distance on Migration." *Journal of Political Economy*, 81-5: 1153-1169.

Sjaastad, Larry A., 1962. "The Costs and Returns to Human Migration." *Journal of Political Economy* 70-5: 80-93.

Temple, Jonathan. 2005. "Growth and Wage Inequality in a Dual Economy." *Bulletin of Economic Research*, 57(2), 145-169.

Temple and Ying 2014. "Life During Structural Transformation." University of Bristol Economics Discussion Paper NO. 14/650.

Todaro, Michael P., 1969. "A Model of Labor Migration and Urban Unemployment in Less Developed Countries." *American Economic Review* 59(1): 138-148.

Ferreira, Francisco H.G., et al., 2016. "A global count of the extreme poor in 2012: data issues, methodology and initial results," *Journal of Economic Inequality*, 14(2): 141-172.

World Bank, 2016. *Poverty and Shared Prosperity 2016. Taking on Inequality*, World Bank: Washington D.C.