Migration, Unemployment and Growth Through Urban Agglomeration:∗
A Harris-Todaro Growth Model

April 2008

Chul-In Lee
SungKyunKwan University (SKKU)

Abstract: Based on agglomeration externalities, we build an endogenous growth model that is consistent with the set of empirical regularities – economic development transitorily involves rural-to-urban migration, urban unemployment, and the resulting urban concentration. Urban agglomeration through migration can generate productivity growth, which attracts further migration. When this self-enforcing feature is combined with the assumption that the urban sector exhibits greater frictions in job matching, we obtain the equilibrium growth path along which the wage and unemployment gaps between rural and urban sectors arise during the urbanization process, the phenomena noted by Harris and Todaro (1970) in their static partial migration equilibrium. We characterize the optimal growth path and draw some policy implications.

Key Words: endogenous growth, urban agglomeration, optimal growth, urban versus non-urban wage and unemployment gaps, migration.

∗ Acknowledgments. The author thanks Charles Brown, Yongsung Chang, Youngjun Chun, Hyungjun Kim, Namhoon Kwon and seminar participants at Yonsei University and the 2007 KPUBE Fall meeting for useful comments and constructive suggestions. Email: leeci@skku.edu. Address: 53, 3-ga, Myongnyun-dong, Chongno-gu, School of Economics, SungKyunKwan Univ., Seoul, 110-745, Republic of Korea; phone: 82-2-760-0421.
I. Introduction

Since Marshall (1890), economists have noted that urban agglomerations create technological externalities: the information spillover benefits in input and output markets of having economic agents in close spatial proximity where information decay over space is very rapid. In addition, the new economic geography develops the idea that spatial proximity involves pecuniary externalities, which reduces the costs of intermediate and final good trade.\(^1\) Recently, the economic growth literature offers further insight into urban agglomeration by viewing cities as promoting human capital accumulation through active learning (e.g., Lucas (1988), and Glaeser and Mare (2001)). In this paper, we show that these well-known “agglomeration benefits” can generate endogenous growth\(^2\) through urban concentration with the feature that wage and unemployment gaps arise between urban and rural areas during the massive rural-urban migration period, a typical developing economy phenomenon noted by Harris and Todaro (henceforth, HT) (1970).

Three prominent aspects of actual economic development are typically omitted from conventional growth models. First, modern economic development has invariably been accompanied by pronounced migration from rural to urban sectors of employment. Furthermore, the growth rate is usually high during this transitory urbanization period while tapering off over time, i.e., a transitory growth effect. Second, for most of the developing countries except for planned economies, urbanization processes take a form of concentration toward a few central cities rather than evenly spreading out new migrants over cities, which generates agglomeration

---

1 See the excellent survey by Henderson (2004).
2 In general, endogenous growth refers to the case where the shares of accumulating factors sum up to unity or above, and thus per capita income growth does not stop even in the long run. As we will see later, our model covers both the case mentioned and the case where per capita income growth eventually stops, i.e., the so-called “semi-endogenous growth.”
It is questionable that urbanization per se brings in permanent growth effects, but agglomeration economies formed with urbanization appear to be essential at least in the process of economic development for many developing economies. Despite the importance of economic growth and urban concentration, the usual optimality question has rarely been raised except in an empirical exploration by Henderson (2003), who noted from international panel data the notion of “optimal urban concentration.” Third, many developing economies in the process of urbanization have a large portion of their labor force facing urban unemployment, often called the Todaro paradox. This regional unemployment pattern usually accompanies the wage gap between agglomerated cities and other areas.

There are many reasons to think that these factors play an important role in the process of economic development. The consequences of rural-urban migration are often dramatic: between 1950 and 1975 the share of urban population in total third-world population rose from 16.7 to 28 percent (Williamson 1988, p. 428). This type of migration raises the extent of urban concentration, and the resulting agglomeration benefits can lead to urban productivity growth. More precisely, agglomeration externalities can be a source of endogenous growth, because urban productivity growth based on agglomeration allows higher wages and greater employment.

---

3 Of course, urbanization is a concept different from urban concentration, but considering the well-established Zipf’s law, a systematic empirical relationship between the size of a city and its ranking, that applies worldwide, including in developing countries, it seems harmless to say that urbanization and urban concentration are highly correlated, lending support to the statement given in the text. While there remains an issue of the “first-city bias,” most countries that either are currently or were formerly at the stage of economic development exhibit fast growth during the urbanization process.

4 Urban concentration sometimes exceeds the optimal degree and is referred to as the so-called “urban bias.” Political reasons such as lobbying and corruption are argued to be sources of such urban bias.

5 Henderson (2003) finds no econometric evidence linking the extent of urbanization to either economic or productivity growth or levels and views urbanization as something that emerges as part of the growth process. In contrast, Gallup, Sacks and Mellinger (1999) suggest that urbanization may cause growth, rather than just emerging as part of the growth process. Our view may be seen as combining these two views. In our endogenous growth model, urbanization emerges as part of the growth process and at the same time, advancement of urbanization brings in agglomeration externalities, which fosters productivity growth to occur transitorily along the transition path.

6 See Hatton and Williamson (1992) for a survey of the typical wage gaps in developing economies.
opportunities and thus induces further migration to the urban sector, which in turn boosts the
degree of urban agglomeration and hence productivity and migration, generating a cycle of self-
enforcing dynamics. In this process, unemployment in the urban sector plays the role of
controlling migration flows, which establishes equilibrium at a point in time with a significant
but gradual migration to the urban sector. This point therefore offers an answer to the following
open question: if human capital accumulates fast in cities (e.g., Lucas (1988) and Glaeser and
Mare (2001) among many others), why doesn’t everyone move to cities?

The agglomeration-based endogenous growth would decline with urban concentration
as geographical migration reaches its limit, presumably because of either congestion
externalities or the literal constraint that urban concentration cannot exceed 100%. Roughly put,
in this model we can conceptualize the optimal growth path by balancing the positive net
externalities from spatial agglomeration with the distortionary uses of scarce resources arising
from inter-regional job matching. As will be shown later, along the endogenous growth path we
can have a transitory growth effect that arises from agglomeration. Our model therefore
provides a theoretical underpinning for the empirical findings of Henderson (2003): (i)
urbanization is a transitory phenomenon, (ii) productivity growth is strongly affected by urban
concentration, and (iii) optimal concentration varies depending on the level of development.

This paper can also be viewed as developing a growth model in line with the well-
known Harris-Todaro (1970) migration model established in the context of economic
development, so our model can be referred to as the Harris-Todaro growth model. The
motivation for HT models comes from accounting for the stylized fact that high urban
unemployment has been widely observed in LDCs together with a substantial wage gap between
urban and non-urban sectors.7

7 The earliest insight on this wage gap can be found in HT’s (1970) rural-urban migration model that has long been
While there are lots of important studies that extended or improved the HT migration equilibrium, what separates this study from previous studies lies in the following. First, viewing agglomeration as the result of accumulated migration, our model can account for the stylized transitional urbanization dynamics that involve the comovements of productivity, rural and urban wage changes, and migration flows, along with the wage and unemployment gaps across sectors. Second, while the original HT model is some sort of a static partial equilibrium model, we extend it to be a dynamic general equilibrium model to address typical growth paths. Our model has general features: (i) flexible urban and rural wages and (ii) free mobility of workers and firms, while still keeping the original motivation of HT as closely as possible. In a fully general dynamic search-matching context, we highlight the exact conditions for the HT equilibrium growth path.

Our view of the urban sector shares with Fujita and Thisse (1996) and Behrman (1999) the following point: compared to non-urban jobs, urban jobs are more heterogeneous and specific, causing search frictions to both firms and workers. A natural implication from this point is that the urban sector faces a more costly job-matching process under imperfect information. We argue that these differences in matching technology between the sectors, when combined with the productivity differential arising from urban agglomeration, permit

---

8 Many studies have also attempted to extend or improve the original HT migration equilibrium by endogenizing urban wages, which are assumed to be fixed at a level higher than non-urban wages. Urban wages in the HT models are assumed to be high and “institutionally” fixed due to minimum wage legislation or the power of labor unions (e.g., Calvo, 1978 and Quibria, 1988). The first type of these models uses the efficiency wage approach under the presumption that the urban sector gives efficiency wages for incentive reasons (e.g., Stiglitz (1974), among many others). The second approach endogenizes the urban wage through a bargaining between a trade union and firms (e.g., Calvo, 1978). The third type applies search-matching models to migration (e.g., Ortega, 2000, among many others). While the first two approaches are certainly appealing, this paper belongs to the third type of approach by viewing migration as a forward-looking behavior that involves unemployment due to search frictions and causes externalities.
equilibrium at a point in time with the wage and unemployment gaps across the sectors. Given the endogenous nature of agglomeration effects, we have the regional equilibrium and the equilibrium gaps that change over time with respect to urban productivity updating. Our two-sector (urban and non-urban) HT growth model with search matching is a natural integration of (i) the original HT, (ii) the classic one-sector macroeconomic search-matching models of Diamond (1984) and Mortensen and Pissarides (1994), and (iii) the large literature on urban agglomeration and the Romer (1990) type endogenous growth models.

In addition to HT models, there are some notable previous studies related to this paper. Lucas (1988) may be the first who recognized the close connection between urban and national economic growth. If urban agglomeration not only generates scale economies but also fosters human capital accumulation through active learning, endogenous growth is possible even in the long run. In a study related to this paper, Bencivenga and Smith (1997) proposed a neoclassical growth model with rural-urban migration and urban underemployment in the informal sector, which arises from a typical adverse selection problem in labor markets for workers with heterogeneous abilities. They show that this migration and underemployment can be a source of development traps and can give rise to a large set of periodic equilibria displaying undampened oscillation. In contrast, our endogenous growth model focuses on agglomeration *externalities* as a source of growth in the economic development process and offers the optimal growth path and policy implications in a search-matching framework. We present growth through urbanization and the key features of the HT equilibrium based on agglomeration externalities and search-matching frictions in the urban sector. Shukla and Stark (1990) and Krichel and Levine (1999) also modeled urban agglomeration economies as a labor pull factor in the regional equilibrium context, but the static nature of their models does not address the endogenous growth arising from agglomeration economies. Compared with the existing urban economics literature, this
paper views agglomeration as the result of accumulated migration over time and examines the growth implications of agglomeration with a more structural interpretation of unemployment in the search-matching context.

The paper consists of the following sections. Section II describes the basic structure of our model of dynamic location choice with free mobility of workers and free entry of firms. To highlight the growth from agglomeration, we focus on agglomeration externalities only, abstracting from other sources of growth, e.g., population growth or savings. Section III presents steady-state equilibrium, key results and interpretation of them, and the optimal growth path along with policy implications. A short summary and conclusions are given in Section IV.

II. The Model

1. Individuals’ decisions

The labor force with its size normalized at unity chooses its location in order to maximize the present value of its future income stream under the economic environment to be explained below. In their job-matching process, workers consider that the urban sector is characterized by search frictions, whereas the non-urban sector is competitive. To discuss its location choice and the conditions for inter-regional labor market equilibrium, we first define three different value functions.

The present value of urban employment at \( t \), \( W(t) \), can be expressed using the Bellman equation as follows. To simplify notations and save space, we henceforth denote the current period’s variable \( x(t) \) simply as \( x \) and the next period’s variable \( x(t + 1) \) by \( x' \) in the Bellman equations.
\[(1 + r)W = w + \{(1 - \lambda) \cdot \max[W', X'] + \lambda \cdot X'\}, \tag{1}\]

where \(w\) is the flow urban wage received at the end of period \(t\) (henceforth, all flow variables are measured at the end of period), \(r\) is the discount factor, \(X' = \max[U', R']\), and \(U\) and \(R\) are the values of urban unemployment and non-urban employment, respectively, to be defined below.

Equation (1) implies that \((1 + r)W\) equals a flow wage rate \(w\) plus the value of the state-contingent choices in the next period: a weighted average of (i) the maximum value among \(W'\), \(R'\) and \(U'\) conditional on the current employment continuing with probability \(1 - \lambda\) and (ii) the greater value between \(R'\) and \(U'\) conditional on the current employment discontinuing with probability \(\lambda\).

Similarly, we can construct the value function for an urban unemployed worker searching for an urban job, \(U(t)\): it is equal to the flow unemployment income \(b(t)\) plus the weighted average of (i) the maximum value conditional on a job offer arriving with probability \(\theta' \cdot q(\theta')\) that is derived from the CRS matching technology\(^9\) and (ii) the maximum value conditional on a job offer not arriving with probability \(1 - \theta'q(\theta')\):

\[(1 + r)U = b + \{\theta'q(\theta') \cdot \max[W', X'] + (1 - \theta'q(\theta'))X'\}. \tag{2}\]

For the sake of simplicity, we will normalize \(b\) at zero. Meanwhile, we assume without loss of generality that rural workers must move to the city if they want to find an urban job. This assumption accords well with the original intuition of HT.\(^{10}\) Taking account of this, the value

---

\(^{9}\) The CRS matching technology is: \(m(sL, vL) = m(s, vL)\), where \(v\) is the aggregate level of vacancy; \(s\) is the aggregate search intensity of urban unemployed workers: \(s = eu_i\); \(e\) is the search intensity of urban unemployed workers with size \(u_i\) (henceforth, \(e\) is normalized at unity); and \(L\) is the urban workforce. From this, we can define a measure of job-market tightness by the ratio of the number of vacancies to that of searchers, \(\theta(t) = v(t)/u(t)\). The probability that a firm with a vacancy will find a worker (i.e., the worker-arrival rate) therefore becomes \(q(\theta(t)) = m(s(t)/v(t), 1)\). By the properties of the matching technology, \(q(\theta)\) is a number between 0 and -1. Meanwhile, urban unemployed workers can obtain an urban job with probability \(\theta(t)q(\theta(t))\). It is often called the unemployment hazard rate or the job-arrival rate. See Pissarides (2000) for details.

\(^{10}\) An earlier version of this paper allows non-urban workers to search for urban jobs while staying in the non-urban sector. It uses a high search efficiency parameter for urban workers and a low efficiency one for rural workers.
function $R(t)$ for non-urban workers is defined using the flow non-urban wage $w_R(t)$ and the value of $X(t+1)$:

$$(1 + r)R = w_R + X'.$$  \hspace{1cm} (3)

2. Firms’ decisions

Both urban and non-urban firms produce an identical good with its price normalized at unity. Their entry is regionally unrestricted. Let $J(t)$ be the present value of the expected profit stream from a filled urban job. Similarly, $V(t)$ is the present value of the expected profit stream from a vacant job. Then, an urban firm with a vacancy filled faces the following value function:

$$(1 + r)J = p - w + [(1 - \lambda) \cdot \max(J', V') + \lambda V'],$$  \hspace{1cm} (4)

where $p(t)$ is the flow output when the match occurs at time $t$, $p(t) - w(t)$ is the flow profit conditional on the match, and the last term is the value of the firm’s state-contingent choices in the next period: a weighted average of (i) a greater value between $J'$ and $V'$ conditional on continuous operation with probability $1 - \lambda$ and (ii) the value $V'$ conditional on shutting down with probability $\lambda$.

Similarly, an urban firm with a vacancy faces the following value function:

$$(1 + r)V = -c + [q(\theta') \cdot \max(J', V') + (1 - q(\theta')) \cdot V'],$$  \hspace{1cm} (5)

where $c$ is the flow opportunity cost of a vacant job and the firm can hire a worker with probability $q(\theta')$ as explained in footnote 8. As will be shown later, perfect competition drives the steady-state value of a vacancy, $V(t)$, down to zero.

Meanwhile, wages for urban matched jobs are derived from the generalized Nash

---

Since the main results do not change, we show the simplest model here. The result for that case is available upon request.
bargaining solution: \( w_i(t) \) maximizes the weighted product of firm \( i \)'s and its workers’ net return from a productive job match. In this bargaining the threat points for the worker and the firm are the opportunity values, \( X(t) = \max[U(t), R(t)] \) and \( V(t) \), respectively. Since the threat point for workers should be the second highest wage (the opportunity wage) at which urban firms can attract workers, we thus use the greater one between \( U(t) \) and \( R(t) \) as a worker’s threat point. Therefore the wage rate for this job satisfies:

\[
w_i(t) = \arg \max \left\{ [W_i(t) - X(t)]^\gamma [J_i(t) - V(t)]^{1-\gamma} \right\},
\]

where \( \gamma \) is the parameter for bargaining power. This equation suggests that there may be a “premium” for matched workers and firms because they are engaged in a bilateral monopoly situation to share the benefit of greater productivity (e.g., arising from agglomeration in the urban sector).

Next, aggregate non-urban production is described by \( A_g(t) = F(u_2(t), z(t)) \), where \( u_2(t) \) is the non-urban population and \( z(t) \) is the size of land. To simplify the discussion, we will use the CRS agricultural production function and inelastic supply of land as follows:

\[
F(u_2(t), z(t)) = u_2(t)^\beta z^{1-\beta}.
\]

If we normalize land size at unity, the aggregate production function is \( f(u_2(t)) = F(u_2(t),1) = u_2(t)^\beta \); the corresponding wage and rent functions are given, respectively, by:

\[
w_g(t) = \beta u_2(t)^{\beta-1}, \quad \text{and}
\]

\[
r_g(t) = (1-\beta)u_2(t)^\beta.
\]

In this formulation, migration to the urban sector reduces the non-urban population, leading to an increase in non-urban labor productivity (i.e., see equation (8-1)). Based on these results, we can express the value of a representative non-urban firm, \( J_g(t) \), as follows:
In equilibrium, non-urban firms will receive a zero profit due to competition combined with the CRS production technology. Also, the value of an urban firm with a vacancy, $V(t)$, equals zero due to inter- and intra-sectoral competition.

3. Agglomeration and urban productivity

Drawing on the literature related to our paper [e.g., Lucas (1988), Romer (1990), Shukla and Stark (1990), Krugman (1991), and Krichel and Levine (1999)], we express urban productivity to depend on the size of urban population: urban agglomeration not only reduces production/transaction costs but also fosters knowledge spillovers in the urban sector, generating positive externalities in terms of productivity. A central feature of this paper is to extend this static idea from the perspective of economic growth by viewing agglomeration in a dynamic context. Urban population/concentration grows over time owing to migration, which generates productivity growth over time. We therefore posit that urban productivity will improve as urban concentration proceeds. We take a general form of urban productivity updating that depends on the agglomeration effects $g(.)$, which is a function of urban population $l_i + u_i$ and existing productivity $p$:

$$p' = p + g(l_i + u_i, p).$$

(10)

It implies that productivity improvement, $p' - p$, arises until the urban agglomeration benefits exceed possible negative congestion externalities.\(^{11}\) Of course, productivity falls when the

\(^{11}\) If individuals expect a specific form of a productivity evolution path independent of the agglomeration effects, we should use that form, but given its arbitrariness, we assume it to be constant for analytical simplicity.
congestion factor dominates agglomeration benefits, which is consistent with Henderson’s (2003) notion of non-optimal urban concentration. An example of the function $g(\cdot)$ would be

$$g(\cdot) = \delta_0 \left[ l_i + u_i \right]^\lambda p^\nu - \delta_1 \left[ l_i + u_i \right]^\nu,$$

where $\lambda$ and $\mu$ are the parameters for qualifying agglomeration benefits and $\nu$ is the parameter for congestion with $\nu > 1$. This form of agglomeration benefits matches closely with the knowledge dynamics of Romer (1990), where new knowledge is created through R&D labor and existing knowledge: e.g.,

$$A = \delta_L A^\lambda A^\nu,$$

where $A$ is the knowledge stock, $L_A$ is R&D labor, and $\lambda$ and $\mu$ here are the parameters for qualifying knowledge growth. Note that $A$ and $L_A$ in Romer are analogous to $p$ and $l_i + u_i$ in our model, respectively. As one would guess at this point, $g$ converges to zero at the long-run interior stationary equilibrium where no further migration occurs. When $g$ remains positive, agglomeration benefits attract migration from the rural sector and the urban concentration rate rises over time.

As in many growth theories with externality (e.g., Arrow (1962), Romer (1990)), we make a comparable assumption about productivity growth: each firm’s or individual’s knowledge is a public good that anyone can access at zero cost, or simply, the *non-rival* nature of production technology. Under this assumption, the benefits of agglomeration are reflected on the production technology and are shared across urban firms at no cost. A parallel assumption is that each agent is *small* enough to neglect its own contribution to aggregate urban productivity. The endogenous productivity gain through migration is therefore unanticipated by private agents when they consider making a migration decision or an entry decision into the urban sector.

---

12 He points out that “the benefits of increasing primacy -- enhanced local scale economies contributing to productivity growth -- against the costs -- more resources diverted away from productive and innovative activities to shoring up the quality of life in congested primate cities.” Based on this, he finds that there is an optimal degree of primacy at each level of development that declines as development proceeds. In this model, it amounts to the case where $\frac{dg(l_i + u_i, p)}{dt} < 0$. 

11
However, once the *ex post* productivity increase is realized, firms can afford to open more vacancies in the next period since firms’ profitability improves. This attracts further migration, which in turn boosts productivity, a self-enforcing dynamic that lasts until net agglomeration externalities vanish.

### III. The Equilibrium and Transition Path

Now, we define equilibrium based on the framework described in Section II. Conceptually, three different long-run equilibria are possible: two corner solutions (equilibrium with no urban population or that with perfect urban concentration) and an interior solution (equilibrium with imperfect urban concentration). Perhaps, the interior solution case seems more realistic, but in any case, we can describe the equilibrium at a point in time along the transition path toward the long-run steady state.

#### 1. Equilibrium at a point in time

Suppose that the current period’s urban productivity is $p$. In our model, economic agents act with the belief that their actions do not affect the economy, while their behaviors in aggregate create dynamics by affecting the level of agglomeration. Taking advantage of this point, we can define some sort of Nash equilibrium in the current period. In accordance with productivity updating in equation (10), we can determine equilibrium in the next period. The differences in urban population and other variables between periods $t$ and $t+1$ are due to migration. Repeating this process gives us equilibrium at a point in time. The following system of equations can describe the HT equilibrium at a particular period under flexible wages and free mobility:
\[ \lambda l = \theta \cdot q(\theta) u \quad (11-1) \]
\[ rV = -c + q(\theta) (J - V) \quad (11-2) \]
\[ rJ = p - w - \lambda J \quad (11-3) \]
\[ rW = w + \lambda (R - W) \quad (11-4) \]
\[ rU = \theta \cdot q(\theta) (W - R) \quad (11-5) \]
\[ rR = w_r (u_2) \quad (11-6) \]
\[ W - R = \gamma (J + W - V - R) \quad (11-7) \]
\[ R = U < W \quad (11-8) \]

where \( I = l + u + u_2 \) (\( l \) is the urban employed population, \( u \) is the urban unemployed population, \( l + u \) is the urban population or the “urbanization rate,” and \( u/(1-u_2) \) is the urban unemployment rate), \( \theta = v/u_1 \), \( V = 0 \), \( b \) is normalized at zero, and \( p \) is the productivity level at a point in time that evolves over time in accordance with equation (10) but that agents perceive to be fixed.

Equation (11-1) is the so-called “Beveridge equation” modified in the context of developing economies with two sectors. With the equilibrium values of \( l \), \( u \) and a measure of market tightness \( \theta \), it implies that the number of workers who enter urban unemployment is equal to that of workers who leave urban unemployment in equilibrium, so the equilibrium urban unemployment rate can be obtained. Equations (11-2) to (11-6) represent value functions in equilibrium that are derived from equations (1) to (5) combined with equation (11-8), which deserves some discussion. Equation (11-8) is the inter-regional labor market equilibrium condition and can also be interpreted as an analogue to HT’s migration equilibrium in a dynamic setting: staying in the non-urban sector is equal in value to migrating to the urban sector as an
unemployed worker searching for a job. This condition also ensures the empirical relevance of the model: workers are allocated in urban and non-urban areas with non-zero urban unemployment. The reason is that given that the value of employed workers exceeds that of unemployed workers \( W(t) > U(t) \), the only possibility of explaining the coexistence of urban unemployment and non-urban employment arises when \( W(t) > R(t) = U(t) \). Using this condition, the complex location choices of mobile workers are expressed simply by (11-2) to (11-6) involving the equilibrium values of \( \theta, w, w_k \) and \( u_2 \). Equation (11-7) shows the result of wage bargaining. Finally, since any firm is free to open a job vacancy and can enter any sectors, the profit from an additional vacancy is set at zero: \( V = 0 \), a zero-profit condition for an additional firm entry.

To illustrate the equilibrium, we begin our discussion by expressing the equations for \( U \), \( R \), and \( W \) as functions of endogenous variables \( w, u_2 \) and \( \theta \):

\[
rW = \frac{w(r + \partial q(\theta))}{r + \lambda + \partial q(\theta)}, \quad rU = \frac{w \partial q(\theta)}{r + \lambda + \partial q(\theta)}, \quad rR = w_k(u_2).
\]

Substituting these values \((W, R \text{ and } U)\) and the value of \( J \) from equation (11-3) into the wage-bargaining equation (11-7), we obtain a “reduced”-form version of the wage bargaining equation:

\[
w = \frac{\gamma p(r + \lambda + \partial q(\theta))}{r + \lambda + \gamma q(\theta)}.
\]

\[13\] In the case of \( W > R > U \) or \( W > U > R \), urban unemployed workers move to the non-urban sector, or non-urban workers move to the urban sector to become unemployed workers, respectively, until \( W(t) > R(t) = U(t) \) is established.
Equation (13) contains general intuition that \( w \) increases as the worker’s bargaining power \( \gamma \), the job finding rate \( \theta q(\theta) \), or the urban productivity \( p \) goes higher\(^{14}\); the opposite is true when the job destruction rate \( \lambda \) or the interest rate \( r \) goes up.

Next, by using \( V=0 \) and equating the \( J \) value from equation (11-2) with the \( J \) value from equation (11-3), we obtain the standard “job creation” equation:

\[
p = w + \frac{(r + \lambda)c}{q(\theta)}. \tag{14}
\]

This equation tells us that the marginal product of urban labor, \( p \), is equal to the wage cost plus the expected capitalized value of the firm’s opportunity cost of a vacant job \( (r + \lambda)c / q(\theta) \).

Solving these two equations (13) and (14) along with the modified Beveridge curve (11-1) for \( \theta, w, \) and \( l_1 \), we obtain the following results:

\[
\theta = \frac{1}{\gamma} \left( \frac{p(1-\gamma)}{c} - \frac{r + \lambda}{q(\theta)} \right), \tag{15}
\]

\[
w = p - \frac{c(r + \lambda)}{q(\theta)}, \tag{16}
\]

\[
l_1 = 1 - \frac{p q(\theta) u_1 (\gamma - 1) + c (r u_1 + (u_1 + \gamma) \lambda)}{c \gamma \lambda}. \tag{17}
\]

To determine the equilibrium values of \( u_1, u_2 \) and \( v \), we use the \( R=U \) condition. Substituting the expressions for \( R \) and \( U \) from (11) into \( R=U \) yields:

\[
w \theta q(\theta) = w_R (r + \lambda + \theta q(\theta)). \tag{18}
\]

\(^{14}\) More precisely, \( dw / dp > 0 \) is always true if \( d\theta q(\theta) / dp \geq 0 \) (see equation (13)). \( d\theta q(\theta) / dp \geq 0 \) is because \( d\theta / dp > 0 \) (see equation (15)) and the elasticity of \( q(\theta) \) with respect to \( \theta \) is a number between 0 and -1 (the CRS matching technology).
Substituting equation (8-1) \( w_r = \beta u_z^{\beta-1} \) into equation (18) gives the solution for \( u_z \) as follows:

\[
\begin{align*}
    u_z &= \left( \frac{\beta (r + \lambda + \beta q(\theta))}{w \theta q(\theta)} \right)^{\frac{1}{1-\beta}}.
\end{align*}
\]  

(19)

Equation (19) provides useful results that are consistent with the stylized fact that the non-urban population declines (or “urbanization” advances) for the following changes: urban wage \( w \) goes up, the labor share in non-urban production \( \beta \) goes down, urban job finding rate \( \theta q(\theta) \) goes up, urban job destruction rate \( \lambda \) goes down, or interest rate \( r \) goes down. Some important features of our model are summarized as follows.

**Proposition 1. Equilibrium along endogenous growth path:** There exists equilibrium in which the economy grows endogenously with urban agglomeration, and the growth rate of urban production slows down as urbanization advances over time.

**Proof.** To present the existence of equilibrium at a particular period, we determine the value of \( \theta \) first for a given value of \( \rho \). \( \theta \) is not solved in an analytically simple form due to the non-linearity of the matching probability function, but Brower’s fixed-point theorem can be used to prove the existence of \( \theta \) from equation (15).\(^{15}\) \( q \) can be determined immediately (i.e., \( q = q(\theta) \)) and then \( w \) is obtained from (16). Next, with the equilibrium values of \( \theta, q \) and \( w \) already determined, we can solve for \( u_1 \) and \( u_2 \) using the Beveridge equation (11-1) and equation (17). The uniqueness of the equilibrium value of \( \theta \) can also be shown by

\(^{15}\)This is because \( q(\theta)<0 \) in the CRS matching technology.
d \left[ \frac{1}{\gamma} \left( \frac{p(1-\gamma)}{c} - \frac{r + \lambda}{q(\theta)} \right) \right] / d\theta = \frac{(r + \lambda)q'(\theta)}{\gamma q(\theta)^2} < 1 \text{ for all } \theta. \text{ Other variables are recursively determined, and we can verify the existence and uniqueness of equilibrium. The next period’s urban productivity is updated by equation (10): } p' = p + g(l_t + u_t, p), \text{ and we can define the next period’s equilibrium by repeating the above process at a given value of } p'. \text{ With a positive agglomeration externality, we obtain a greater } p' \text{ with } p' > p. \text{ Appendix A shows that a greater } p' \text{ leads to a decrease in the non-urban population, suggesting a higher urban concentration. The differences in urban population and other variables between periods } t \text{ and } t+1 \text{ are due to migration. The other periods’ equilibria, including the possible long-run steady state, are similarly defined (the economics behind the transitional growth path will be discussed in the next subsection).}

Next, the growth rate slows down if the following condition is satisfied:

\[ \frac{d((p' - p)/p)}{dt} = \frac{d(g/p)}{dt} < 0. \]

That is, if the ratio of \( g \) to \( p \) decreases over time (e.g., marginal agglomeration benefits \( g \) declines or stays constant over time), the growth rate of urban production slows down. If the property of \( g \) is such that it stays positive and converges to zero, then productivity rises with time, so does urbanization, which proves the latter part of Proposition 1: the growth rate of urban production slows down as urbanization advances over time. ■

The diminishing \( g \) over time seems intuitively plausible because congestion also grows with agglomeration. Furthermore, urban productivity growth based on further agglomeration faces its limit because the urban concentration rate cannot exceed 100%.\(^{16}\) In short, rural-urban.

\(^{16}\) This is because “additional” technical and pecuniary externalities get smaller with urban concentration while additional congestion externalities get larger. At the long-run interior urbanization rate, if any, \( g(l_t + u_t, p') = 0 \) should hold, implying no further agglomeration-based growth.
migration proceeds in response to the wage gap as in HT, which in turn boosts productivity through agglomeration benefits and generates economic growth. However, this type of growth is likely to face its limit as urbanization advances. If this is the case, we can say that agglomeration generates a “transitory” growth effect. Looked at another way, our model can account for the HT equilibrium along the equilibrium growth path defined above.

**Proposition 2. Productivity, wage and unemployment:** Agglomeration-based high urban productivity combined with matching frictions in the urban sector can generate the equilibrium urban vs. non-urban wage and unemployment gaps under flexible wages and free mobility.

**Proof.** Equation (18) can be rearranged to the following equation: 

\[ \frac{w_R}{w} = \frac{\partial \eta(\theta)}{r + \lambda + \partial \eta(\theta)}, \]

which is bounded between zero and unity. From this, we obtain the wage gap result, \( w(t) > w_R(t) \). The wage gap arises on the consumption side from job destruction and the subsequent time-consuming search, which are the sources of the unemployment gap.

Next, we show that behind the wage gap there lies the sectoral productivity differential arising from agglomeration externalities on the production side. Proposition 2 says that the gap is also based on the following implicit condition that urban average labor productivity \( p(t) \) is greater than its comparable non-urban counterpart \( \frac{\beta f(u_z(t))}{u_z(t)} : p(t) > \frac{\beta f(u_z(t))}{u_z(t)} = w_R(t) \). To see this, we substitute \( w \) from equation (16) into equation (18) to obtain:

\[
\frac{w_R}{p - c(r + \lambda) / q(\theta)} = \frac{\partial \eta(\theta)}{r + \lambda + \partial \eta(\theta)}.
\]

(18')

After inverting the above equation, we rearrange it as follows to show that the urban-to non-urban labor productivity ratio is greater than unity, i.e., \( p(t) > w_R(t) \):
This exercise shows that behind the wage differential, there exists the sectoral productivity differential. Rearranging the first line of equation (20), we can express the wage differential as a function of the labor productivity differential, \( p / w_R \):

\[
\frac{p}{w_R} = \frac{r + \lambda + \theta q(\theta)}{\theta q(\theta)} + \frac{c(r + \lambda)}{w_R} \tag{20} \\
= 1 + \left( \frac{r + \lambda}{\theta q(\theta)} + \frac{c(r + \lambda)}{w_R} \right) > 1.
\]

Since the last term in the second line of (20') is negative, the wage differential should be smaller than the labor productivity differential, implying that the productivity differential is the underlying source of the wage gap. The productivity differential is accounted for by agglomeration benefits.

Next, the unemployment gap across sectors arises because migrants search for an urban job and there exist search frictions in urban job matching, so frictional unemployment occurs.\(^{17}\) Migration from the rural to urban sector is motivated by the wage gap, which is again based on the productivity gap. Migration continues until the equilibrium gap is established. The productivity gap between sectors is based on the agglomeration externality, and this gap is an equalizing gap that is supported by migration and wage and unemployment gaps.

\(^{17}\) If there is no unemployment gap, every rural worker moves to the urban sector and finds a job immediately, which is not possible due to search frictions in the urban sector.
Lemma 1. Ordering of Incomes: In equilibrium the current period incomes, \( w(t), \ w_R(t), \) and \( b(t), \) can be ordered as follows:

\[
w(t) > w_R(t) > b(t).
\] (21)

**Proof.** The inequality \( w(t) > w_R(t) \) holds according to Proposition 1: compensating differentials always lead to \( w(t) > w_R(t) \). From the Nash bargaining theory, it is obvious that \( w(t) > b(t) \). Now, what is left is whether the condition \( w_R(t) > b(t) \) really holds. To show this directly, we re-express equation (11-5) with the benefit level \( b \) un-normalized as follows:

\[
r U = b + \theta \cdot q(\theta) (W - R).
\] (11 – 5’)

Substituting equation (11 – 5’) and \( r R = w_R(u_z) \) (equation (11-6)) into the inter-regional equilibrium condition \( U = R \), we can derive the equation: \( b + \theta \cdot q(\theta) (W - R) = w_R(u_z) \), which is rearranged to give an inequality:

\[
\frac{w_R - b}{\theta \cdot q(\theta) (W - R)} > 0.
\] (11-5’’)

Given that the job finding rate, \( \theta \cdot q(\theta) \), is positive and \( W - R \) is also positive in equilibrium (i.e., having an urban job is better than having a non-urban job; see (11-8)), we find that \( w_R(t) > b(t) \) in equilibrium. Putting the results together, our model can account for the empirical regularity of \( w(t) > w_R(t) > b(t) \).

In short, the empirical regularity of \( w(t) > w_R(t) > b(t) \) is consistent with our model. Looking at this pattern, the HT model postulates that migration proceeds in response to differences in urban and rural “expected” incomes. In their model, the wage gap can thus be interpreted as a “compensating wage differential” for high urban unemployment. Our model, however, can explain the same phenomenon in a more structural framework with less restrictive
assumptions. The next subsection conceptualizes the transition path toward the long-run steady state.

2. The transition path

Given the nature of productivity growth through urban agglomeration, we can obtain the equilibrium growth path or the transition path toward the long-run equilibrium. Defining the equilibrium path requires connecting the competitive equilibrium points described earlier in this paper.

In this subsection, we discuss the economics behind the transition path. As in Arrow (1962), each agent is assumed to be small enough to neglect its contribution to aggregate urban productivity. The endogenous productivity gain through urbanization is therefore unanticipated by private agents at the time of their decision making, but given their non-rival nature, agglomeration benefits are eventually shared by agents at no cost in the form of improvement in production technology.

Based on these features, we first define equilibrium for a given level of productivity starting at (the beginning of) a particular period $t$, and then define new equilibrium at a point in time for later periods as follows. Because agglomeration leads to a productivity improvement, we need to redefine the new equilibrium in the next period in accordance with productivity updating.\(^{18}\) We summarize the discussion above using Proposition 3, the proof of which is shown in Appendix A in the proof of Proposition 1.

**Proposition 3. Urbanization and productivity:** In our model urbanization advances with

\(^{18}\) Here, we implicitly assume that urban productivity at the initial period is higher than the rural counterpart. Otherwise, endogenous technological growth is not possible.
urban productivity.

**Proof.** See the last part of Appendix A. Equation (A-6) establishes that the urban population is an implicit function of urban productivity, \(1 - u_2(l_i + u_i) = h(p)\) with \(h'(p) > 0.\)

This result can be taken erroneously as if urbanization *per se* causes productivity growth. In our model, it is agglomeration benefits that bring such a result.

Overall, the results in Section III offer an opportunity to view labor migration, unemployment and economic growth in a more structural setting. Apparently, the wage gap is a reason for labor mobility, but our model studies why and how urban and non-urban wages change and predicts that migration occurs as an adjustment toward the long-run equilibrium in response to unanticipated agglomeration-driven productivity changes. Further, unemployment arises due to search frictions, and it also determines the size of migration to the urban sector.

3. The steady state

Without loss of generality, we can focus on the interior solution where the long-run equilibrium urbanization rate is less than 100%. The long-run steady state is attained where productivity growth stops and no further migration to the urban sector occurs. This corresponds to the case where \(g(l_i + u_i)^*, p^*\) = 0 at \(0 < (l_i + u_i)^* < 1\) in the long run. At this level of \(p^*\), we can define the rest of the steady-state conditions as follows.

Using Proposition 3, we can define a long-run stationary equilibrium under some conditions on the agglomeration externality \(g\). To show this, we can express \(g\) as a function of \(p\), \(g = G(p): g = g(l_i + u_i, p) = g(h(p), p) = G(p)\) where we make use of Proposition 3: that the urban population is a function of urban productivity in equilibrium. Applying the contraction mapping principle to the urban productivity equation (10), we obtain that there exists the long-
run stationary equilibrium corresponding to $p^*$ if $\Delta G(p) < \Delta p' - \Delta p$ where $\Delta$ is the time difference operator,\(^{19}\) for all domain of $p$ with $G(p^*) = 0$ and $0 < l_i(p^*) + u_i(p^*) < 1$. Roughly put, if the changes in agglomeration benefits, $\Delta G(p)$, are smaller than the changes in productivity growth, $\Delta p' - \Delta p$, then the long-run stable equilibrium can arise.

However, if $g$ stays positive at any level of urban concentration, the economy will move to the virtually perfect urban concentration rate $l_i + u_i = 1$, where we can have a long-run equilibrium path with a non-zero growth rate, as in the usual endogenous economic growth theory that focuses on either a one-sector economy or cities only.

**Proposition 4. Conditions for the long-run stationary equilibrium:** The long-run interior stationary competitive equilibrium at $p^*$ with imperfect urban concentration arises if $\Delta G(p) < \Delta p' - \Delta p$ for all domain of $p$ with $G(p^*) = 0$ and $0 < l_i(p^*) + u_i(p^*) < 1$; if $G(p) > 0$ for all domains of $p$, the perfect urban concentration (i.e., $l_i + u_i = 1$) with a non-zero growth rate occurs.

**Proof.** A straightforward application of the contraction mapping principle proves the first part of Proposition 4. For the latter part, $G(p) > 0$ for all domains of $p$ implies that $p$ continues to grow according to (10), generating non-zero endogenous growth indefinitely. This leads to a corner solution of the perfect urban concentration, $l_i + u_i = 1$, with a non-zero growth rate. ■

4. **Optimal growth path**

Discussions on the social optimality involve (i) internalizing externalities in our model and (ii) correcting for the possible distortions in the competitive resource allocations. First, conceptually there are two different kinds of externalities in the model: the first one is the

\(^{19}\) The continuous time version is $dG/dt < d^2p/dt^2$. 

23
agglomeration externality, and the second one is the two-sided search externalities that a
worker’s job search affects other workers’ search due to congestion externality; and similarly, a
firm’s search for the right worker creates a similar kind of congestion for other firms. Second,
differences in the job-matching process across sectors imply sectoral differences in job-
matching costs and entail underutilization of scarce resources in the urban sector, i.e.,
unemployment, although unemployment is a natural outcome of (voluntary) migration.

To address the optimal growth path, we propose the utilitarian social planner’s
problem, \(^\text{20}\) which is a standard dynamic programming problem to model an economy’s optimal
production in a multi-period setting. In the Bellman equation below, we express the net (flow)
production as the sum of urban output \(y\) and non-urban output \(u_2^\beta\) minus the recruitment cost
to the firm, \(cv = c\theta u_1\).

\[
V(y, u_1) = \max_{\theta, u_2} \left\{ \frac{1}{1 + r} \left[ y + u_2^\beta - c\theta u_1 + V(y', u'_1) \right] \right\}
\]

subject to (i) \(y' = \partial q(\theta)u_1 p' + (1 - \lambda) y \left( \frac{p'}{p} \right)\),

(ii) \(u'_1 = \lambda \cdot (1 - u_1 - u_2) + (1 - \partial q(\theta)) \cdot u_1\),

(iii) \(p' = p + g(1 - u_2, p)\).

Since the right-hand side of the Bellman equation is a contraction map for all \(r > 0\), a unique
solution exists for the value function \(V(y, u_1)\). The first constraint is the dynamic of urban
output, suggesting that the next period’s urban output is equal to the newly starting firms’ output,
\(\partial q(\theta)u_1 p'\), plus surviving firms’ output that grows due to agglomeration economies arising from
added workers in accordance with the productivity updating constraint (iii), \((1 - \lambda) y(p'/p)\).

Constraint (ii) is the urban unemployment dynamic, which includes job destruction and job

---

\(^{20}\) Given that individuals maximize their wealth, which comes from production in the general equilibrium setting, the utilitarian social optimum is achieved when we maximize the present value of overall production in all periods.
The first-order conditions are given by:

\[
(V_y p' - V_{u_1}) \frac{d\theta q(\theta)}{d\theta} - c = 0, \quad (23-1)
\]

\[
V_y \left[ \theta q(\theta)u_1 \frac{dp'}{du_2} + (1 - \lambda) y \frac{d(g/p)}{du_2} \right] - \lambda V_{u_1} + \beta u_2^{\beta - 1} = 0. \quad (23-2)
\]

And the envelope conditions are:

\[
(1 + r)V_y = 1 + (1 - \lambda)(1 + g/p)V_y, \quad (24-1)
\]

\[
(1 + r)V_{u_1} = -c \theta + V_{u_1} \left[ -\lambda + (1 - \theta q(\theta)) \right] + V_y \left[ \theta q(\theta) p' \right]. \quad (24-2)
\]

Next, we manipulate the equations to derive some meaningful equations. Using \( V_y = V_{y'} \), we obtain

\[
V_y = \frac{p}{p(r + \lambda) - (1 - \lambda)g}. \quad \text{Again, using } V_{u_1} = V_{u'_1}, \text{ we obtain}
\]

\[
V_{u_1} = -pg_{u_2} - c\theta \left[ (1 + \lambda)g + p(r + \lambda) \right] + \frac{p^2 \theta q(\theta)}{[(1 + \lambda)g + p(r + \lambda)]r + \lambda + \theta q(\theta)]. \quad \text{And then, substituting these equations into}
\]

the first-order conditions (23-1) and (23-2), we finally obtain:

\[
c = \left[ \frac{g_{u_2} + p(r + \lambda)}{[(1 + \lambda)g + p(r + \lambda)][r + \lambda + \theta q(\theta) - \theta (d\theta q(\theta)/d\theta)]} \right], \quad (25-1)
\]

\[
u_2 = \left[ \frac{\beta((1 - \lambda)g + p(r + \lambda))}{(1 - \lambda)g_{u_2} + \frac{\lambda(c(1 - \lambda)g\theta - p(g_{u_2} + c(r + \lambda)\theta - p\theta q(\theta)))(1 - \lambda)g_{u_2} + \frac{\lambda(c(1 - \lambda)g\theta - p(g_{u_2} + c(r + \lambda)\theta - p\theta q(\theta)))}{r + \lambda + \theta q(\theta)}} }{1 - \beta} \right]. \quad (25-2)
\]

Along with the optimal values of \( \theta \) and \( u_2 \) in (25-1) and (25-2), the other two variables \( y \) and \( u_1 \) are described by the first two constraints in (22), and we thus have four variables and four equations. To attain the socially optimal allocation of resources across regions, we solve for \{ \theta, u_2 \}.  

\footnote{This is because the discount factor is equal to the interest rate in our model (see Bellman equation (22)).}
Implications of the social optimality condition

By comparing equations (25-1) and (25-2) with equations (14) and (20), respectively, we see the obvious differences between the social optimum and the competitive equilibrium. The optimality conditions given in (25-1) and (25-2) take into account not only the two-sided search externalities within the urban labor market but also both agglomeration externalities and underutilization of resources in the urban sector due to the costly matching process. However, the competitive equilibrium is derived without consideration of them. To attain the socially optimal values of \{ \theta, u_t \} from the competitive equilibrium, we thus need to implement a combination of two policy instruments \{ t, s \}; positive values of \( t \) and \( s \) represent a tax and a subsidy, respectively. With a tax-subsidy scheme, we can improve social welfare by (i) internalizing the externalities in both agglomeration and job matching, and (ii) correcting for the underutilization of resources as well.

We obtain the following equations when we implement a tax \( t \) on the competitive output \( p \) of the urban firm and a subsidy \( s \) to competitive non-urban wages.

\[
c = \frac{(1-t^*) p (1-r)}{r + \lambda + \gamma \partial \theta \theta}, \quad (26-1)
\]

\[
u_2 = \left[ \frac{(1+s^*) \beta (r + \lambda + \partial \theta \theta)}{p \partial \theta \theta - c \theta (r + \lambda)} \right]^{1-\beta} \cdot (26-2)
\]

While we can derive the exact closed-form solution for an optimal tax-subsidy scheme, the solution looks complicated, so we rather attempt to describe its properties at some interesting points and draw useful implications. We will first examine the optimal long-run stationary equilibrium where agglomeration externalities are exhausted.
a. The case of no agglomeration economies: the long-run stationary equilibrium

The optimality conditions shown above can be surprisingly simplified in the optimal long-run stationary equilibrium where agglomeration economies converge to zero, so that we are essentially at a static setting. Then we do not have to consider agglomeration externality $g$, so we get rid of the terms involving $g$ and $g_{u2}$ from (25-1) and (25-2). The resulting optimality equations (25-1) and (25-2) above are simplified to:

\[
c = \frac{p\eta}{r + \lambda + (1 - \eta)\partial q(\theta)}, \quad \eta \equiv d \ln[\partial q(\theta)]/d \ln(\theta); \quad (25-1')
\]

\[
u_2 = \left[\frac{\beta(r + \lambda + \partial q(\theta))}{\lambda(-c\theta + p\partial q(\theta)/(r + \lambda))}\right]^\frac{1}{1-\beta}.
\]

The above result suggests that the competitive equilibrium described above can achieve the socially optimal allocation of resources when (i) $\eta = 1 - \gamma$ with $\eta \equiv d \ln[\partial q(\theta)]/d \ln(\theta)$, i.e., the elasticity of the job finding rate with respect to $\theta$ equals $1 - \gamma$ – the worker’s share of match surplus, and (ii) the wage subsidy to the non-urban sector at the rate of $s^* = r/\lambda$ with no tax on urban output, $t^* = 0$. Put differently, condition (i), the so-called Hosios (1990) condition, combined with the non-urban wage subsidy, can achieve social optimality. Therefore, even when the search and recruitment externalities in the urban sector cancel out under the usual Hosios condition $\eta = 1 - \gamma$, the competitive equilibrium involves inefficiently many unemployed workers and inefficiently many firms with job vacancies as a result, a reason for the subsidy for non-urban wages. The intuition is this. In the competitive equilibrium, non-urban workers make their migration decisions knowing the probabilistic state-contingent possibilities of being matched or unmatched to an urban firm. But to search for an urban job, they have to move to the
urban sector as unemployed workers; unemployed job searchers cannot produce at all until next period, which, from the perspective of the social planner, is seen as a waste of scarce resources. For this reason, the social planner can improve overall efficiency by relocating some of the unmatched urban workers to the non-urban sector. Given that the discussion so far is based on the long-run steady state with no agglomeration effects, the previous point implies in turn that the competitive long-run stationary equilibrium involves over agglomeration: more specifically, over unemployment and under employment.

\[ b. \text{The case of agglomeration economies: along the transition path} \]

With agglomeration externalities present in the process of economic development, however, the equilibrium transition path depends on the specific form of function \( g \). An interesting transition path among others (henceforth, the “usual” case) would be that the economy moves toward the competitive long-run stationary point such that the following standard conditions are satisfied: (i) agglomeration externalities exist but decrease over time with \( g \to 0 \), \( G'(p) < 0 \) and \( G''(p) < 0 \) for all domains of \( p \), (ii) positive agglomeration externalities outweigh other negative externalities and underutilization of resources that arise from migration in the early stage of development, while this inequality reverses in the later stage (so interior LR equilibrium exists), and (iii) the Hosios (1990) condition is satisfied for brevity in discussion.

With the Hosios (1990) condition satisfied, we can focus on the relative importance of the two counteracting factors only, i.e., the net positive agglomeration externalities and the underutilization of scarce resources. Compared with the case of no agglomeration externalities, here we should consider the salient feature of the early development process where \( g \) is sufficiently large. The optimal fiscal policy should therefore be to subsidize not only the non-
urban sector but also the urban migrants/firms because we are better off with internalizing the agglomeration externalities from migration into the urban sector and also with correcting for the underutilization of scarce resources due to costly job matching in the urban sector. Thus, we may allow some moderate level of seemingly excessive migration and unemployment when urban productivity increases rapidly through migration during the early development process.

Discussions

With an optimal tax-subsidy scheme properly implemented, the equilibrium transition path can reach an optimal urban concentration rate at a point in time where the social marginal gains from migration to the urban sector balance with the social marginal costs, i.e., forgone rural production and the underutilization of labor due to costly urban job matching.

Along the equilibrium transition path toward the competitive long-run stationary point, migration generates externalities but migrants and new urban firms are not compensated for their mobility and entry, respectively, in a decentralized economy. Not only wage subsidies to the non-urban sector but also subsidies to the migrants/urban firms for the additional agglomeration benefits caused by them would be warranted, such that the economy can grow along the optimal transitional growth path. This would hasten the growth path through urban concentration, as the turnpike theorem suggests. Looked at another way, not only can we improve the competitive long-run equilibrium but we can also choose the fastest growth path. In this context, we can define the exact scope of government intervention, but it does not imply that a large urban sector per se generates economic growth.\(^{22}\) The optimal agglomeration level depends on the economic stage, labor market conditions, agglomeration effects, etc. This paper

---

\(^{22}\) The implication for growth happens to be intuitively similar in spirit to the infant industry protection argument: subsidy to new businesses may be warranted in the early stages of development when the growth rate is high with migration-driven urban agglomeration, but it may not be necessary in a fully developed stage of the economy.
suggests that it is more appropriate to create an environment where agglomeration benefits can be maximized so that aggregate productivity grows fast with urban agglomeration.

We compare the policy implications from our paper with those from the literature. The common feature of the optimal policies in the early HT models is that because a high urban wage causes too high urban unemployment, it is necessary to implement policies that attract firms to the urban sector in order to increase urban employment opportunities and reallocate urban job searchers to the rural sector. Specifically, Harris and Todaro (1970) argue that no single policy measure can attain the socially optimal resource allocation, and they propose a policy combination: a wage subsidy policy in the urban sector and a labor-mobility restriction policy. Later, Bhagwati and Srinivasan (1974) show that a uniform wage subsidy without migration restriction can yield the optimal, first best solution. A subsequent study by Basu (1980) relaxes the condition for the Bhagwati and Srinivasan result to hold. Compared with these existing studies, this paper lends some partial support to the prediction that the rural population in the long-run competitive equilibrium tends to be lower than the optimal level, so we need to subsidize rural labor; however, we may also need to subsidize urban labor in the development process where positive agglomeration externalities are substantially high. We summarize the discussions so far by Proposition 5.

**Proposition 5. Differential agglomeration policy warranted in different economic stages:** With the Hosios condition satisfied, the competitive long-run stationary equilibrium involves overagglomeration, so subsidies to the non-urban sector would be needed. However, in the development process where positive agglomeration externalities are substantially high, not only subsidies to the non-urban sector but also subsidies to migrants and urban firms for fostering urban agglomeration would improve social welfare.
IV. Summary and Conclusions

Based on agglomeration externalities, we build an endogenous growth model that is consistent with a set of empirical regularities – economic growth transitorily involves rural-to-urban migration, an inter-regional wage gap, urban unemployment, and the resulting urban concentration. Higher urban agglomeration through migration can provide transitory urban productivity growth. When this feature is combined with the assumption that the urban sector exhibits the properties of greater frictions in job matching, the wage and unemployment gaps between rural and urban sectors arise during the urbanization process.

Our analysis of the optimal growth path provides some implications regarding urban concentration. This paper theorizes the optimal concentration for economic growth, a notion explored in the empirical work by Henderson (2003). Approaching the optimal concentration issue as the result of accumulated migration to the urban sector, we show that rural-to-urban migration is not socially optimal if the negative externalities from urban job search by migrants combined with the subsequent underutilization of scarce resources dominate the positive externalities from recruitment and agglomeration as well.

Based on these findings, we may be able to propose two policy recommendations. First, given that agglomeration externalities vary depending on the level of urban concentration, moderately rapid migration to the urban sector would be justified in the early stages of economic development where marginal benefits from agglomeration economies are relatively high. Second, if the existing high level of urban agglomeration is fostered by aggressive policy interventions but the level of agglomeration is stagnant with little productivity growth, the current agglomeration can be said to be excessively high.

To better explain the nature of migration and urban agglomeration, perhaps we need to
include the housing market in the model because the high price of urban housing is also a crucial element of regional migration and is left for future research. In a similar context, including the capital markets in the model seems a promising avenue. Finally, it should be noted that this paper does not consider consumption-side issues such as amenities and disamenities associated with agglomerated cities.
References


Squire, Lyn, Employment Policy in Developing Countries: A Survey of Issues and Evidence, Published for the World Bank by Oxford University Press, 1981.


Williamson, J., “Migration and Urbanization,” in Handbook of Development Economics, vol. 1,
Appendix A. Proof for part of Proposition 1: a greater \( p' \) leads to a decrease in non-urban population, suggesting a higher urban concentration.

First, using equation (15) we obtain the following derivative:
\[
\frac{d\theta}{dp} = \frac{1 - \gamma}{\gamma c} / A > 0, \quad \text{where} \quad A \equiv 1 - \frac{(r + \lambda)q'(\theta)}{\gamma q(\theta)^2} > 0. \tag{A-1}
\]

This property in turn leads to an increase in the job arrival rate \( \theta q(\theta) \) and a decrease in the worker arrival rate \( q(\theta) \). Using these results, we can verify that \( w \) rises with \( p \) in equation (13) (see (A-2) below). Finally, we can find from equation (19) that the non-urban population goes down as \( p \) rises. To prove this in (A-5), we derive the partial derivatives (A-3) and (A-4) from equation (19) and use (A-1) and (A-2) as well.

\[
\frac{dw}{dp} = \gamma \left( 1 + c \cdot \frac{d\theta}{dp} \right) > 0, \tag{A-2}
\]

\[
\frac{\partial u_2}{\partial \theta} = -\frac{1}{1 - \beta} u_2^{\beta} \left( \frac{\beta w(r + \lambda) \frac{d\theta}{d\theta}}{[w\theta q(\theta)]^2} \right) < 0 \tag{A-3}
\]

\[
\frac{\partial u_2}{\partial w} = -\frac{1}{1 - \beta} u_2^{\beta} \left( \frac{\beta(r + \lambda + \theta q(\theta))}{w^2 \theta q(\theta)} \right) < 0 \tag{A-4}
\]

\[
\frac{du_2}{dp} = \frac{\partial u_2}{\partial \theta} \frac{d\theta}{dp} + \frac{\partial u_2}{\partial w} \frac{dw}{dp} < 0 \tag{A-5}
\]

---

23 See footnote 9.

24 \( d\theta q(\theta) / d\theta = q(\theta)(1 + \varepsilon_{q,\theta}) > 0 \) is derived because the elasticity of \( q(\theta) \) is bounded between -1 and 0 in the CRS matching technology: \( -1 \leq \varepsilon_{q,\theta} \leq 0 \).
\[ \frac{d(1-u_2)}{dp} = -\frac{du_2}{dp} > 0. \] (A-6)

The last line proves Proposition 3 that urbanization advances with urban productivity.