CIRJE-F-99

Labor Mobility and Economic Geography

Takatoshi Tabuchi
University of Tokyo

Jacques-François Thisse
CORE, Université catholique de Louvain

January 2001

Labor Mobility and Economic Geography*

Takatoshi Tabuchi† and Jacques-François Thisse‡

15 January 2001

Abstract

This paper investigates the impact of the heterogeneity of the labor force on the spatial distribution of activities. This goal is achieved by applying the tools of discrete choice theory to an economic geography model. We show that taste heterogeneity acts as a strong dispersion force. We also show that the relationship between the spatial distribution of the industry (the wage differential) and trade costs is smooth and $\cap$-shaped. Finally, while Rawlsian equity leads to the dispersion of industry, our analysis reveals that efficiency leads to a solution close to the market outcome, although the latter is likely to involve too much agglomeration compared to the former.

1 Introduction

Ever since the work of Krugman (1991), there has been a growing literature stressing the fact that the secular decline in trade costs (Bairoch, 1988), broadly defined to include all impediments to the exchange of goods, favors the emergence of a core-periphery structure in which all manufacturing firms would be geographically concentrated (see, e.g. Fujita, Krugman and Venables, 1999, for an extensive overview). Yet, this result has been criticized

*We thank Yasusada Murata for helpful comments. The second author acknowledges support from the Ministère de l'éducation, de la recherche et de la formation (Communauté française de Belgique), Convention 00/05-262.

†Faculty of Economics, University of Tokyo.

‡CORE, Université catholique de Louvain (Belgium), CERAS, Ecole nationale des ponts et chaussées (France) and CEPR.

by many because the setting used in this literature suffers from several severe limitations that cast doubt on the predictive power of the model. More precisely, by ignoring most of the costs imposed by space, the core-periphery model would remain in the tradition of international trade theory and, thus, would fail to provide an accurate description of the working of a spatial economy.

This state of affairs has led to several extensions. Some of them are reviewed in Ottaviano and Puga (1998), while Ottaviano, Tabuchi and Thisse (2001) use an alternative model allowing for the study of a broader set of issues. However, one critical assumption common to all existing models of economic geography is that individuals have the same preferences. Although this assumption is not uncommon in economic modeling, *it seems highly implausible that all potentially mobile individuals will react in the same way to a given "gap" between regions*. First of all, it is well known that some people show a high degree of attachment to the region where they are born. They will stay put even though they may guarantee to themselves higher living standards in other places. In the same spirit, life-time considerations such as marriage, divorce and the like play an important role in the decision to migrate (Greenwood, 1985). Second, regions are not similar and exhibit different natural and cultural features. Clearly, people value differently local amenities and such differences in attitudes are known to affect the migration process (Roback, 1982). The simple recognition that individuals are heterogenous in their perception of regional differences is, therefore, sufficient to invite us to revisit the core-periphery model.

Although the standard assumption of a priori identical regions is convenient to isolate the pure effects generated by the interplay between the agglomeration and dispersion forces, it does not permit us to study *the impact of differential amenities*. This is our second modification of the core-periphery model: amenity levels need not be the same across regions. Indeed, our model allows for a simple determination of the market outcome even when regions have different amenities. In such a context, both the market outcome and the optimum are asymmetric and it is worthwhile exploring their difference.

To sum up, we consider a setting in which *potentially mobile workers may choose to stay put because of extraneous considerations*. These considerations are fundamental ingredients of the migration decision and should be accounted for explicitly in workers' preferences. Even though the personal motivations may be quite diverse and, therefore, difficult to model at the individual level, Miyao (1978), Ginsburgh, Papageorgiou and Thisse (1985) and Tabuchi (1986) have argued that it is possible to identify their aggregate impact on the spatial distribution of economic activities by using discrete choice theory. Specifically, we assume in this paper that the "matching" of mobile workers' with regions is expressed through the *binary logit* (McFadden, 1974).[1] This assumption turns out to be empirically relevant in migration modeling (see, e.g. Anderson and Papageorgiou, 1994), while it is analytically convenient without affecting the qualitative nature of our main results.[2] Finally, regions need not be a priori identical in terms of amenities. Yet, as will be seen, it will be useful to consider the case of symmetric regions because it yields very neat results.

Previewing our main results, we show that *taste heterogeneity is a strong dispersion force* that drastically affects the conclusions inferred from the core-periphery model. More precisely, as trade costs steadily decrease to zero, the equilibrium configuration typically involves, first, dispersion, then partial agglomeration and, last re-dispersion. In other words, the relationship between the spatial distribution of the industry and trade costs is $\cap$-shaped. Furthermore, we will see that the more heterogenous the population of mobile workers, the more likely the dispersed configuration. In particular, even when agglomeration occurs, it involves a weaker degree of concentration of firms and workers than the core-periphery model, while the agglomeration process is gradual instead of exhibiting a bang-bang behavior. All these results strike us as being more plausible than existing ones and show how heterogeneity in workers' attitudes toward migrations have a profound impact on the spatial distribution of industry and, therefore, on the nature of trade. Still, our model yields the usual core-periphery structure, but only in the limiting case in which mobile workers are homogenous. At the other extreme, dispersion is the only market outcome when heterogeneity among workers is strong. Finally, our welfare analysis reveals that the efficient configuration displays a pattern similar to that of the market equilibrium, although both excessive or insufficient agglomeration may arise. However, the general trend looks like

[1]The logit model has already been used by Anas (1990) to model residential choice within a city.

[2]It is worth noting that this modeling strategy provides us with a first reconciliation between economic geography and spatial interaction theory. Although this branch of regional science was relegated by Krugman (1995) in the "five lost traditions" of economic geography, it is our contention that it remains a lively and promising research domain. Thus, connecting the two fields is likely to be relevant from both the theoretical and empirical points of view.

too much agglomeration in equilibrium.

We find these results especially relevant for the following reason. Once individual welfare level gets sufficiently high through the steadily increase of income, workers tend to pay more attention to the non-market attributes of their environment (see also Glaeser, Kolko and Saiz, 2001). Typically, they exhibit idiosyncratic tastes about such attributes. In this context, our results suggest that the spatial pattern of production in post-industrial societies might well differ from what it has been in the industrial societies we have known (Geyer and Kontuly, 1996; MacKellar and Vining, 1995). We return to this important topic in the concluding section.

The remaining of the paper is organized as follows. The model is presented in the subsequent section. The market outcome, with asymmetric and symmetric regions, is analyzed in Section 3. Interregional differential in nominal wages are examined in relation to the empirical literature in Section 4. The comparison between the equilibrium outcome and the first best optimum is conducted in Section 5 for the cases of asymmetric and symmetric regions. Section 6 concludes.

2 The model

The economic space is made of two regions, denoted H and F. There are two factors, denoted A and L. Factor A is evenly distributed across regions and is spatially immobile. Factor L is mobile between the two regions and $\lambda \in [0, 1]$ denotes the share of this factor located in region H. For expositional purposes, we refer to sector A as 'agriculture' and sector L as 'manufacturing'. Accordingly, we call 'farmers' the immobile factor A and 'workers' the mobile factor L.[3] There are two goods in the economy. The first good is homogenous. Consumers have a positive initial endowment of this good which is also produced using factor A as the only input under constant returns to scale and perfect competition. This good can be traded freely between regions and is chosen as the numéraire. The other good is a horizontally differentiated product; it is supplied by using L as the only input under increasing returns to scale and monopolistic competition.

[3] The reader who finds this interpretation nowadays irrelevant may think of the economy as being formed by a modern sector and a traditional sector, using respectively high-skilled and low-skilled workers.

Preferences about the differentiated product and the numéraire are identical across individuals. They are described by a quasi-linear utility with a quadratic subutility symmetric in all varieties:

$$U(q_0; q(i), i \in [0, N]) = \alpha \int_0^N q(i)di - \frac{\beta - \gamma}{2} \int_0^N [q(i)]^2 di \qquad (1)$$
$$- \frac{\gamma}{2} \left[\int_0^N q(i)di \right]^2 + q_0$$

where $q(i)$ is the quantity of variety $i \in [0, N]$ and q_0 the quantity of the numéraire. The parameters are such that $\alpha > 0$ and $\beta > \gamma > 0$. In (1), α expresses the intensity of preference for the differentiated product, whereas $\beta > \gamma$ means that consumers have a love of variety. Finally, for a given value of β, the parameter γ expresses the substitutability between varieties: the higher γ, the closer substitutes the varieties. We use a quasi-linear utility that abstracts from general equilibrium income effects for analytical convenience. Although this modeling strategy gives our framework a fairly strong partial equilibrium flavor, it does not remove the interaction between product and labor markets, thus allowing us to develop a full-fledged model of agglomeration formation, independently of the relative size of the industrial sector.

Any worker is endowed with one unit of labor and $\overline{q}_0 > 0$ units of the numéraire. Her budget constraint can then be written as follows:

$$\int_0^N p(i)q(i)di + q_0 = w + \overline{q}_0$$

where w is the worker's wage and $p(i)$ the price of variety i. The initial endowment $\overline{q}_0$ is supposed to be sufficiently large for the equilibrium consumption of the numéraire to be positive for each individual. This assumption allows us to focus on interior solutions only and is consistent with the idea that each worker is interested in consuming both types of goods.

Turning to the supply side, technology in agriculture requires one unit of A in order to produce one unit of output. With free trade in agriculture, the choice of this good as the numéraire implies that in equilibrium the wage of the farmers is equal to one in both regions, that is, $w_H^A = w_F^A = 1$. Technology in manufacturing requires ϕ units of L in order to produce any amount of a variety, i.e. the marginal cost of production of a variety is set equal to zero. This simplifying assumption, which is standard in many models of industrial organization, entails no loss of generality when firms' marginal costs are incurred in the numéraire. We assume that there is continuum N of potential firms and denote by λ_r the fraction of workers (L) living in region r, with $\lambda_H + \lambda_F = 1$.

There are no scope economies so that, due to increasing returns to scale, there is a one-to-one relationship between firms and varieties. Each variety can be traded at a positive cost of τ units of the numéraire for each unit transported from one region to the other, regardless of the variety, where τ accounts for all the impediments to trade. Since each firm sells a differentiated variety, it faces a downward sloping demand. Each firm has a negligible impact on the market outcome in the sense that it can ignore its influence on, and hence reactions from, other firms. However, aggregate market conditions of some kind (here average price across firms) affects any single firm. This provides a setting in which individual firms are not competitive (in the classic economic sense of having infinite demand elasticity) but, at the same time, they have no strategic interactions with one another. The demand for variety $i \in [0, N]$ may be shown to be given by:

$$\begin{aligned} q(i) &= a - bp(i) + c\int_0^N [p(j) - p(i)]dj \\ &= a - (b + cN)\, p(i) + cP \end{aligned} \tag{2}$$

where

$$P \equiv \int_0^N p(i)di$$

which can be interpreted as the price index in the modern sector, while $a \equiv \alpha/[(\beta + (N-1)\gamma]$, $b \equiv 1/[\beta + (N-1)\gamma]$ and $c \equiv \gamma/(\beta - \gamma)[\beta + (N-1)\gamma]$. The indirect utility corresponding to the demand system (2) is as follows:

$$\begin{aligned} V(w; p(i),\, i \in [0, N]) = \frac{a^2 N}{2b} - a\int_0^N p(i)di + \frac{b + cN}{2}\int_0^N [p(i)]^2 di \\ - \frac{c}{2}\left[\int_0^N p(i)di\right]^2 + w + \overline{q}_0 \end{aligned}$$

Let n_r be the number of firms in region r. Labor market clearing implies that

$$n_r = \frac{\lambda_r L}{\phi} \tag{3}$$

hence,

$$L = \phi N$$

so that the total number of firms in the economy is $N = 1/\phi$. Since workers and firms move together, immigrants do not displace native workers but start new businesses and create jobs for the region. As in Krugman (1991), however, the model allows for the endogenous determination of the spatial distribution of the industry but not for its total size.

In accordance with empirical observations, we assume that each firm is able to set a price specific to the market in which its product is sold, that is, markets are segmented (Greenhut 1981; Head and Mayer, 2000; McCallum, 1995). Hence, the profits made by a firm in region $r = H, F$ are defined as follows:

$$\Pi_r = p_{rr}q_{rr}(p_{rr})(A/2 + \lambda_r L) + (p_{rs} - \tau)q_{rs}(p_{rs})(A/2 + \lambda_s L) - \phi w_r$$

where w_r stands for the wage prevailing in region r.

Each firm i in region r maximizes its profit Π_r, assuming accurately that its price choice has no impact on the regional price indices

$$P_r \equiv \int_0^{n_r} p_{rr}(i)di + \int_0^{n_s} p_{sr}(i)di \qquad s \neq r$$

Since, by symmetry, the prices selected by the firms located within the same region are identical, the result is denoted by $p^*_{rr}(P_r)$ and $p^*_{rs}(P_s)$. Clearly, it must be that

$$n_r p^*_{rr}(P_r) + n_s p^*_{sr}(P_r) = P_r$$

Given (3), it is then readily verified that the equilibrium prices are as follows:

$$p^*_{rr} = \frac{1}{2}\frac{2a + \tau c n_s}{2b + cN} \qquad s \neq r \tag{4}$$

$$p^*_{rs} = p^*_{ss} + \frac{\tau}{2} \qquad s \neq r \tag{5}$$

For these prices to be meaningful, trade of varieties must be profitable to both firms and consumers. In other words, it must be that τ does not exceed some threshold given by

$$\tau_{trade} \equiv \frac{2a\phi}{2b\phi + cL}$$

Finally, entry and exit are free so that profits are zero in equilibrium. Hence, (3) implies that any change in the population of workers located in one region must be accompanied by a corresponding change in the number of firms. The equilibrium wage w_r^* of the workers living in region r is obtained from the zero profit condition evaluated at the equilibrium prices (4) and (5):

$$w_r^* = [(b + cN)(p_{rr}^*)^2(A/2 + \lambda_r L) + (b + cN)(p_{rs}^* - \tau)^2(A/2 + \lambda_s L)]/\phi$$

3 Agglomeration or dispersion under heterogenous tastes

Workers are free to live in either region. However, as discussed in the introduction, they are heterogenous in their perception of the attributes and characteristics associated with a particular region, while they are affected by different sorts of extraneous considerations. Such personal variations in tastes being unobservable, discrete choice theory suggests to model individual idiosyncrasies by assuming that the actual matching value between a worker and region $r = H, F$ is the realization of a random variable ε_r (Anderson, de Palma and Thisse, 1992, ch.3). In what follows, we assume that the ε_r are identically and independently distributed across individuals according to the double exponential with zero mean and a variance equal to $\pi^2\mu^2/6$. Assuming that the ε_r are i.i.d. implies that choices are governed by the same probability distribution whereas tastes are stochastically uncorrelated. However, actual choices may differ across workers. The fact that the distribution is the double exponential involves little restriction in the case of two regions, while allowing for simple and neat expressions.

From now on, it is convenient to set $\lambda_H \equiv \lambda$ and $\lambda_F \equiv 1 - \lambda$. Let $V_r(\lambda)$ be the indirect utility associated with the differentiated product and the numéraire in region r. Then, the probability that a worker will choose to reside in region r is given by the logit formula:

$$\mathbb{P}_r(\lambda) = \frac{\exp[V_r(\lambda)/\mu]}{\exp[V_r(\lambda)/\mu] + \exp[V_s(\lambda)/\mu]} \tag{6}$$

In (6), μ expresses the dispersion of individual tastes: the larger μ, the more heterogenous the workers' tastes about their living place. When $\mu = 0$, workers are homogenous and behave as in Ottaviano *et al.* (2001).

Assume also that all workers agree that region H has a higher level of amenities (d_H) than region F (d_F). Without loss of generality, we assume that the amenity differential between the two regions is such that:

$$d \equiv d_H - d_F \geq 0$$

For reasons that will become clear later on, H (F) is called the large (small) region.

We know from Ottaviano *et al.* that $V_r(\lambda)$ is a parabola whose quadratic term is the same for both regions. As a consequence, after some tedious calculations, we obtain

$$\Delta V(\lambda) \equiv V_H(\lambda) - V_F(\lambda) = C^* \tau(\tau^* - \tau)(\lambda - 1/2) + d$$

in which

$$C^* \equiv [2b\phi(3b\phi + 3cL + cA) + c^2 L(A+L)] \frac{L(b\phi + cL)}{2\phi^2(2b\phi + cL)^2} > 0$$

$$\tau^* \equiv \frac{4a\phi(3b\phi + 2cL)}{2b\phi(3b\phi + 3cL + cA) + c^2 L(A+L)} > 0$$

In what follows, we assume that $\tau_{trade} > \tau^*$ in order to be able to describe a complete agglomeration process. This amounts to assuming

$$A/L > \frac{6b^2\phi^2 + 8bc\phi L + 3c^2L^2}{cL(2b\phi + cL)} > 3 \tag{7}$$

where the second inequality holds because $b/c = \beta/\gamma - 1 \in (0, +\infty)$. Indeed, when (7) does not hold, the population of workers gets more and more dispersed as trade costs keep decreasing, as in Helpman (1998), so that no agglomeration occurs.

In the present setting, it should be clear that the population of workers changes according to the following equation of motion:

$$\frac{d\lambda}{dt} = (1-\lambda)\mathbb{P}_H(\lambda) - \lambda \mathbb{P}_F(\lambda) \tag{8}$$

where the first term in the RHS of (8) stands for the fraction of people migrating into region H, while the second term represents those leaving this region for region F.

A spatial equilibrium λ^* arises when $d\lambda/dt = 0$. Since the denominator is the same in both $\mathbb{P}_r(\lambda)$, $d\lambda/dt = 0$ is equivalent to $(1-\lambda)\exp[V_H(\lambda)/\mu] = \lambda \exp[V_F(\lambda)/\mu]$. Taking the logarithm of both sides, we may rewrite the equilibrium condition associated with (8) as follows:

$$\begin{aligned} J(\lambda; \tau) &\equiv \Delta V(\lambda) - \mu \log \frac{\lambda}{1-\lambda} \\ &= C^* \tau(\tau^* - \tau)\left(\lambda - \frac{1}{2}\right) + d - \mu \log \frac{\lambda}{1-\lambda} = 0 \end{aligned} \tag{9}$$

As a result, since $\mathrm{sgn}(d\lambda/dt) = \mathrm{sgn}[J(\lambda;\tau)]$, the stability condition of the system (8) is equivalent to

$$\partial J(\lambda^*; \tau)/\partial \lambda < 0$$

3.1 The case of asymmetric regions

The equation

$$\frac{\partial J(\lambda;\tau)}{\partial \lambda} = C^* \tau(\tau^* - \tau) - \frac{\mu}{\lambda(1-\lambda)} = 0$$

has two solutions with respect to λ. When both of them are real, denote the smaller solution by $\widetilde{\lambda}(\tau)$. Then, we have,

$$\widetilde{\lambda}(\tau^*/2) = \frac{1 - \sqrt{1 - \mu/\mu^*}}{2}$$

where

$$\mu^* \equiv \frac{C^*(\tau^*)^2}{16}$$

Assume that $J(\widetilde{\lambda}(\tau^*/2); \tau^*/2) \leq 0$ holds. Since $J(\widetilde{\lambda}(\tau); \tau)$ can be shown to be increasing (decreasing) in τ for $\tau > \tau^*/2$ ($\tau < \tau^*/2$), and since $\lim_{\tau \to 0} J(\widetilde{\lambda}(\tau); \tau) = \lim_{\tau \to \tau^*} J(\widetilde{\lambda}(\tau); \tau) = +\infty$, there exist τ_a^* and τ_b^* such that $J[\widetilde{\lambda}(\tau_a^*); \tau_a^*] = J[\widetilde{\lambda}(\tau_b^*); \tau_b^*] = 0$ with $0 < \tau_a^* \leq \tau^*/2 \leq \tau_b^* < \tau^*$.

The following result whose proof of which is given in appendix will be useful in characterizing the evolution of stable equilibria when τ decreases.

Lemma 1 *Consider two asymmetric regions ($d > 0$). Then, two cases may arise.*

(i) If

$$\mu^* > \mu \tag{10}$$

or

$$J(\widetilde{\lambda}(\tau^*/2); \tau^*/2) < 0 \tag{11}$$

does not hold, there exists a unique stable equilibrium belonging to $(1/2, 1)$.

(ii) If both (10) and (11) hold, the outcome is as follows: if $\tau > \tau_b^*$ or $\tau < \tau_a^*$, then there exists a unique stable equilibrium belonging to $(1/2, 1)$*; if $\tau_a^* < \tau < \tau_b^*$ there exist two stable equilibria, one belonging to* $(0, 1/2)$ *and the other to* $(1/2, 1)$.

In accordance with intuition, this lemma says that the region with the higher amenity level (i.e. region H) is usually the region with the larger industrial share. However, the region with the lower amenity level may also end up the larger industrial share. This multiplicity of equilibria arises when (i) trade costs take intermediate values ($\tau_a^* < \tau < \tau_b^*$) and (ii) both the degree of heterogeneity and the amenity differential are sufficiently small for the two regions not to be much differentiated in the workers' eyes. As in the homogenous case ($\mu = 0$), when trade are not high, agglomeration may arise in either region under weak heterogeneity (by a continuity argument). Nevertheless, when trade costs are sufficiently low, the existence of an amenity differential suffices to prevent an equilibrium to occur in the region endowed with low amenities. Finally, a low amenity differential means that both regions are close to being symmetric. All these results are illustrated in Figure 1, where the heavy curves describe the stable equilibria and the broken line the unstable equilibria.

Insert Figure 1 about here

Insofar as local stability is satisfied, the initially larger region H is always larger throughout the process of falling trade costs.

Proposition 1 *Consider two asymmetric regions ($d > 0$) and assume that the initial size of region H exceeds that of region F. Then, the size of region*

H is always larger than that of region F for any continuous decreases in trade costs. When τ steadily decreases, the size of region H grows for $\tau > \tau^*/2$ and declines for $\tau < \tau^*/2$.

Proof. We know from Lemma 1 that there exists exactly one equilibrium $\lambda^*(\tau)$ in the interval $(1/2, 1)$ for any $\tau \in [0, \tau_{trade}]$. Since $\lambda = 1/2$ is never an equilibrium and since $\lambda^*(\tau)$ changes continuously with τ, $\lambda^*(\tau)$ is always in $(1/2, 1)$.

The second part of the statement can be derived from the following inequalities:

$$\text{sgn}\left(\frac{\partial \lambda^*}{\partial \tau}\right) = \text{sgn}\left(-\frac{\partial J(\lambda^*;\tau)/\partial \tau}{\partial J(\lambda^*;\tau)/\partial \lambda}\right) = \text{sgn}\left(\frac{\partial \Delta V(\lambda^*)}{\partial \tau}\right) = \text{sgn}\,(\tau^* - 2\tau)\left(\lambda^* - \frac{1}{2}\right)$$

since $\partial J(\lambda^*, \tau)/\partial \lambda$ is always negative at any stable equilibrium. $\square$

As long as μ is positive, the existence of a stable equilibrium such as $1/2 < \lambda^* < 1$ implies that *the advantage in amenities matters and is reflected by the fact that the corresponding region is always larger than the other.* Furthermore, workers and firms never fully agglomerate within a single region. Hence, even a low degree of heterogeneity is sufficient to prevent the emergence of a standard core-periphery structure. Among other things, this implies the existence of intraindustry trade between the two regions, but flows are unequal since more firms are located in region H. In addition, λ^* does not exhibit any flat spot so that the change in size of the large (small) region is *smooth* and $\cap$-shaped ($\cup$-shaped). Thus, when workers are heterogenous, *the economy does not exhibit any catastrophic change such as those shown within the standard core-periphery model.*

Since $\Delta V(\lambda^*) \geq 0$ always holds, we have

$$\lambda^*|_{\tau=0} = \frac{\exp(d/\mu)}{1 + \exp(d/\mu)} \geq \frac{1}{2}$$

where the equality holds when $d = 0$. Moreover, it is readily verified that if the initial size of region H exceeds $1/2$, then

$$\lambda^*|_{\tau>0} > \lambda^*|_{\tau=0}$$

implying that the degree of agglomeration takes its lowest value when trade costs are zero. This result clearly shows the role of these costs in shaping the economic landscape.

Finally, we have:

$$\frac{\partial \lambda^*}{\partial \mu} < 0 \qquad \frac{\partial \lambda^*}{\partial L} > 0$$

Consequently, *the regional size differential gets narrower as the degree of heterogeneity is higher and the number of workers larger*, while the asymmetry in trading the differentiated product varies with the industrial share of region H. The former inequality can be explained as follows: more heterogeneity within workers means that prices, variety and wage matter less to them, thus fostering more dispersion since workers' matching values are drawn independently from the same distribution. In the limiting case where $\mu \to \infty$, we always have full dispersion because workers are willing to choose their place according to a fifty-fifty random choice rule. Stated differently, heterogeneity always benefits the region with the larger endowment, but simultaneously prevents this region from accommodating all workers. The latter inequality reflects the following idea: the larger the mass of workers within the economy, the more important their location for their well-being. Hence, as expected, the large region is larger when the mass of workers increases.

3.2 The case of symmetric regions

The case of symmetric regions leads to very neat results and is, therefore, worth considering. Clearly, symmetry ($\lambda^* = 1/2$) is always a steady-state for (8) since $J(1/2;\tau) = 0$. From (18), $J(\lambda;\tau)$ is convex (concave) with respect to λ for all $\lambda \in (0, 1/2)$ ($\lambda \in (1/2, 1)$), implying that there exists at most one asymmetric equilibrium (up to a permutation).

Computing $\partial J(\lambda;\tau)/\partial\lambda$, we obtain

$$C^*\tau(\tau^* - \tau) \gtreqless 4\mu \qquad \Longleftrightarrow \qquad \frac{\partial J(1/2;\tau)}{\partial \lambda} \gtreqless 0 \tag{12}$$

Hence, for a given admissible value of τ, the symmetric equilibrium is stable (unstable) if and only if the heterogeneity in tastes is sufficiently strong (weak). When $\mu = 0$, (12) implies that there is full dispersion when τ exceeds τ^*, while there is full agglomeration when τ is lower than τ^*. In other words, we get the standard core-periphery model when workers are homogenous.

Solving the quadratic equation $C^*\tau(\tau^* - \tau) - 4\mu = 0$ with respect to τ, we are able to determine the symmetry breaking threshold as follows. When $\mu = \mu^*$, the discriminant of this equation equals zero. Accordingly, if the heterogeneity is small in that $\mu < \mu^*$, then the quadratic equation $C^*\tau(\tau^* - \tau) - 4\mu = 0$ has two real and distinct real roots given by

$$\tau_1^*, \tau_2^* = \frac{\tau^*}{2} \pm \sqrt{\frac{(\tau^*)^2}{4} - \frac{4\mu}{C^*}}$$

It is easy to check that $0 < \tau_1^* \leq (\tau_1^* + \tau_2^*)/2 = \tau^*/2 \leq \tau_2^* < \tau^* < \tau_{trade}$. We thus have the following result.

Lemma 2 *In the case of two symmetric regions ($d = 0$), there exists a unique stable equilibrium (up to a permutation). It involves dispersion when $\tau \geq \tau_2^*$ or $\tau \leq \tau_1^*$ and partial agglomeration when $\tau_1^* < \tau < \tau_2^*$.*

This lemma implies that, in the intervals $[0, \tau_1^*]$ and $[\tau_2^*, \tau_{trade}]$, workers are dispersed when $\mu > 0$, whereas they would be agglomerated in the case of a homogenous population ($\mu = 0$) because $\tau_1^* = 0$ and $\tau_2^* = \tau^*$. Thus, *symmetry is more likely to occur when workers are heterogenous in their regional matching*.

We now describe how the stable equilibrium changes with the level of trade costs.

Proposition 2 *Consider two symmetric regions ($d = 0$) and let the trade costs to decrease steadily.*

(i) Assume that $0 < \mu < \mu^*$. When $\tau \geq \tau_2^*$, the economy involves full dispersion of the industry. When $\tau_2^* > \tau > \tau_1^*$, partial agglomeration of the industry arises. For $\tau_2^* > \tau > \tau^*/2$ the gap between the two regions widens, but narrows for $\tau^*/2 > \tau > \tau_1^*$. Finally, when $\tau \leq \tau_1^*$, the industry is again fully dispersed.

(ii) Assume that $\mu \geq \mu^*$. Then, the industry is fully dispersed for all admissible τ.

When heterogeneity μ is weak, the industry displays a three-stage pattern: dispersion, partial agglomeration, and re-dispersion, as shown in Figure 2. By contrast, when μ is large enough, there is always dispersion. This is already enough to show that *taste heterogeneity is a strong dispersion force*. But we can say more. It is readily verified that $\partial\tau_1^*/\partial\mu > 0$ and $\partial\tau_2^*/\partial\mu < 0$.

Hence, a higher degree of heterogeneity implies that the symmetry breaking threshold τ_1^* (τ_2^*) arises at a higher (lower) value of trade costs, thus implying that the domain for which partial agglomeration arises shrinks.

Insert Figure 2 about here

Case (i) in Proposition 2 shows the existence of a $\cap$-*shaped relationship* between the regional concentration of the industry and the level of trade costs. Such a relationship has already been obtained by Fujita *et al.* (1999, ch.7) and Ottaviano *et al.* (2001) but in different contexts so that it is worth comparing results. In the former case, the agglomeration of the manufacturing sector within, say, region H generates large imports of the agricultural good from region F. When transport costs in the manufacturing sector becomes sufficiently low, the price indices of this good is about the same in the two regions. Then, the relative price of the agricultural good in H rises as its transport cost remains unchanged. This in turn lowers region F's nominal wage that guarantees the same utility level in both regions to the workers. When the transport costs within the manufacturing sector decrease sufficiently, the factor price differential becomes strong enough to induce firms to move away from H to F. In other words, as transport costs in the manufacturing sector keep decreasing from high to very low values while transport costs in the agricultural sector remain constant, the manufacturing sector is first fully dispersed, then fully agglomerated and, last, re-dispersed.

In the latter, it is assumed that the concentration of firms and workers within a region takes place within a monocentric city: firms are located at the city center, while workers are dispersed around this center. In this context, workers also consume land (the lot size is supposed to be fixed and equal across workers, and the land rents go to absentee landlords) and commute to the city center where they work. If θ denotes the unit commuting cost, whether dispersion or agglomeration arises is determined by the sign of the expression $C^*\tau(\tau^*-\tau)-\theta L/2$ (compare this expression and (12)). Hence, the existence of commuting costs leads to re-dispersion for sufficiently low values of τ. By contrast, it is the sign of $C^*\tau(\tau^*-\tau)-4\mu$ that matters in the present model. When μ is sufficiently large, full dispersion arises because, trade costs being low enough, the matching of workers with a particular region matters more than anything else.

Replacing a, b and c by their values, we also see that $\mu^* = 0$ when there is no increasing returns ($\phi = 0$) as in Ottaviano *et al.* (2001). Thus, here also, we need increasing returns in order to trigger an agglomeration process.

4 Interregional wage differential

A straightforward computation yields the nominal wage differential:

$$w_H^* - w_F^* = \frac{(b\phi + cL)\,[2b\phi + c(A+L)]L}{2\phi^2\,(2b\phi + cL)}\tau(\tau^o - \tau)(\lambda^* - 1/2) \qquad (13)$$

where

$$\tau^o \equiv \frac{4a\phi}{2b\phi + c(A+L)}$$

The amenity differential d does not directly appear in (13). However, since λ^* is a function of d, the amenity differential affects the nominal wage differential.

If

$$\frac{\tau^*}{2} < \tau^o < \tau_{trade}$$

holds, then the nominal wage in the large region is lower (higher) than in the small one for $\tau > \tau^o$ ($\tau < \tau^o$). This leads to the following sequence. (i) When $\tau \in (\tau^o, \tau_{trade})$, the large region attracts workers although wages are higher in the small region; (ii) when $\tau \in (\tau^*/2, \tau^o)$, the large region still attracts workers, a result which seems more intuitive because wages are now higher in region H; (iii) when $\tau \in (0, \tau^*/2)$, one gets an unexpected result because the large region lose workers although wages in H are higher than in the small region, which now accommodates more workers. Such counter-intuitive flows have often been observed in the real world (Vining and Kontuly, 1978). However, they should not come as a real surprise since what drives migration is the utility differential and not the wage differential, which is just one part of the former. This also shows how misleading might be the comparison of regional living standards based on average incomes only.

Let us now focus on the nominal wage differential (13). We know from the RHS of (13) that the term $\tau(\tau^o - \tau)$ is increasing (decreasing) in $\tau < \tau^o/2$ ($\tau > \tau^o/2$), while $(\lambda^* - 1/2)$ is increasing (decreasing) in $\tau < \tau^*/2$ ($\tau > \tau^*/2$) from Proposition 1. Since $\tau^o/2 < \tau^*/2 < \tau_{trade}$ holds, we may conclude that for $\tau \in (\tau^o/2, \tau^*/2)$, when trade costs τ decrease over time, *the interregional nominal wage differential, first, increases for large values of* τ *and, then, decreases for small values of* τ.

This last result sheds some light on an old debate dealing with the spatial implications of economic development. In the development literature, a high degree of urban concentration together with a widening wage differential is expected to arise during the early phases of economic growth; as development proceeds, spatial deconcentration and a narrowing wage differential should occur because (i) the initial urban giants become highly congested and (ii) the mechanism of factor price equalization should proceed (Williamson, 1965; Alonso, 1980; Wheaton and Shishido, 1981). Hence, our results above provide a formal proof of the existence of such $\cap$-shaped relationships between economic development and the spatial distribution of activities, as well as between economic development and the interregional wage differential.

5 Efficiency vs. regional equity

Since preferences are quasi-linear and profits are zero, we may evaluate efficiency by using the sum of individual welfare across workers and farmers. In the homogenous case, the consumer surplus is simply given by the sum of indirect utilities over the two regions. However, once we introduce heterogeneity across workers, we must account for the fact that they now benefit from *intrinsic differentiation between regions*. To do so, we use Proposition 3.7 by Anderson *et al.* (1992), which gives us the utility level of a worker as a function of λ:

$$\mathbf{V}_0(\mathbb{P}_H) = \mathbb{P}_H V_H(\lambda) + \mathbb{P}_F V_F(\lambda) - \mu\left(\mathbb{P}_H \log \mathbb{P}_H + \mathbb{P}_F \log \mathbb{P}_F\right)$$

where the first two terms stands for the expected utility derived from living in region r with probability $\mathbb{P}_r$, whereas the last term corresponds to a "premium" associated with the presence of heterogeneity (observe that the terms in parentheses are negative). Since the probability $\mathbb{P}_H$ is the same for all workers, it must be that $\mathbb{P}_H = \lambda$ and $\mathbb{P}_F = 1 - \lambda$. This in turn implies that the maximum utility level of a worker may be rewritten as follows:[4]

$$\mathbf{V}(\lambda) = \lambda V_H(\lambda) + (1-\lambda) V_F(\lambda) - \mu\left[\lambda \log \lambda + (1-\lambda)\log(1-\lambda)\right] \tag{14}$$

[4]Equation (14) can also be derived from the maximization of the following individual welfare (see, e.g. Small and Rosen, 1981):

$$\mathbf{V}(\lambda) = \mu \log\left\{\exp[V_H(\lambda)/\mu] + \exp[V_F(\lambda)/\mu]\right\}$$

through interregional income transfers, Subject to the equilibrium condition $(1-\lambda)\mathbb{P}_H = \lambda\mathbb{P}_F$.

As a result, the global efficiency level associated with λ is given by

$$W(\lambda) \equiv L\mathbf{V}(\lambda) + \frac{A}{2}\left[V_H(\lambda) + V_F(\lambda)\right] \qquad (15)$$

The first best outcome may then be obtained by maximizing $W(\lambda)$ with respect to λ, all prices being set equal to marginal costs ($p_{rr} = 0$ and $p_{rs} = \tau$).

The first-order condition for efficiency ($W'(\lambda) = 0$) is given by

$$C^o \tau(\tau^o - \tau)\left(\lambda - \frac{1}{2}\right) + d - \mu \log \frac{\lambda}{1-\lambda} = 0 \qquad (16)$$

where

$$C^o = \frac{[2b\phi + c(A+L)]L}{\phi^2}$$

The optimality condition (16) is therefore similar to the equilibrium condition (9) except for the parameters C^o and τ^o. As a result, (16) has exactly one interior solution in the interval of $[1/2, 1)$, which is a local maximizer of (15). Moreover, since $d \geq 0$, it is readily verified that

$$\begin{array}{ll} W(1/2 + x) \geq W(1/2 - x) & \qquad \forall x \in [0, 1/2] \\ W'(1/2) \geq 0 & \qquad \forall \tau \in [0, \tau_{trade}] \end{array}$$

Thus, the social optimum λ^o is uniquely determined in the interval of $[1/2, 1)$ (up to a permutation in the special case where $d = 0$).

5.1 The case of asymmetric regions

When $d > 0$, the efficient fraction of workers in region H must be such that

$$\lambda^o > 1/2 \qquad \forall \tau \in (0, \tau_{trade}]$$

thus implying partial agglomeration in region H for any τ. As expected, the existence of an amenity differential is sufficient to prevent the symmetric configuration from being efficient. In addition, it is readily shown that $W'(0) = \infty$ and $W'(1) = -\infty$. Hence, the maximizer is necessarily interior. Put differently, *even when the amenity differential d is very large, full agglomeration is never socially desirable* insofar as the degree of heterogeneity μ is positive.

We also have

$$\frac{\partial W'(\lambda)}{\partial \tau} = 2C^o L(\tau^o/2 - \tau)(\lambda - 1/2)$$

Summarizing the above, we have the following result.

Proposition 3 *Assume that both regions are asymmetric ($d > 0$) and that workers are heterogenous ($\mu > 0$). Then, the efficient size of region H is larger than that of region F for all admissible τ. When τ steadily decreases, the efficient size of region H grows, without reaching the value 1, as long as $\tau > \tau^o/2$ and declines for $\tau < \tau^o/2$.*

Clearly, both Proposition 1 and Proposition 3 are similar but, as illustrated in Figure 3, the two patterns are not necessarily identical. The maximum size arises at $\tau^o/2$ in the efficiency case and at $\tau^*/2$ in the equilibrium case with $\tau^o/2 < \tau^*/2$.

Insert Figure 3 about here

Next, we compare the equilibrium and efficient distributions. From the LHS of (9) and of (16), we find that λ^* coincides with λ^o at $\tau = 0$ and $\tau = \tau^c$, where

$$\tau^c \equiv \frac{C^o\tau^o - C^*\tau^*}{C^o - C^*} > 0$$

Since $\operatorname{sgn}(\tau^c - \tau_{trade}) = \operatorname{sgn}(B)$, where

$$B \equiv 10b^3\phi^3 - 2b^2\phi^2 c(3A - 5L) - b\phi c^2 L(5A - L) - c^3 L^2(A + L)$$

$\lambda^* > \lambda^o$ must hold for $B > 0$. This occurs for high degree of product differentiation ($1/c$) and high fixed costs (ϕ). Hence, we have:

Proposition 4 *(i) When varieties are sufficiently differentiated and increasing returns are sufficiently high ($B > 0$), the equilibrium configuration is more concentrated than the optimal one for all admissible τ.*

(ii) When varieties are close substitutes and/or increasing returns are sufficiently low ($B < 0$), the equilibrium configuration is less concentrated than the optimal one for $\tau \in (\tau^c, \tau_{trade}]$ but more concentrated for $\tau \in (0, \tau^c)$.

Both cases (i) and (ii) are illustrated in Figure 3 where the dotted curve shows the optimum, and the heavy ones the stable equilibria. It is worth noting that the parameters τ^c, τ^*, τ^o and B are independent of μ, while the distributions λ^* and λ^o depend on the degree of heterogeneity μ. In modern economies, the degree of the product differentiation and the degree of scale economies tend to be large, whereas trade costs tend to be low. Thus, it seems fair to conclude that, in our economies, *the market outcome is likely to involve too much agglomeration.*

Finally, since $\Delta V(\lambda) > 0$ holds for all $\lambda \geq 1/2$ and $d > 0$, the welfare of farmers residing in region H is always higher than that of farmers living in F. This leads us to focus on an alternative social welfare function that has retained a lot of attention in regional planning, namely *spatial equity.* This amounts to maximizing the lowest welfare level in the economy:

$$\max_{\lambda} \min\{V_H(\lambda), V_F(\lambda)\} \tag{17}$$

given that the prices are set equal to marginal costs. Since $\Delta V(\lambda) > 0$ holds for any parameter values with $\lambda \geq 1/2$, we know that (17) is maximized at $\lambda^R < 1/2$. *When the social objective is to maximize a Rawlsian welfare function, there must be more workers in the region with less amenities.* Such a result is a sharp contrast to what we obtained above in the efficiency case. In a sense, the implementation of interregional income transfers, which are carried out in several industrialized countries, may be considered as an attempt at providing some reconciliation between efficiency and spatial equity.

5.2 The case of symmetric regions

When $d = 0$, we always have

$$W'(1/2) = 0$$

Solving the quadratic equation $W''(1/2) = 0$ with respect to τ, we obtain the following two roots:

$$\tau_1^o, \tau_2^o = \frac{\tau^o}{2} \pm \sqrt{\frac{(\tau^o)^2}{4} - \frac{4\mu}{C^o}}$$

with $\tau_1^o \leq \tau^o/2 \leq \tau_2^o$. They are real and positive if and only if

$$\mu < \mu^o \equiv \frac{C^o(\tau^o)^2}{16}$$

Hence, *when taste heterogeneity is sufficiently broad, dispersion is always socially efficient*. In other words, μ^o is the symmetry breaking threshold in the efficiency case. As in the equilibrium case, it is easy to see that $\mu^o = 0$ when $\phi = 0$. Thus, we need increasing returns for partial agglomeration to be socially efficient.

It is readily verified that $W(\lambda)$ reaches a local maximum (minimum) at $1/2$ if and only if $\tau > \tau_2^o$ or $0 \leq \tau < \tau_1^o$ ($\tau_1^o < \tau < \tau_2^o$). Furthermore, as shown in the previous subsection, there is a unique local maximum (up to a permutation) different from $1/2$ when $\tau_1^o < \tau < \tau_2^o$. We may then establish Proposition 5 for the efficiency case in the same way we did for Proposition 2, by replacing the superscript $*$ by o.

Proposition 5 *Consider two symmetric regions ($d = 0$) and let the trade costs to decrease steadily.*

(i) Assume that $0 < \mu < \mu^o$. When $\tau \geq \tau_2^o$, the efficient configuration involves full dispersion of the industry. When $\tau_2^o > \tau > \tau_1^o$, the industry is partial agglomerated. For $\tau_2^o > \tau > \tau^o/2$ the gap between the two regions widens, but narrows for $\tau^o/2 > \tau > \tau_1^o$. Finally, when $\tau \leq \tau_1^o$, the optimum is again fully dispersed.

(ii) Assume that $\mu \geq \mu^o$. The optimum involves full dispersion for all admissible τ.

The efficient configuration in case (i) as a function of τ is qualitatively similar to the one displayed in Figure 2. Unfortunately, the comparison between the various distributions in case (ii) is not straightforward. Yet, we know that both $(\tau_1^o + \tau_2^o)/2 = \tau^o < \tau^* = (\tau_1^* + \tau_2^*)/2$ and $\tau_1^o < \tau_2^*$ hold. As a result, both the market equilibrium and the efficient outcome involve symmetry for sufficiently large degrees of heterogeneity μ (case (ii)) or for sufficiently low trade costs τ (case (i)). Otherwise, the market outcome is more dispersed or more agglomerated than the optimum, depending on the parameter values.

Finally, as expected, maximizing the Rawlsian welfare function (17) always yields full dispersion in the case of symmetric regions.

6 Concluding remarks

We have shown that *taste heterogeneity is a strong dispersion force* that dramatically affects the core-periphery structure. Typically, it gives rise to a

$\cap$-shaped relationship between the spatial concentration of industry and the level of trade costs. In our setting, the dispersed equilibrium is generally asymmetric and the region with the high level of amenity is larger than the other one. In addition, the evolution of the equilibrium pattern of the industrial sector no longer involves a catastrophic change. As a consequence, the global economy would follow a three-stage process, involving dispersion, agglomeration, and re-dispersion, which is continuous with respect to variations in the trade cost.

In the introduction, we have argued that welfare differentials based on market consumption goods and wages have decreased across industrial regions as economic development has proceeded. This implies that nonmarket factors, such as the attachment of people to their region of origin or the presence of specific amenities, have a growing importance in individual decisions to migrate (formally, the values of μ and d rise). Indeed, once they have reached some living thresholds, workers are less willing to trade their family and social environment against more individual consumption. This is especially well illustrated by the fact that, within the European Union, traditional migration flows such as those from Southern to Northern Italy have stopped during the last decades. By contrast, there has been an increase in the mobility of high skilled workers within the European Union and Japan. This can be explained by the following two factors. First, education generates human capital which is easily transferable to another region and eases the search for job, residency and a social environment to live in. In other words, the marginal rate of substitution between the social environment and individual consumption would be larger for the skilled than for the unskilled (in our setting, the skilled would have a lower μ than the unskilled). Second, regional wage differentials are likely to be larger for the skilled than for the unskilled (Black, 2000). These two forces put together make migration more attractive for the former than for the latter. All of this then suggests an explanation for the fundamental changes observed in migration patterns within post-industrial economic areas such as the European Union or Japan. By the same token, this would provide us with an explanation for the rising concentration of high-level services within large metropolitan regions (Amiti, 1998; Fujita and Tabuchi, 1997).

To sum-up, this paper has shown that *taste heterogeneity is likely to be a critical force in shaping modern economic spaces.* Thus, omitting it in our analyses might lead to rough and incomplete pictures.

References

[1] Alonso, W. (1980) Five bell shapes in development, *Papers of the Regional Science Association* 45, 5-16.

[2] Amiti, M. (1998) New trade theory and industrial location in the EU: a survey of evidence, *Oxford Review of Economic Policy* 14, 45-53.

[3] Anas, A. (1990) Taste heterogeneity and spatial urban structure: the logit model and monocentric theory reconciled, *Journal of Urban Economics* 28, 318-35.

[4] Anderson, S.P., de Palma A. and J.-F. Thisse (1992) *Discrete Choice Theory of Product Differentiation*, Cambridge (Mass.), MIT Press.

[5] Anderson, W.P. and Y.Y. Papageorgiou (1994) An analysis of migration streams for the Canadian regional system, 1952-1983. 1. Migration probabilities, *Geographical Analysis* 26, 15-36.

[6] Bairoch, P. (1988) *Cities and Economic Development: From the Dawn of History to the Present*, Chicago, University of Chicago Press.

[7] Black, D. (2000) Local knowledge spillovers and inequality, University of California at Irvine, Department of Economics, mimeo.

[8] Fujita, M., P. Krugman and A.J. Venables (1999) *The Spatial Economy. Cities, Regions and International Trade*, Cambridge (Mass.), MIT Press.

[9] Fujita, M. and T. Tabuchi (1997) Regional growth and in postwar Japan, *Regional Science and Urban Economics* 27, 643-670.

[10] Geyer, H.S. and T.M. Kontuly (1996) *Differential Urbanization: Integrating Spatial Models*, London, Arnold.

[11] Ginsburgh, V., Y.Y. Papageorgiou and J.-F. Thisse (1985) On existence and stability of spatial equilibria and steady-states. *Regional Science and Urban Economics* 15, 149-158.

[12] Glaeser E.L., J. Kolko and A. Saiz (2001) Consumer city, *Journal of Economic Geography* 1, 27-50.

[13] Greenhut, M.L. (1981) Spatial pricing in the U.S.A., West Germany, and Japan, *Economica* 48, 79-86.

[14] Greenwood, M.J. (1985) Human migration: theory, models and empirical studies, *Journal of Regional Science* 25, 521-544.

[15] Head, K. and T. Mayer (2000) Non-Europe. The magnitude and causes of market fragmentation in the EU, *Weltwirtschaftliches Archiv* 136, 284-314.

[16] Helpman, E. (1998) The size of regions, in: D. Pines, E. Sadka and I. Zilcha, eds., *Topics in Public Economics. Theoretical and Applied Analysis*, 33-54, Cambridge, Cambridge University Press.

[17] Krugman, P. (1991) Increasing returns and economic geography, *Journal of Political Economy* 99, 483-499.

[18] Krugman P. (1995) *Development, Geography, and Economic Theory*, Cambridge (Mass.), MIT Press.

[19] MacKellar, F.-L. and D.R. Vining (1995) Population concentration in less developed countries: new evidence, *Papers in Regional Science* 74, 259-293.

[20] McCallum, J. "National Borders Matter: Canada-US Regional Trade Patterns," *American Economic Review* 85 (1995), 615-623.

[21] McFadden, D. (1974) Conditional logit analysis of qualitative choice behavior, in: P. Zarembka, ed., *Frontiers in Econometrics*, 105-142, New York, Academic Press.

[22] Miyao, T. (1978) A probabilistic model of location choice with neighborhood effects, *Journal of Economic Theory* 19, 347-358.

[23] Ottaviano, G.I.P. and D. Puga (1998) Agglomeration in the global economy: a survey of the 'new economic geography', *World Economy* 21, 707-731.

[24] Ottaviano, G.I.P., T. Tabuchi and J.-F. Thisse (2001) Agglomeration and trade revisited, *International Economic Review*, forthcoming.

[25] Roback J. (1982) Wages, rents, and the quality of life, *Journal of Political Economy* 90, 1257-1278.

[26] Small, K.A. and H.S. Rosen (1981) Applied welfare economics with discrete choice models, *Econometrica* 49, 105-130.

[27] Tabuchi, T. (1986) Existence and stability of city-size distribution in the gravity and logit models, *Environment and Planning A* 18, 1375-1389.

[28] Vining, D.R. and T. Kontuly (1978) Population dispersal from major metropolitan regions: an international comparison, *International Regional Science Review* 3, 49-73.

[29] Wheaton, W. and H. Shishido (1981) Urban concentration, agglomeration economies, and the level of economic development, *Economic Development and Cultural Change* 30, 17-30.

[30] Williamson, J. (1965) Regional inequality and the process of national development, *Economic Development and Cultural Change* 14, 3-45.

Proof of Lemma 1

Consider first $\lambda \in [1/2, 1]$, which corresponds to case (i) and the first part of case (ii). Since $J(1/2, \tau) > 0 > \lim_{\lambda\to 1} J(\lambda, \tau)$, $\lambda = 1/2$ and $\lambda = 1$ are not equilibria. From

$$\frac{\partial^2 J(\lambda, \tau)}{\partial \lambda^2} = \frac{\mu}{\lambda^2} - \frac{\mu}{(1-\lambda)^2} \tag{18}$$

it follows that $J(\lambda, \tau)$ is concave with respect to λ for all $\lambda \in (1/2, 1)$. Hence, there is exactly one equilibrium λ_3^* in the interval $(1/2, 1)$. Furthermore, it is stable since $\partial J(\lambda_3^*, \tau)/\partial\lambda < 0$.

Next, consider $\lambda \in [0, 1/2)$, which is the second part of case (ii). Since $\lim_{\lambda\to 0} J(\lambda, \tau) > 0$, $\lambda = 0$ is not an equilibrium. Consider the derivative

$$\frac{\partial J(\lambda, \tau)}{\partial \lambda} = C^*\tau(\tau^* - \tau) - \frac{\mu}{\lambda(1-\lambda)} = 0$$

Since the first term is maximized at $\tau = \tau^*/2$, and the second term is minimized at $\lambda = 1/2$, multiple equilibria are possible only if $\partial J(1/2, \tau^*/2)/\partial\lambda > 0$ holds, which is (10).

Since $\lim_{\lambda\to 0} J(\lambda, \tau^*/2) > 0$ and $J(1/2, \tau^*/2) > 0$, multiple equilibria arise only if $J(\widetilde{\lambda}(\tau^*/2), \tau^*/2) \leq 0$. From (18), $J(\lambda, \tau)$ is convex with respect to λ for all $\lambda \in (0, 1/2)$. Hence, there exist two equilibria, denoted by λ_1^* and λ_2^* with $0 < \lambda_1^* \leq \lambda_2^* < 1/2$, if $J(\widetilde{\lambda}(\tau^*/2), \tau^*/2) \leq 0$.

When $J(\widetilde{\lambda}(\tau^*/2), \tau^*/2) = 0$, the repeated root $\widetilde{\lambda}(\tau^*/2)$ is an unstable equilibrium since $\partial J(\widetilde{\lambda}(\tau^*/2), \tau^*/2)/\partial\lambda = 0$ and $\partial^2 J(\widetilde{\lambda}(\tau^*/2), \tau^*/2)/\partial\lambda^2 > 0$. Therefore, (11) is necessary for the equilibrium to be stable.

Similarly, when $J(\widetilde{\lambda}(\tau_a^*), \tau_a^*) = 0$ ($J(\widetilde{\lambda}(\tau_b^*), \tau_b^*) = 0$), the repeated root $\widetilde{\lambda}(\tau_a^*)$ ($\widetilde{\lambda}(\tau_b^*)$) is also an unstable equilibrium. Hence, two distinct equilibria exist in the interval $(0, 1/2)$ only if $\tau \in (\tau_a^*, \tau_b^*)$ under (10) and (11), where λ_1^* is stable and λ_2^* is unstable. $\square$

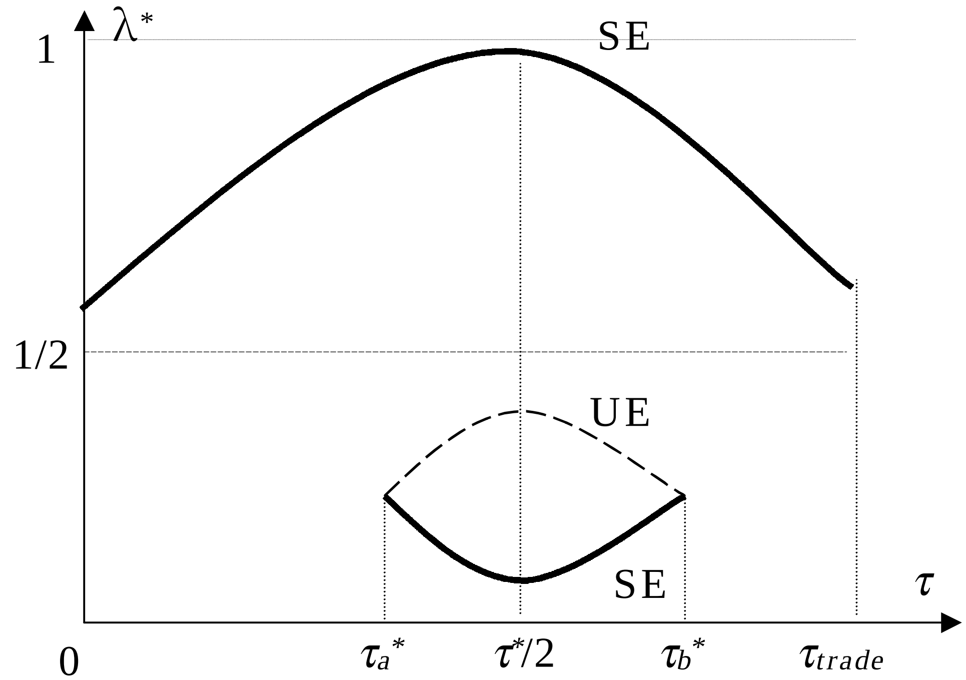

Figure 1 Equilibrium when $d > 0$

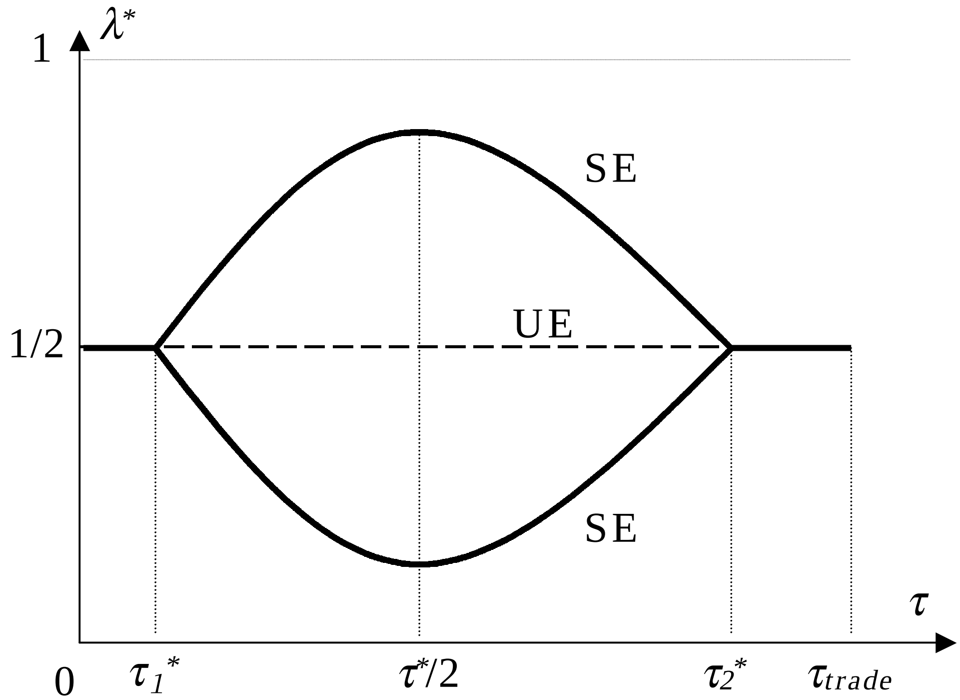

Figure 2 Equilibrium when $d = 0$

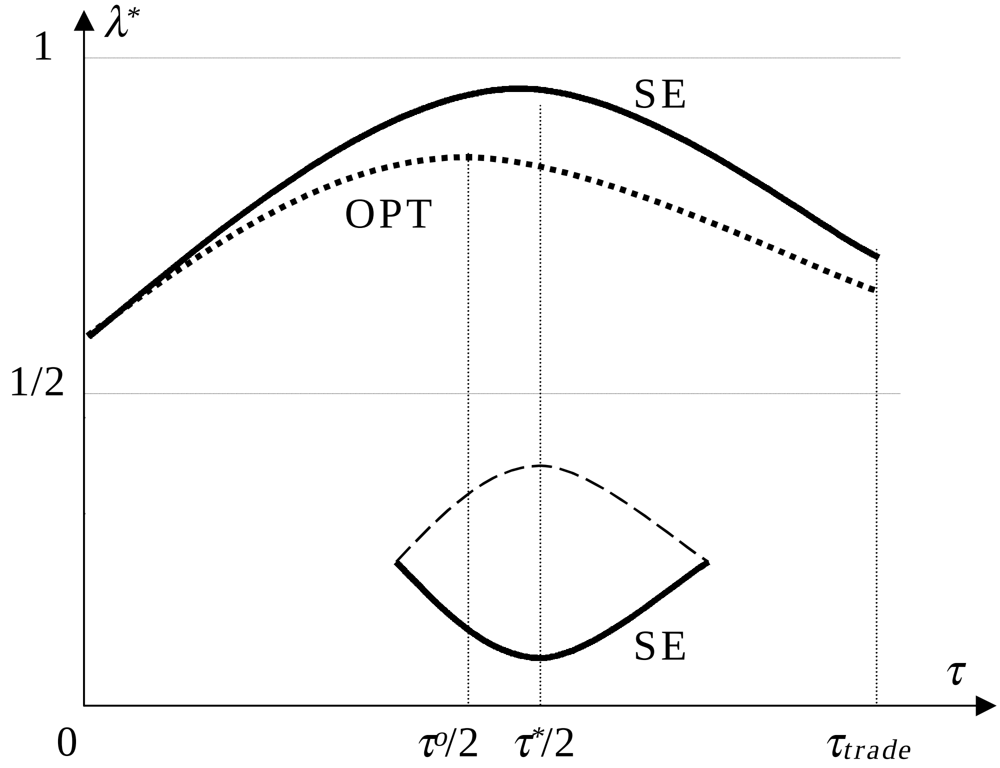

Figure 3(i) Equilibrium and social optimum when $d > 0$ and $B > 0$

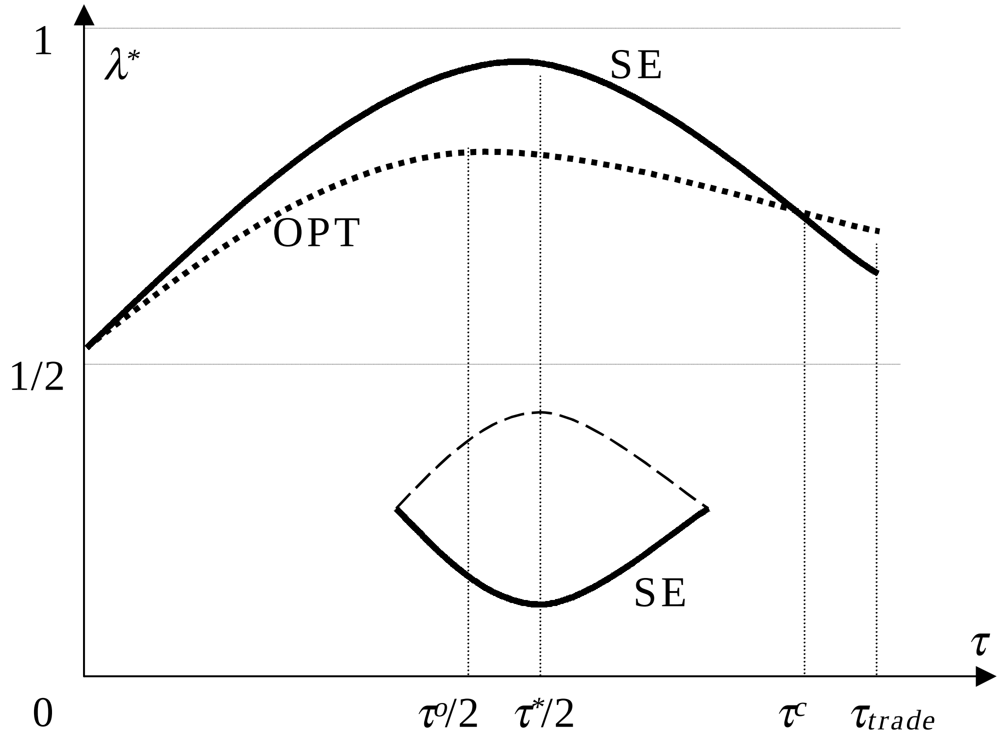

Figure 3(ii) Equilibrium and social optimum when $d > 0$ and $B < 0$