CIRJE-F-65

Agglomeration and Trade Revisited

Gianmarco Ottaviano
Università degli Studi di Bologna

Takatoshi Tabuchi
University of Tokyo

Jacques-François Thisse
Université catholique de Louvain

December 1999

Agglomeration and Trade Revisited*

Gianmarco Ottaviano[†] Takatoshi Tabuchi[‡]

Jacques-François Thisse[§]

7 December 1999

Abstract

The purpose of this paper is twofold. First, we present an alternative model of agglomeration and trade that displays the main features of the recent economic geography literature while allowing for the derivation of analytical results by means of simple algebra. Second, we show how this framework can be used to permit (i) a welfare analysis of the agglomeration process, (ii) a full-fledged forward looking analysis of the role of history and expectations in the emergence of economic clusters, and (iii) a simple analysis of the impact of urban costs on the spatial distribution of economic activities.

Keywords: agglomeration, trade, monopolistic competition, self-fulfilling expectations, urban costs.

J.E.L. Classification: F12, L13, R13.

*We should like to thank four referees for their comments as well as R. Baldwin, G. Duranton, M. Fujita, P. Jehiel, A. Linbeck, K. Middelfart-Knarvik, T. Mori, D. Peeters, P. Picard, D. Pines, K. Stahl, participants to seminars at Kyoto University, IUI in Stockholm, CEPR ERWIT in Rotterdam, SITE in Stanford, and ASSET in Bologna. The first author is grateful to the EU TMR programme, the second author to the Japan Society for the Promotion of Science through the Future Program, and the first and third authors to FNRS (Belgium) for funding. The revision was made when the first author was Jean Monnet Fellow at EUI Florence (Italy).

[†]Università degli Studi di Bologna, CORE, Université catholique de Louvain, and CEPR.

[‡]Faculty of Economics, University of Tokyo.

[§]CORE, Université catholique de Louvain, CERAS, Ecole nationale des ponts et chaussées, and CEPR.

1 Introduction

The agglomeration of activities in a few locations is probably the most distinctive feature of the economic space. Despite some valuable early contributions made by Hirschman, Perroux or Myrdal, this fact remained unexplained by mainstream economic theory for a long time. It is only recently that economists have become able to provide an analytical framework explaining the emergence of economic agglomerations in an otherwise homogenous space. As argued by Krugman (1995), this is probably because economists lacked a model embracing both increasing returns and imperfect competition, the two basic ingredients of the formation of the economic space, as shown by the pioneering work of Hotelling (1929), Lösch (1940) and Koopmans (1957).

However, even though several modeling strategies are available to study the emergence of economic agglomerations (Fujita and Thisse, 1996), their potential has not been really explored as recognized by Krugman (1998, p.164) himself:

> To date, the new economic geography has depended heavily on the tricks summarized in Fujita, Krugman and Venables (1999) with the slogan "Dixit-Stiglitz, icebergs, evolution, and the computer"

The slogan of the new economic geography is explained by the following methodology (see Fujita *et al.*, 1999). First, the main tool used in the new economic geography is a particular version of the Chamberlinian model of monopolistic competition developed by Dixit and Stiglitz (1977) in which consumers love variety and firms have fixed requirements for limited productive resources (hence, *"Dixit-Stiglitz"*). Love of variety is captured by a CES utility function which is symmetric in a bundle of differentiated products. Each firm is assumed to be a negligible actor in that it has no impact on overall market conditions. Second, transportation is modeled as a costly activity that uses the transported good itself: in other words, a certain fraction of the good melts on the way (hence, *"icebergs"*). Taken together, these assumptions yield a demand system in which the own-price elasticities of demands are constant, identical to the elasticities of substitutions and equal to each other across all differentiated products. This entails equilibrium prices that are independent of the spatial distribution of firms and consumers. Though convenient from an analytical point of view, such a result conflicts with research in spatial competition which shows that demand elasticity varies with distance while prices change with the level of demand and the intensity of competition. Moreover, the iceberg assumption also implies that any increase in the price of the transported good is accompanied by a proportional increase in its trade cost, which is unrealistic. Third, the stability analysis used to select spatial equilibria rests on myopic adjustment processes in which the location of mobile factors is driven by differences in current returns. Despite some analogy with evolutionary game theory (hence, *"evolution"*), this approach neglects the role of expectations (Krugman, 1991a; Matsuyama, 1991), which may be crucial for locational decisions since they are often made once-and-for-all.[1] Last, notwithstanding their simplifying assumptions, the models of the new economic geography are often beyond the reach of analytical resolution so that authors have to appeal to numerical investigations (hence, *"the computer"*).[2]

The purpose of this paper is to propose a complementary modeling strategy which allows us to go beyond some of the current limits of the new economic geography. In particular, we do that by presenting a model of agglomeration and trade that, while displaying the main features of the core-periphery model by Krugman (1991b), differs under several major respects. First, preferences are not CES in that we adopt an alternative specification of the preference for variety, namely the *quadratic utility* model, which is also popular in industrial organization (Dixit, 1979; Vives, 1990), in international trade (Anderson *et al.,* 1995; Krugman and Venables, 1990) as well as in demand analysis (Phlips, 1983). Moreover, while firms are still considered as negligible actors, we adopt a broader concept of equilibrium than the one in Dixit and Stiglitz (1977). Second, trade costs are assumed to absorb resources that are different from the transported good itself. Taken together, our specifications allow us to disentangle the economic meanings of the various parameters thus leading to clear-cut comparative static results that are likely to be easier to test than those based on Dixit and Stiglitz (1977).[3] They also entail elasticities of demand and substitution that vary with prices, while equilibrium prices now depend on all the fundamentals of the market.

[1] See, however, Ottaviano (1999) for the analysis of a special case.

[2] The most that can be obtained within this framework without resorting to numerical solutions has been probably achieved by Puga (1999). See also some chapters of Fujita *et al.* (1999).

[3] As an additional example, in Krugman (1991b) the same parameter turns out to measure not only the elasticities of demand and substitution but also (inversely) the returns to scale that remain unexploited in equilibrium.

Using this framework, we are able to derive analytically the results obtained by Krugman (1991b). Going beyond them, our setting allows us to provide a neat welfare analysis of agglomeration. While natural due to the many market imperfections that are present in new economic geography models, such an analysis is seldom touched due to the limits of the standard approach.[4] What we show is that the market yields agglomeration for values of the trade costs for which it is socially desirable to keep activities dispersed. Hence, while they coincide for high and low values of the trade costs, *the equilibrium and the optimum differ for a domain of intermediate values.* In this case, there is room for regional policy interventions grounded on both efficiency and equity considerations.

In addition, our framework allows us to study forward-looking location decisions and to determine the exact domain in which expectations matter for agglomeration to arise. Specifically, we show that expectations influence the agglomeration process in a totally unsuspected way in that they have an influence on the emergence of a particular agglomeration for intermediate values of the trade cost only. For such values, and only for them, *if* (for whatever reason) *workers expect the lagging region to become the leading one, their expectations will reverse the dynamics of the economy provided that the difference in initial endowments between the two regions is not too large.*

Finally, our model is sufficiently flexible to establish a bridge between the new economic geography and urban economics. We show that it can be easily extended to accommodate urban costs (Fujita, 1989). *This trade-off leads to a set of results richer than the core-periphery model since a $\cap$-shaped relationship emerges when the manufactured goods' trade costs decrease.*[5] Such a result confirms and extends preliminary explorations undertaken by Helpman (1998), Tabuchi (1998), and Puga (1999). It also agrees with the observations according to which some developed economies (especially the UK) would experience re-dispersion (Geyer and Kontuly, 1996).

The organization of the paper reflects what we have said in the foregoing. The model is presented in the next section while the equilibrium prices and wages are determined in Section 3 for any given distribution of firms

[4]The utilitarian approach is difficult to justify in the CES case because farmers and workers have different incomes, hence a different marginal utility for the numéraire. See, however, Krugman and Venables (1995) and Helpman (1998) for some numerical developments about the welfare implications of agglomeration in related models.

[5]Other reasons leading to a $\cup$-shaped relationship are discussed in Ottaviano and Puga (1998) and Fujita *et al.* (1999).

and workers. The process of agglomeration is analyzed in Section 4 by using the standard myopic approach in selecting the stable equilibria. In Section 5, we compare the optimum and market outcomes. In Section 6, we introduce forward-looking behavior and show how our model can be used to compare history (in the sense of initial endowments) and expectations in the emergence of an agglomeration. The impact of urban costs associated with the formation of an agglomeration is investigated in Section 7. Section 8 concludes.

2 The model

The economic space is made of two regions, called H and F. There are two factors, called 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. We want to stress the fact, however, that *the role of factor A is to capture the idea that some inputs* (such as land or some services) *are nontradable while some others have a very low spatial mobility* (such as low-skilled workers). Hence our model, as Krugman's one, should not necessarily be interpreted as an agriculture-oriented model.

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 imperfect competition.

Each firm in the manufacturing sector has a negligible impact on the market outcome in the sense that it can ignore its influence on, and hence reactions from, other firms. To this end, we assume that there is a continuum N of potential firms, so that all the unknowns are described by density functions. There are no scope economies so that, due to increasing returns to scale, there is a one-to-one relationship between firms and varieties. Since each firm sells a differentiated variety, it faces a downward sloping demand.

Formally, since there is a continuum of firms, each firm is negligible and the interaction between any two firms is zero. 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.

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.

Preferences are identical across individuals and described by a *quasi-linear utility with a quadratic subutility* which is supposed to be symmetric in all varieties (see the appendix for more details):

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

where $q(i)$ is the quantity of variety $i \in [0,N]$ and q_0 the quantity of the numéraire. The parameters in (1) are such that $\alpha > 0$ and $\beta > \gamma > 0$. In this expression, α expresses the intensity of preferences for the differentiated product, whereas $\beta > \gamma$ means that consumers are biased toward a dispersed consumption of varieties. Suppose, indeed, that an individual consumes a total mass of Nq of the differentiated product. If consumption is uniform on $[0,x]$ and zero on $(x,N]$, then the density on $[0,x]$ is Nq/x. Equation (1) evaluated for this consumption pattern is

$$\begin{aligned} U &= \alpha\int_0^x \frac{Nq}{x}di - \frac{\beta-\gamma}{2}\int_0^x \left(\frac{Nq}{x}\right)^2 di - \frac{\gamma}{2}\left[\int_0^x \left(\frac{Nq}{x}\right) di\right]^2 + q_0 \quad (2) \\ &= \alpha Nq - \frac{\beta-\gamma}{2x}N^2q^2 - \frac{\gamma}{2}N^2q^2 + q_0 \end{aligned}$$

which is strictly increasing in x since $\beta > \gamma$. Hence, regardless of the values of q and N, (2) is maximized at $x = N$ where variety consumption is maximal. We may then conclude that *the quadratic utility function exhibits love of variety as long as* $\beta > \gamma$. Finally, for a given value of β, the parameter γ expresses the substitutability between varieties: the higher γ, the closer substitutes the varieties.[6]

[6]When $\beta = \gamma$, substitutability is perfect. Indeed, (1) degenerates into a utility funtion that is quadratic in total consumption $\int_0^N q(i)di$, which is exactly what one would expect with a homogeneous product.

We use a quasilinear 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 manufacturing sector.

Any individual is endowed with one unit of labor (of type A or L) and $\overline{q}_0 > 0$ units of the numéraire. His budget constraint can then be written as follows:

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

where y is the individual's labor income, $p(i)$ is the price of variety i while the price of the agricultural good is normalized to one. The initial endowment $\overline{q}_0$ is supposed to be large enough for the residual consumption of the numéraire to be strictly positive for each individual. Observe that this assumption is consistent with any expenditure share on the differentiated product strictly smaller than one.

Solving the budget constraint for the numéraire consumption, plugging the corresponding expression into (1) and solving the first order conditions with respect to $q(i)$ yields

$$\alpha - (\beta - \gamma)q(i) - \gamma \int_0^N q(j)dj = p(i), \quad i \in [0, N]$$

Therefore, the demand for variety $i \in [0, N]$ is:

$$q(i) = a - bp(i) + c \int_0^N [p(j) - p(i)]dj \tag{3}$$

where $a \equiv \alpha/[(\beta + (N-1)\gamma]$, $b \equiv 1/[\beta + (N-1)\gamma]$ and $c \equiv \gamma/(\beta - \gamma)[\beta + (N-1)\gamma]$.[7] Increasing the degree of product differentiation among a given set of varieties amounts to decreasing c. However, assuming that all prices are identical and equal to p, we see that the aggregate demand for the differentiated product equals $aN - bpN$, which is independent of c. Hence (3) has the desirable property that the market size in the industry does not change

[7]Notice that when $\gamma = \beta$ the indirect demand cannot be inverted in terms of each variety's quantity $q(i)$ but only in terms of total quantity $\int_0^N q(i)di$. Once more, this is what one would expect with homogeneous products and explains why parameter c degenerates to infinity as γ tends to β.

when the substitutability parameter c varies. More generally, it is possible to decrease (increase) c through a decrease (increase) in the parameter γ in the utility U while keeping the other structural parameters a and b of the demand system unchanged. The own price effect is stronger (as measured by $b + cN$) than each cross price effect (as measured by c) as well as the sum of all cross price effects (cN), thus allowing for different elasticities of substitution between pairs of varieties as well as for different own elasticities at different prices.

The indirect utility corresponding to the demand system (3) is as follows:

$$V(y;p(i),\, i \in [0,N]) = -a\int_0^N p(i)di + \frac{b+cN}{2}\int_0^N [p(i)]^2 di - \frac{c}{2}\left[\int_0^N p(i)di\right]^2 + y + \overline{q}_0 \tag{4}$$

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. Clearly, ϕ is a measure of the degree of increasing returns in the manufacturing sector.

Labor market clearing implies that

$$n_H = \lambda L/\phi \tag{5}$$

and

$$n_F = (1-\lambda)L/\phi \tag{6}$$

Consequently, the total mass of firms (varieties) in the economy is fixed and equal to $N = L/\phi$. This means that, in equilibrium, ϕ can also be interpreted as an inverse measure of the mass of firms. As $\phi \to 0$ (or $L \to \infty$), the mass of varieties becomes arbitrarily large. In addition, (5) and (6) show that the region with the larger labor market is also the region accommodating the larger proportion of firms. In addition, (5) and (6) imply that any change in the population of workers located in one region must be accompanied by a corresponding change in the mass of firms.

As to equilibrium wages, they are determined as follows. Due to free entry and exit, profits are zero in equilibrium. As in Krugman (1991b), the equilibrium wages corresponding to (5) and (6) are determined by a bidding process between firms for workers, which ends when no firm can earn a strictly positive profit at the equilibrium market prices. In other words, all operating profits are absorbed by the wage bills.

Since trade costs are positive, firms have the ability to segment markets, that is, each firm is able to set a price specific to the market in which its product is sold. Indeed, even for very low trade costs, there are many good reasons to believe that firms want and will succeed to price discriminate between segmented markets (Horn and Shy, 1996; Thisse and Vives, 1988), as confirmed by empirical work on international markets (Head and Mayer, 2000; McCallum, 1995; Wei, 1996).

In the sequel, we focus on region H. Things pertaining to region F can be derived by symmetry. Using the assumption of symmetry between varieties and (3), demands faced by a representative firm located in H in region H (q_{HH}) and region F (q_{FF}) are given respectively by:

$$q_{HH} = a - (b + cN)p_{HH} + cP_H \tag{7}$$

and

$$q_{HF} = a - (b + cN)p_{HF} + cP_F \tag{8}$$

where

$$P_H \equiv n_H p_{HH} + n_F p_{FH}$$

$$P_F \equiv n_H p_{HF} + n_F p_{FF}$$

Clearly, P_H/N and P_F/N are the average prices prevailing in region H and F so that P_H and P_F can be interpreted as the corresponding price indices since N is fixed. Finally, the profits made by a firm in H are defined as follows:

$$\Pi_H = p_{HH}q_{HH}(p_{HH})(A/2 + \lambda L) + (p_{HF} - \tau)q_{HF}(p_{HF})[A/2 + (1 - \lambda)L] - \phi w_H \tag{9}$$

where $A/2$ stands for the number of farmers in each region and w_H for the wage prevailing in region H.

3 Short-run price equilibria

In this section, we study the process of competition between firms for a given spatial distribution of workers. Prices are obtained by maximizing profits while wages are determined as described above by equating the resulting profits to zero. Since we have a continuum of firms, each one is negligible in the sense that its action has no impact on the market. Hence, when choosing its prices, a firm in H accurately neglects the impact of its decision over the two price indices P_H and P_F. In addition, because firms sell differentiated varieties, each one has some monopoly power in that it faces a demand function with finite elasticity.

When Dixit and Stiglitz use the CES, the same assumption implies that each firm is able to determine its price independently of the others because the price index enters the demand function as a multiplicative term. This no longer holds in our model because the price index now enters the demand function as an additive term (see (7) and (8)). Stated differently, a firm must account for the distribution of the firms' prices through some aggregate statistics, given here by the price index, in order to find its equilibrium price. As a consequence, our market solution is given by a Nash equilibrium with a continuum of players in which prices are interdependent: *each firm neglects its impact on the market but is aware that the market as a whole has a non-negligible impact on its behavior.* As a result, the equilibrium prices will depend on key aspects of the market instead of being given by a simple relative mark-up rule.

Since profit functions are concave in own price, solving the first order conditions for profit maximization with respect to prices yields the equilibrium prices (denoted by *):

$$p^*_{HH} = \frac{1}{2}\frac{2a + \tau c(1-\lambda)N}{2b + cN} \tag{10}$$

$$p^*_{FF} = \frac{1}{2}\frac{2a + \tau c\lambda N}{2b + cN} \tag{11}$$

$$p^*_{HF} = p^*_{FF} + \frac{\tau}{2} \tag{12}$$

$$p^*_{FH} = p^*_{HH} + \frac{\tau}{2} \tag{13}$$

Consequently, *the equilibrium prices under monopolistic competition depend on the demand and firm distributions between regions.* In particular, both the prices charged by local and foreign firms fall when the mass of local firms increases (because price competition is fiercer) but the impact is weaker when τ is smaller. In the limit, when τ is negligible, the relocation of firms in H, say, has almost no impact on market prices. In this case, prices are 'independent' of the way firms are distributed between the two regions.

Equilibrium prices also rise when the relative desirability of the differentiated product with respect to the numéraire, evaluated by a, gets larger or when the degree of product differentiation, inversely measured by c, increases provided that trade occurs (see (14) below). All these results are in accordance with what is known in industrial organization and spatial pricing theory.

Furthermore, *there is freight absorption since only a fraction of the trade cost is passed on to the consumers.* Indeed we have:

$$p^*_{HF} - p^*_{HH} = \tau \frac{b + c\lambda N}{2b + cN}, \text{ ...which is equal to...} \frac{\tau}{2} \text{ when } \lambda = 1/2$$

$$p^*_{FH} - p^*_{FF} = \tau \frac{b + c(1-\lambda)N}{2b + cN}, \text{...which is equal to...} \frac{\tau}{2} \text{ when } \lambda = 1/2$$

It is well known that a monopolist facing a linear demand absorbs exactly one-half of the trade cost. By contrast, we see that monopolistic competition leads to more (less) freight absorption than monopoly when the foreign market is the small (large) one: in their attempt to penetrate the distant market, competition leads firms to a price gap that varies with the relative size of the home and foreign markets.

By inspection, it is readily verified that p^*_{HH} (p^*_{FF}) is increasing in τ because the local firms in H (F) are more protected against foreign competition while $p^*_{HF} - \tau$ ($p^*_{FH} - \tau$) is decreasing because it is now more difficult for these firms to sell on the foreign market. Finally, our demand side happens to be consistent with identical demand functions at different locations but different price levels, as in standard spatial pricing theory.

Deducting the unit trade cost τ from the prices set on the distant markets, i.e. (12) and (13), we see that firms' prices net of trade costs are positive regardless of the workers' distribution if and only if

$$\tau < \tau_{trade} \equiv \frac{2a\phi}{2b\phi + cL} \tag{14}$$

The same condition must hold for consumers in F (H) to buy from firms in H (F), i.e. for the demand (8) evaluated at the prices (10) and (11) to be positive for all λ. From now on, condition (14) is assumed to hold. Consequently, there is intra-industry trade and reciprocal dumping, as in Anderson *et al.* (1995). However, *there must be increasing returns for trade to occur.* Indeed, when $\phi = 0$ all potential varieties are produced in each region that becomes autarchic.

It is easy to check that the equilibrium gross profits earned by a firm established in H on each separated market are as follows:

$$\Pi^*_{HH} = (b + cN)(p^*_{HH})^2(A/2 + \lambda L) \qquad (15)$$

where Π^*_{HH} denotes the profits earned in H while the profits made from selling in F are

$$\Pi^*_{HF} = (b + cN)(p^*_{HF} - \tau)^2[A/2 + (1 - \lambda)L] \qquad (16)$$

Increasing λ has two opposite effects on Π^*_{HH}. First, due to tougher competition, the equilibrium price (10) falls as well as the quantity of each variety bought by each consumer living in region H. At the same time, the total population of consumers residing in this region increases so that the profits made by a firm located in H on local sales might rise. What is at work here is *an aggregate local demand effect due to the increase in the local population that may compensate firms for the adverse price effect as well as for the individual demand effect generated by a wider array of competing varieties.*

The individual consumer surplus S_H in region H associated with the equilibrium prices (10)-(13) is then as follows (a symmetric expression holds in region F):

$$\begin{aligned} S_H(\lambda) &= -\frac{aL}{\phi}\left[\lambda p^*_{HH} + (1-\lambda)p^*_{FH}\right] + \frac{(b\phi + cL)L}{2\phi^2}\left[\lambda(p^*_{HH})^2 + (1-\lambda)(p^*_{FH})^2\right] \\ &\quad -\frac{cL^2}{2\phi^2}\left[\lambda p^*_{HH} + (1-\lambda)p^*_{FH}\right]^2 \end{aligned}$$

which can be shown to be concave in λ. Furthermore, (14) implies that $S_H(\lambda)$ is always increasing in λ over the interval $[0, 1]$.

The equilibrium wage prevailing in region H may be obtained by evaluating $w^*_H(\lambda) = \Pi^*_H/\phi$, thus yielding the following expression:

$$w^*_H(\lambda) = \frac{b\phi + cL}{4(2b\phi + cL)^2\phi^2}\left\{[2a\phi + \tau cL(1-\lambda)]^2\left(\frac{A}{2} + \lambda L\right)\right.$$

$$+ [2a\phi - 2\tau b\phi - \tau cL(1-\lambda)]^2 \left[\frac{A}{2} + (1-\lambda)L\right]\Big\}$$

which, after simplifying, turns out to be quadratic in λ. Standard, but cumbersome, investigations reveal that $w_H^*(\lambda)$ is concave and increasing (convex and decreasing) in λ when ϕ is large (small) as well as when τ, c, A, and L are small (large). This implies that both $S_H(\lambda)$ and $w_H^*(\lambda)$ increase with λ when τ is small, while they go in opposite directions when τ is large.

4 When do we observe agglomeration?

The distribution $\lambda \in [0,1]$ is a *spatial equilibrium* when no worker may get a higher utility level by changing location. Given that the indirect utility in region H is as follows:

$$V_H(\lambda) = S_H(\lambda) + w_H + \overline{q}_0$$

a spatial equilibrium arises at $\lambda \in (0,1)$ when

$$\Delta V(\lambda) \equiv V_H(\lambda) - V_F(\lambda) = 0$$

or at $\lambda = 0$ when $\Delta V(0) \leq 0$, or at $\lambda = 1$ when $\Delta V(1) \geq 0$.

In order to study the stability of a spatial equilibrium, we assume that local labor markets adjust instantaneously when some workers move from one region to the other. More precisely, the number of firms in each region must be such that the labor market clearing conditions (5) and (6) remain valid for the new distribution of workers. Wages are then adjusted in each region for each firm to earn zero profits everywhere. For now, we assume a myopic adjustment process, that is, the driving force in the migration process is workers' current utility differential between H and F:

$$\dot{\lambda} \equiv d\lambda/dt = \lambda(1-\lambda)\Delta V(\lambda) \quad (17)$$

when t is time. Clearly, a spatial equilibrium implies $\dot{\lambda} = 0$. If $\Delta V(\lambda)$ is positive, some workers will move from F to H; if it is negative, some will go in opposite direction.

A spatial equilibrium is *stable* for (17) if, for any marginal deviation from the equilibrium, this equation of motion brings the distribution of workers back to the original one. Therefore, the agglomerated configuration is always stable when it is an equilibrium while the dispersed configuration is stable if and only if the slope of $\Delta V(\lambda)$ is nonpositive in a neighborhood of this point.

The forces at work are similar to those found in the core-periphery model. First, the immobility of the farmers is a centrifugal force, at least as long as there is trade between the two regions. The centripetal force finds its origin in a demand effect generated by the preference for variety. If a larger number of firms are located in region H, there are two effects at work. First, less varieties are imported. Second, (10) and (13) imply that the equilibrium prices of all varieties sold in H are lower. Observe that the latter effect does not appear in Krugman's model. This, in turn, induces some consumers to migrate toward this region. The resulting increase in the number of consumers creates a larger demand for the industrial good in the corresponding region, which therefore increases operating profits (hence, wages) and leads to more firms to move there. In other words, both backward and forward linkages are present in our model.

It is readily verified that the indirect utility differential can be written as follows:

$$\begin{aligned}\Delta V(\lambda) &\equiv V_H(\lambda) - V_F(\lambda) \equiv S_H(\lambda) - S_F(\lambda) + w_H(\lambda) - w_F(\lambda) \\ &= C\tau(\tau^* - \tau) \cdot (\lambda - 1/2) \qquad (18)\end{aligned}$$

where

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

and

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

It follows immediately from (18) that $\lambda = 1/2$ is always an equilibrium. Since $C > 0$, for $\lambda \neq 1/2$ the indirect utility differential has always the same sign as $\lambda - 1/2$ if and only if $\tau < \tau^*$, otherwise it has the opposite sign. In particular, *when there are no increasing returns in the manufacturing sector* ($\phi = 0$), the coefficient of $(\lambda - 1/2)$ is always negative since $\tau^* = 0$ so that *dispersion is the only (stable) equilibrium.* This shows once more the importance of increasing returns for the possible emergence of an agglomeration.

It remains to determine when τ^* is lower than τ_{trade}. This is so if and only if

$$A/L > \frac{6b^2\phi^2 + 8bc\phi L + 3c^2L^2}{cL(2b\phi + cL)} > 3 \qquad (19)$$

where the second inequality holds because $b/c = \beta/\gamma - 1 \in (0, +\infty)$. This inequality means that the population of farmers is large relative to the population of workers. When (19) does not hold, the coefficient of $(\lambda - 1/2)$ in (18) is always positive for all $\tau < \tau_{trade}$.

When $\tau < \tau^*$, the symmetric equilibrium is unstable and workers agglomerate in region H (F) provided that the initial fraction of workers residing in this region exceeds $1/2$. In other words, *agglomeration arises when the trade cost is low enough*, as in Krugman (1991b) and for similar reasons. In contrast, for large trade costs, that is, when $\tau > \tau^*$, it is straightforward to see that the symmetric configuration is the only stable equilibrium. Hence, the threshold τ^* corresponds to both the critical value of τ at which symmetry ceases to be stable (the 'break point') and the value below which agglomeration is stable (the 'sustain point'); this follows from the fact that (18) is linear in λ.

Proposition 1 *Assume that $\tau < \tau_{trade}$. Two cases may arise:*

(i) When (19) holds, we have the following: if $\tau > \tau^*$, then the symmetric configuration is the only stable spatial equilibrium with trade; if $\tau < \tau^*$ there are two stable spatial equilibria corresponding to the agglomerated configurations with trade; if $\tau = \tau^*$ then any configuration is a spatial equilibrium.

(ii) When (19) does not hold, any stable spatial equilibrium involves agglomeration.

The reverse of (19) plays a role similar to the 'black hole' condition in Krugman and Venables (1995) and Fujita *et al.* (1999). As in their case, more product differentiation (lower c) and stronger increasing returns (higher ϕ) make the black hole condition more likely. Although the size of the industrial sector is captured here through the relative population size A/L and not through its share in consumption, the intuition is similar: the ratio A/L must be sufficiently large for the economy to display different types of equilibria according to the value of τ. Our result does not depend on the expenditure share on the manufacturing sector because of the absence of general equilibrium income effects: small or large sectors in terms of expenditure share may either be agglomerated when τ is small enough. This does not strike us as being implausible.

Furthermore, when $\tau_{trade} > \tau^*$, trade occurs regardless of the type of equilibrium that is stable. However, the nature of trade varies with the type of configuration emerging in equilibrium. In the dispersed configuration, there is only intra-industry trade in the differentiated product; in the agglomerated equilibrium, the region accommodating the manufacturing sector only imports the homogenous good from the other region.

When increasing returns are stronger, as expressed by higher values of ϕ, τ^* rises since $d\tau^*/d\phi > 0$. This means that *the agglomeration of the manufacturing sector is more likely, the stronger are the increasing returns at the firm's level.* In addition, τ^* increases with product differentiation since $d\tau^*/dc < 0$. In words, *more product differentiation fosters agglomeration.* In particular, c very small implies that $\tau_{trade} < \tau^*$ so that agglomeration always arises under trade.

The best way to convey the economic intuition behind Proposition 1 is probably to make use of a graphical analysis.[8] Figure 1 depicts the aggregate inverse demand in region H for a typical local firm after choosing, for simplicity, the units of L so that $b + cN = 1$:

$$p_{HH} = a + cP_H(n_H, \tau) - \frac{Q_{HH}}{A/2 + \phi n_H} \tag{20}$$

where Q_{HH} is the aggregate local demand of a firm located in H and, because $p_{FH} > p_{HH}$ and the total number of firms is fixed by labor market clearing, the price index P_H is a decreasing function of n_H at a rate which increases with τ:

$$\frac{\partial P_H(n_H, \tau)}{\partial n_H} < 0, \quad \left| \frac{\partial^2 P_H(n_H, \tau)}{\partial n_H \partial \tau} \right| > 0 \tag{21}$$

[Insert Figure 1 about here]

The horizontal and vertical intercepts of (20) are respectively $[a+cP_H(n_H, \tau)] \cdot (A/2 + \phi n_H)$ and $[a + cP_H(n_H, \tau)]$. The equilibrium values of Q_{HH} and p_{HH} are shown as Q'_{HH} and p'_{HH}. They are found by setting marginal revenue equal to marginal cost. The operating profits are shown by the shaded rectangle and accrue to the workers while, as usual, the above triangle represents the consumer surplus enjoyed by both types of workers.

[8]For illustrative purposes, we neglect the impact of relocation on the firm's profit in F since this one is typically smaller than the impact on its profit in H.

Although its depicts a partial equilibrium argument, Figure 1 is a powerful learning device to understand the forces at work in our model. To see why, start from an initial situation where regions are identical ($n_H = n_F$). Suppose that some firms move from the foreign to the home region so that n_H rises and n_F falls. For these firms to want to stay in the home region, operating profits, thus wages, have to increase. Indeed, were this not the case, the firms would rather go back to the foreign region.

As revealed by Figure 1, an increase in n_H has two opposite effects on operating profits, hence on wages. First, as new firms enter the home region, the price index $P_H(n_H, \tau)$ decreases. Ceteris paribus, this would shift the inverse demand (20) toward the origin of the axes and operating profits, hence wages, would shrink. This effect is due to increased competition in the home market and stems from the fact that fewer firms now face trade costs when supplying the home market. But this negative competition effect is not the only effect. For some firms to move to the home region, some workers have to follow since $n_H = \lambda L/\phi$. This means that, as n_H increases, λ also goes up so that the market of the home region expands. Ceteris paribus, the horizontal intercept of the inverse demand would move away from the origin and operating profits, thus wages, would expand. This is a positive demand effect which is induced by the linkage between the locations of firms and workers' expenditures.

Since the two effects oppose each other, the net result is a priori ambiguous. But we can say more than that. In particular, we can assess which effect prevails depending on parameter values. Start with the competition effect that goes through $[a + cP_H(n_H, \tau)]$. This effect is strong if c is large, i.e. if varieties are good substitutes. It is also strong if $|\partial P_H(n_H, \tau)/\partial n_H|$ is large. As shown in (21), this happens if τ is large, because, when obstacles to trade are high, competition from the other region is weak and home firms care a lot about their competitors being close rather than distant. As to the demand effect, it will be strong if ϕ is large because each new firm brings along many workers, and if A is small because immigrants have a large impact on the local market size.

We can therefore conclude that the demand effect dominates the competition effect, when goods are bad substitutes (c small), increasing returns are intense (ϕ large), the farmers are unimportant (A small) and trade costs are low (τ small). Under such circumstances, the entry of new firms in one region would raise the operating profits of all firms, hence wages. *Higher operating profits and wages would attract more firms and workers, thus gen-*

erating circular causation among locational decisions. Agglomeration would then be sustainable as a spatial equilibrium.

This argument establishes a sufficient condition for agglomeration. Since the impact of firms' relocation on consumer surplus is always positive, agglomeration could still arise even when operating profits, hence wages, decrease with the size of the local market because the demand effect is dominated by the competition effect. Furthermore, it is likely to hold for most downward sloping demand functions.

5 Optimality vs. equilibrium

We now wish to determine whether or not such an agglomeration is socially optimal. To this end, we assume that the planner is able (i) to assign any number of workers (or, equivalently, of firms) to a specific region and (ii) to use lump sum transfers from all workers to pay for the loss firms may incur while pricing at marginal cost. Observe that no distortion arises in the total number of varieties since N is determined by the factor endowment (L) and technology (ϕ) in the manufacturing sector and is, therefore, the same at both the equilibrium and optimum outcomes. Because our setting assumes transferable utility, the planner chooses λ in order to maximize the sum of individual indirect utilities:

$$W(\lambda) \equiv \frac{A}{2}S_H(\lambda)+\lambda L[S_H(\lambda)+w_H(\lambda)]+\frac{A}{2}S_F(\lambda)+(1-\lambda)L[S_F(\lambda)+w_F(\lambda)]+A \tag{22}$$

in which all prices have been set equal to marginal cost:

$$p^o_{HH} = p^o_{FF} = 0 \qquad \text{and} \qquad p^o_{HF} = p^o_{FH} = \tau$$

thus implying by (15) and (16) that operating profits are zero, and hence $w^o_H(\lambda) = w^o_F(\lambda) = 0$ for every λ so that firms do not incur any loss. Hence (22) becomes:

$$W(\lambda) = C^o\tau(\tau^o - \tau)\lambda(\lambda - 1) + \text{constant} \tag{23}$$

where

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

and

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

The function (23) is strictly concave in λ if $\tau > \tau^o$ and strictly convex if $\tau < \tau^o$. Furthermore, since the coefficients of λ^2 and of λ are the same (up to their sign), this expression has always an interior extremum at $\lambda = 1/2$. As a result, the optimal choice of the planner is determined by the sign of the coefficient of λ^2, that is, by the value of τ with respect to of τ^o.

Hence we have:

Proposition 2 *If $\tau > \tau^o$, then the symmetric configuration is the optimum; if $\tau < \tau^o$ any agglomerated configuration is the optimum; if $\tau = \tau^o$ any configuration is an optimum.*

In accordance with intuition, it is socially desirable to agglomerate the manufacturing sector into a single region once trade costs are low, increasing returns in the manufacturing sector are strong enough and/or the output of this sector is sufficiently differentiated. Note also that the optimum is always dispersed when increasing returns vanish ($\phi = 0$).

A simple calculation shows that $\tau^o < \tau^*$. This means that *the market yields an agglomerated configuration for a whole range ($\tau^o < \tau < \tau^*$) of trade cost values for which it is socially desirable to have a dispersed pattern of activities.* Accordingly, when trade costs are low ($\tau < \tau^o$) or high ($\tau > \tau^*$) no regional policy is required from the efficiency point of view, although equity considerations might justify such a policy when agglomeration arises. On the contrary, for intermediate values of trade costs ($\tau^o < \tau < \tau^*$), the market provides excessive agglomeration, thus justifying the need for an active regional policy in order to foster the dispersion of the modern sector on both the efficiency and equity grounds.

This discrepancy may be explained as follows. First, the individual demand elasticity is much lower at the optimum (marginal cost pricing) than at the equilibrium (Nash equilibrium pricing), so that regional price indices are less sensitive to a decrease in τ. The fall in trade costs must therefore be sufficiently large to make the agglomeration of workers socially desirable. Second, workers do not internalize the negative external effects they impose on the farmers who stay put. Hence, even though the workers have individual incentives to move, these incentives do not reflect the social value of their move.

6 The impact of workers' expectations on the agglomeration process

The adjustment process (17) is often used in new economic geography. Yet, the underlying dynamics is myopic because workers care only about their current utility level, thus implying that only history matters. This is a fairly restrictive assumption to the extent that *migration decisions are typically made on the grounds of current and future utility flows.* In addition, this approach has been criticized because it is not consistent with fully rational forward-looking behavior (Matsuyama, 1991). In this section, we want to see how the model presented above can be used to shed more light on the interplay between history and expectations in the formation of the economic space when migrants maximize the intertemporal value of their utility flows.

Since workers have perfect foresight, the easiest way to generate a non bang-bang migration behavior is to assume that, when moving from one region to the other, workers incur a utility loss that depends on the rate of migration $\dot{\lambda}$ (Mussa, 1978). In other words, a migrant imposes a negative externality on the others by congesting the migration process. Specifically, we follow Krugman (1991a) and assume that the utility loss for a migrant is equal to $\left|\dot{\lambda}\right|/\delta$, where $\delta \in (0, +\infty)$ is the speed of adjustment.

Following Fukao and Bénabou (1993) as well as Ottaviano (1999), we now define the utility of a worker residing in region H as

$$v_H(t) = \int_t^T e^{-\rho(s-t)} V_H(s) ds + e^{-\rho(T-t)} v_H(T) \qquad (24)$$

where T is the first time when all workers are established into a single region and ρ the discount rate, while a similar expression holds for $v_F(t)$.

Since workers are free to choose where to reside and since the individual migration cost at time t is equal to $\left|\dot{\lambda}\right|/\delta$, for a worker in region H to stay put it must be that the utility in this region (weakly) exceeds the utility in region F minus the migration cost:

$$v_H(t) \geq v_F(t) - \left|\dot{\lambda}(t)\right|/\delta \quad \text{where the equality holds when } \dot{\lambda}(t) < 0 \qquad (25)$$

whereas a similar expression holds for someone living in region F. This means that $v_H(t) - v_F(t)$ stands for the private value for a worker to be in H instead of F.

Assuming an interior solution for (25), we easily get

$$\dot{\lambda} = \delta \Delta v \equiv \delta(v_H - v_F) \tag{26}$$

while differentiating $v_H(t) - v_F(t)$ yields

$$\dot{\Delta v} = \rho \Delta v - \Delta V \tag{27}$$

where $\Delta V \equiv V_H - V_F$ stands for the instantaneous indirect utility differential flow given by (18). Hence we obtain a system of two linear differential equations instead of the first order differential equation (17).

Since $\lambda = 1/2$ implies $\Delta V = 0$, the system has always an interior steady state at $(\lambda, \Delta v) = (1/2, 0)$ which corresponds to the dispersed configuration. Consider now its stability. The eigenvalues of the Jacobian matrix of the system evaluated at $(1/2, 0)$ are given by

$$\frac{\rho \pm \sqrt{\rho^2 - 4\delta C\tau(\tau^* - \tau)}}{2} \tag{28}$$

When $\tau > \tau^*$, the two eigenvalues are real and have opposite signs. Then, the steady state is a saddle point so that the system always converges towards the dispersed configuration, thus implying that neither history nor expectations matter for the final outcome.

Consider now the case in which $\tau < \tau^*$. Two scenarios may arise. In the first one, $\rho > \sqrt{C\delta}\tau^*$ so that the two eigenvalues are still real but both positive. The steady state $(1/2, 0)$ is an unstable node and there are two trajectories that steadily go to the endpoints, $(0, 0)$ or $(1, 0)$, depending on the initial spatial distribution of workers, say λ_0. In this case, *only history matters*: from any initial $\lambda_0 \neq 1/2$, there is a single trajectory that goes towards the closer endpoint as in the case where the dynamics is given by (17). In this case, the myopic adjustment process studied in the previous section provides a good approximation of the qualitative evolution of the economy under forward looking behavior.

Things turn out to be quite different in the second scenario in which $\rho < \sqrt{C\delta}\tau^*$. Since $C\tau(\tau^* - \tau) = 0$ at both $\tau = 0$ and $\tau = \tau^*$, the equation $C\tau(\tau^* - \tau) - \rho^2/4\delta = 0$ has two positive real roots in τ, denoted τ_1^e and τ_2^e, that are smaller than τ^*:

$$\tau_1^e \equiv \frac{\tau^* - D}{2} \qquad \text{and} \qquad \tau_2^e \equiv \frac{\tau^* + D}{2}$$

where

$$D \equiv \sqrt{(\tau^*)^2 - \rho^2/C\delta}$$

stands for the size of the domain of values of τ for which expectations matter; it shrinks as the discount rate ρ increases or as the speed of adjustment δ decreases. Indeed, for $\tau \in (0, \tau_1^e)$ as well as for $\tau \in (\tau_2^e, \tau^*)$, both eigenvalues are real positive numbers and the steady state $(1/2, 0)$ is an unstable node as before. However, for $\tau \in (\tau_1^e, \tau_2^e)$, they become complex numbers with a positive real part so that the steady state is an unstable focus. The two trajectories spiral out from $(1/2, 0)$. Therefore, for any λ_0 close enough to, but different from, $1/2$, there are two alternative trajectories going in opposite directions. It is in such a case that expectations decide along which trajectory the system moves. In other words, *expectations matter for λ close enough to $1/2$, while history matters otherwise.* The corresponding domains are now described.

The range of values for which expectations matter, called the *overlap* by Krugman (1991a), can be obtained as follows. As observed by Fukao and Bénabou (1993), the system must be solved backwards in time starting from the terminal points $(0, 0)$ and $(1, 0)$. The first time the backward trajectories intersect the locus $\Delta v = 0$ allows for the identifications of the endpoints of the overlap:

$$\lambda^L \equiv \frac{1}{2}(1 - \Lambda) \qquad\qquad \lambda^H \equiv \frac{1}{2}(1 + \Lambda)$$

where

$$\Lambda \equiv \exp\left(-\frac{\rho\pi}{\sqrt{4\delta C\tau(\tau^* - \tau) - \rho^2}}\right)$$

is the width of the overlap, which is an interval centered around $\lambda = 1/2$.

The overlap is nonempty as long as $\tau \in (\tau_1^e, \tau_2^e)$. Thus, the width of the overlap is increasing in δ, C, and τ^*, while it decreases with ρ. Moreover, it is $\cap$-shaped with respect to τ, reaching a maximum at $\tau = \tau^*/2$. Since $C\tau(\tau^* - \tau) > 0$ is the slope of ΔV and since this one measures the strength of the forward and backward linkages pushing towards agglomeration, we see that expectations matter more when such linkages are stronger. Consequently, we have shown the following result:

Proposition 3 *Let λ_0 be the initial spatial distribution of workers. Given $\rho < \sqrt{C\delta}\tau^*$, there exist $\tau_1^e \in (0, \tau^*/2)$, $\tau_2^e \in (\tau^*/2, \tau^*)$, $\lambda^L \in (0, 1/2)$, and*

$\lambda^H \in (1/2, 1)$ *such that workers' beliefs about their future earnings influence the process of agglomeration if and only if* $\tau \in (\tau_1^e, \tau_2^e)$ *and* $\lambda_0 \in [\lambda^L, \lambda^H]$.

Hence, history alone matters when τ and λ are large enough or small enough. In other words, the agglomeration process evolves as if workers were short-sighted when obstacles to trade are high or low and when regions are initially quite different. Instead, as long as obstacles to trade take intermediate values and regions are not initially too different, *the equilibrium is determined by workers' expectations and not by history.*

The existence of the range (τ_1^e, τ_2^e) may be explained as follows. Suppose, indeed, that the economy is such that $\lambda_0 > 1/2$ and ask what is needed to reverse an ongoing agglomeration process leading towards $\lambda = 1$. If the evolution of the economy were to change direction, workers would experience falling instantaneous indirect utility flows for some time period as long as $\lambda > 1/2$. The instantaneous indirect utility flows would start growing only after λ becomes smaller than $1/2$. Accordingly, workers would first experience utility losses followed by utility gains. Since the losses would come before the gains, they would be less discounted. This provides the root for the intuition behind Proposition 3. When the forward and backward linkages lead to substantial wage rises (that is, for intermediate values of τ), the benefits of agglomerating at $\lambda = 0$ can compensate workers for the losses they incur during the transition phase, thus making the reversal of migration possible. On the contrary, when these linkages get weaker (that is, for low or high values of τ), the benefits of agglomerating at $\lambda = 0$ do not compensate workers for the losses. As a consequence, *the reversal in the migration process may occur only for intermediate values of* τ.

As to the remaining comparative static properties of the overlap, they are explained by fact that proximity to the endpoint increases the time period over which workers bear losses, large rate of time preference gives more weight to them, and a slow speed of adjustment extends the time period over which workers' well being is reduced.

7 The impact of urban costs

So far, we have assumed that the agglomeration of workers into a single region does not involve any agglomeration costs. Yet, it is reasonable to believe that a growing settlement in a given region will often take the form of an urban area, typically a city. In order to deal with such an aspect of the process of agglomeration, we extend the core-periphery model by adding the central variables suggested by urban economics (Fujita, 1989). In order to keep the analysis short, we go back to the myopic adjustment process of Section 4.

Space is now continuous and one-dimensional. Each region has a spatial extension and involves a linear city whose center is given but with a variable size. The city center stands for a central business district (CBD) in which all firms locate once they have chosen to set up in the corresponding region (see Fujita and Thisse, 1996, for various arguments explaining why firms want to be agglomerated in a CBD). The two CBDs are two remote points of the location space. Interregional trade flows go from one CBD to the other.

Housing is a new good in our economy and is described by the amount of land used by workers. While firms are assumed not to consume land, workers, when they live in a certain region, are urban residents who consume land and commute to the regional CBD in which manufacturing firms are located. Hence, unlike Krugman (1991b) but like Alonso (1964), Helpman (1998) and Tabuchi (1998), *each agglomeration has a spatial extension that imposes commuting and land costs on the corresponding workers.* For simplicity, workers consume a fixed lot size normalized to unity, while commuting costs are linear in distance, the commuting cost per unit of distance being given by $\theta > 0$ units of the numéraire. Without loss of generality, the opportunity cost of land is normalized to zero.

When λL workers live in H, they are equally distributed around the H-CBD. In equilibrium, since all workers residing in region H earn the same wage, they reach the same utility level. Furthermore, since they all consume one unit of land, the equilibrium land rent at distance $x < \lambda L/2$ from the H-CBD is given by

$$R^*(x) = \theta(\lambda L/2 - x)$$

Hence, a worker located at the average distance $\lambda L/4$ from the H-CBD bears a commuting cost equal to $\theta\lambda L/4$ and pays the average land rent $\theta\lambda L/4$ (Fujita, 1989; Papageorgiou and Pines, 1999). When the land rents go to absentee landlords, individual urban costs, defined by commuting cost plus land rent at each residence x, are given by $\theta\lambda L/2$. In order to avoid working with absentee landlords, we assume as in Helpman (1998) that all the land rents are collected and equally redistributed among the H-city workers. Consequently, the individual urban costs after redistribution are equal to

$\theta\lambda L/4$.

Since the urban costs prevailing in regions H and F are not equal, the incentive to move from one region to the other are no longer given by (18). Indeed, we must account for the difference in urban costs between H and F, namely

$$\lambda\theta L/4 - (1-\lambda)\theta L/4 = (\lambda - 1/2)\theta L/2$$

which must be subtracted from (18) to obtain the actual utility differential:

$$\Delta V_u(\lambda) = [C\tau(\tau^* - \tau) - \theta L/2] \cdot (\lambda - 1/2)$$

As in Section 4, agglomeration is a spatial equilibrium when the slope of $\Delta V_u(\lambda)$ is positive. This is the case as long as τ falls within the two values

$$\tau_1^u \equiv \frac{\tau^* - E}{2} \qquad \text{and} \qquad \tau_2^u \equiv \frac{\tau^* + E}{2}$$

where $\tau_1^u, \tau_2^u \in (0, \tau^*)$ and

$$E \equiv \sqrt{(\tau^*)^2 - 2\theta L/C}$$

measures the domain of values of τ for which $\Delta V_u(\lambda) > 0$. Notice that such domain shrinks as the commuting cost per unit of distance θ as well as the total mobile population L increases. Consequently, we have:

Proposition 4 *If $2\theta L < C(\tau^*)^2$, there exist $\tau_1^u \in (0, \tau^*/2)$, $\tau_2^u \in (\tau^*/2, \tau^*)$ such that agglomeration (dispersion) is the only stable equilibrium if and only if $\tau \in (\tau_1^u, \tau_2^u)$ $(\tau \notin (\tau_1^u, \tau_2^u))$. For $\tau = \tau_1^u$ or $\tau = \tau_2^u$, any distribution of workers is a spatial equilibrium. If $2\theta L > C(\tau^*)^2$, dispersion is the only spatial equilibrium.*

Thus, *the existence of positive commuting costs within the regional centers is sufficient to yield dispersion when the trade costs are sufficiently low.* This implies that the relationship between trade costs and the degree of agglomeration is $\cap$-shaped, as argued in the introduction. An increase (decrease) in the commuting costs fosters dispersion (agglomeration) by widening (shrinking) the left range of τ-values for which dispersion is the only spatial equilibrium. Also, sufficiently high commuting costs always yield dispersion.

It is interesting to point out that while dispersion arises for both high and low trade costs, this happens for very different reasons. In the former case, firms are dispersed as a response to the high trade costs they would incur by supplying farmers from a single agglomeration. In the latter, firms are dispersed as a response to the high urban costs workers would bear within a single agglomeration.

Finally, *when the mass of farmers A is equal to zero*, it is readily verified that $\tau_{trade} < \tau^*/2 < \tau_2^u$ so that dispersion does not arise when trade costs are high. *The economy moves from agglomeration to dispersion when trade costs fall*, thus confirming the numerical results obtained by Helpman (1998).

8 Concluding remarks

Recent years have seen the proliferation of applications of the 'Dixit-Stiglitz, iceberg, evolution, and the computer' framework for studying the impact of trade costs on the spatial distribution of economic activities. While these applications have produced valuable insights, they have often been criticized because they rely on a very particular research strategy.

We have proposed *a different framework that is able not only to confirm those insights but also to produce new results that could barely be obtained within the standard one.* Specifically, we have used this framework to deal with the following issues: (i) the welfare properties of the core-periphery model, (ii) the impact of expectations in shaping the economic space, and (iii) the effects of urban costs on the interregional distribution of activities. This suggests that our framework is versatile enough to accommodate other extensions.

So, we have shown that the main results in the literature do not depend on the specific modeling choices made, as often argued by their critics. In particular, the robustness of the results obtained in the core-periphery model against an alternative formulation of preferences and transportation seems to point to the existence of a whole class of models for which similar results would hold. However, we have also shown that those modeling choices can be fruitfully reconsidered once the aim is to shed light on different issues.

The model used in this paper still displays some undesirable features to which it should be remedied in future research. First, there is a fixed mass of firms regardless of the consumer distribution. Furthermore, by ignoring income effects, our setting has a strong partial equilibrium flavor.

References

[1] Alonso, W. *Location and Land Use* (Cambridge (Mass.): Harvard University Press, 1964).

[2] Anderson, S.P., N. Schmitt and J.-F. Thisse "Who Benefits from Antidumping Legislation ?" *Journal of International Economics* 38 (1995), 321-337.

[3] Dixit, A.K. "A Model of Duopoly Suggesting a Theory of Entry Barriers," *Bell Journal of Economics* 10 (1979), 20-32.

[4] Dixit, A.K. and J.E. Stiglitz "Monopolistic Competition and Optimum Product Diversity," *American Economic Review* 67 (1977), 297-308.

[5] Fujita, M. *Urban Economic Theory. Land Use and City Size* (Cambridge: Cambridge University Press, 1989).

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

[7] Fujita, M. and J.-F. Thisse "Economics of Agglomeration," *Journal of the Japanese and International Economies* 10 (1996), 339-378.

[8] Fukao, K. and R. Bénabou "History versus Expectations: A Comment," *Quarterly Journal of Economics* 108 (1993), 535-542.

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

[10] Head, K. and T. Mayer "Non-Europe. The Magnitude and Causes of Market Fragmentation in the EU," *Weltwirtschaftliches Archiv*, forthcoming, 2000.

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

[12] Horn, H. and O. Shy "Bundling and International Market Segmentation," *International Economic Review* 737 (1996), 51-69.

[13] Hotelling, H. "Stability in Competition," *Economic Journal* 39 (1929), 41-57.

[14] Koopmans, T.C. *Three Essays on the State of Economic Science* (New York: McGraw-Hill, 1957).

[15] Krugman, P. "History versus Expectations," *Quarterly Journal of Economics* 106 (1991a), 651-667.

[16] Krugman, P. "Increasing Returns and Economic Geography," *Journal of Political Economy* 99 (1991b), 483-499.

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

[18] Krugman, P. "Urban Concentration: The Role of Increasing Returns and Transport Costs," *International Regional Science Review* 19 (1996), 5-30.

[19] Krugman, P. "Space: The Final Frontier," *Journal of Economic Perspectives* 12 (1998), 161-174.

[20] Krugman, P. and A.J. Venables "Integration and the Competitiveness of Peripheral Industry," in C. Bliss and J. Braga de Macedo, eds., *Unity with Diversity in the European Community* (Cambridge: Cambridge University Press, 1990), 56-75.

[21] Krugman, P. and A.J. Venables "Globalization and the Inequality of Nations", *Quarterly Journal of Economics* 110 (1995), 857-880.

[22] Lösch, A. *Die Räumliche Ordnung der Wirtschaft* (Jena: Gustav Fischer Verlag, 1940). English translation: *The Economics of Location* (New Haven (Conn.): Yale University Press, 1954).

[23] Matsuyama, K. "Increasing Returns, Industrialization, and Indeterminacy of Equilibrium," *Quarterly Journal of Economics* 106 (1991), 617-650.

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

[25] Mussa, M. "Dynamic Adjustment in the Heckscher-Ohlin-Samuelson Model," *Journal of Political Economy* 86 (1978), 775-791.

[26] Ottaviano, G.I.P. "Integration, Geography and the Burden of History," *Regional Science and Urban Economics* 29 (1999), 245-256.

[27] Ottaviano, G.I.P. and D. Puga "Agglomeration in the Global Economy: A Survey of the 'New Economic Geography'," *World Economy* 21 (1998), 707-731.

[28] Papageorgiou, Y.Y. and D. Pines *An Essay in Urban Economic Theory* (Dordrecht: Kluwer Academic Publishers, 1999).

[29] Phlips, L. *Applied Consumption Analysis* (Amsterdam: North-Holland, 1983).

[30] Puga, D. "The Rise and Fall of Regional Inequalities," *European Economic Review* 43 (1999), 303-334.

[31] Tabuchi, T. "Agglomeration and Dispersion: A Synthesis of Alonso and Krugman," *Journal of Urban Economics* 44 (1998), 333-351.

[32] Thisse, J.-F. and X. Vives "On the Strategic Choice of a Spatial Price Policy", *American Economic Review* 78 (1988), 122-137

[33] Vives, X. "Trade Association Disclosure Rules, Incentives to Share Information, and Welfare," *Rand Journal of Economics* 21 (1990), 409-430.

[34] Wei, S.-J. "Intra-national versus International Trade: How Stubborn are Nations in Global Integration," National Bureau of Economic Research Working paper n° 5531, 1996.

APPENDIX

In the case of two varieties, the symmetric quadratic utility is given by:

$$U(q_1, q_2) = \alpha(q_1 + q_2) - (\beta/2)(q_1^2 + q_2^2) - \gamma q_1 q_2 + q_0$$

In the case of $n > 2$ varieties, this expression is extended as follows:

$$\begin{aligned} U(q) &= \alpha \sum_{i=1}^{n} q_i - (\beta/2) \sum_{i=1}^{n} q_i^2 - (\gamma/2) \sum_{i=1}^{n} \sum_{j \neq i}^{n} q_i q_j + q_0 \\ &= \alpha \sum_{i=1}^{n} q_i - [(\beta - \gamma)/2] \sum_{i=1}^{n} q_i^2 - (\gamma/2) \sum_{i=1}^{n} \left(q_i \sum_{j=1}^{n} q_j \right) + q_0 \\ &= \alpha \sum_{i=1}^{n} q_i - [(\beta - \gamma)/2] \sum_{i=1}^{n} q_i^2 - (\gamma/2) \left(\sum_{i=1}^{n} q_i \right)^2 + q_0 \end{aligned}$$

When $\gamma \to \beta$, this utility boils to a standard quadratic utility for a homogenous good. Letting $n \to \infty$ and $q_i \to 0$, we obtain (1) in which N stands for the mass of varieties.

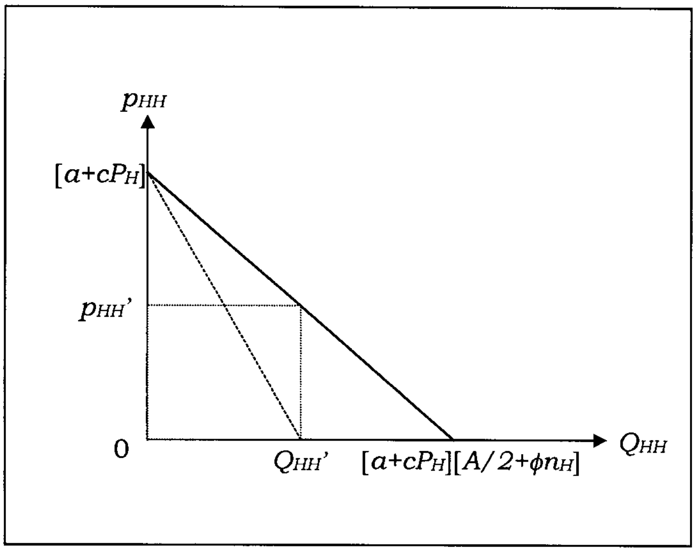

Figure 1 – Inverse demand