Entrepreneurs’ Location Choice and Public Policies: a Survey of the New Economic Geography

Fabien Candau

To cite this version:


HAL Id: hal-02514085
https://hal-univ-pau.archives-ouvertes.fr/hal-02514085

Submitted on 21 Mar 2020

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L’archive ouverte pluridisciplinaire HAL, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d’enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.
Abstract. The aim of this paper is to survey what has been done by the New Economic Geography (NEG) on a regional scale in order to answer the three following questions: what are the predictions of the NEG concerning the future of regions in the triad? Are these predictions robust? What can be the optimal public policy on a regional and national scale in a world characterized by agglomeration, trade liberalization and mobility of entrepreneurs? In surveying the most recent contributions in this area, the paper sheds light on several shortcomings of the NEG literature in order to point out new directions for further research, with particular reference to studies concerning welfare and tax competition.

Keywords. Economic geography; Taste heterogeneity; Welfare; F12, R13

1 Introduction

"Consider a general equilibrium model in which an arbitrary number of goods is produced either as inputs or for final consumption. The only nonproduced goods are land and labor, each of which is assumed to be homogeneous. Assume that each production function has constant returns to scale and that all input and output markets are competitive [...] Under these circumstances, consumers would spread themselves over the land at a uniform density to avoid bidding up the price of land above that of land available elsewhere. Adjacent to each consumer would be all industries necessary directly or indirectly to satisfy the demands of that customer. Constant returns assures us that production could take place at an arbitrary small scale without loss of efficiency. In this way, all transportation costs could be avoided without any need to agglomerate economic activity » Edwin S. Mills (1967, p198)

In this quote extracted from a section entitled "A World without Cities", Edwin Mills intuitively describes what is going to be called some years later the Spatial Impossibility Theorem. This theorem, which we owe to Starrett (1978), simply demonstrates that agglomeration of activities cannot arise in a world of pure and perfect competition where space is homogeneous. Thus, in order to understand why
activities are agglomerated in some places authors have focused on localized externalities (Marshall (1890)), on heterogeneity of space (von Thünen (1826), Ricardo (1917), Heckscher (1918) and Ohlin (1933)) and on imperfect competition (Hotelling (1929)). More recently, a new approach born with Fujita (1988), Krugman (1991a) and Venables (1996), has revisited this problematics of "Who does What, Where and Why?". That literature which makes the heterogeneity of space endogenous by integrating pecuniary externalities and imperfect competition starts from the same observation that activities are unevenly spread in space. As a very crude approximation, Figure 1 illustrates how obvious that statement is, an increasingly significant part of the world being agglomerated in huge cities.

Figure 1 – Urban population and growth rate of urbanization.

Source: Author’s calculation based on the World Urbanization Prospects (Department of Economic and Social Affairs).

The questions raised by these observations are linked to the conditions of agglomeration: 1) what are they?; 2) Can these conditions vanish in the future?; 3) What is the place of governments in a world characterized by agglomeration, trade liberalization and mobility of firms? We are going to answer these questions in four stages, in which we will focus on models that deal exclusively with entrepreneurs’ mobility². The first part presents models where entrepreneurs do not change sectors but move geographically, where they change sectors but not regions and lastly where they change regions and are at the source of endogenous growth. Many of these models are presented in detail since they are a basis for recent literature’s building of public policy. They are also models which have been questioned by empirical works since they all display two particularly strong results, namely that trade liberalization leads to agglomeration and that depending on the trade liberalization
level, two or three different location configurations are stable equilibria. Then, in
the second part we survey empirical articles that test the agglomeration trend and
the existence of multiple equilibria. In this part, we will study the European Union,
the USA and Japan separately in order to understand if the results found in one part
of the triad can be generalized to any other. Furthermore we will also distinguish
between different geographical scale in order to find out whether empirical results
are identical at national, regional and urban levels. These empirical investigations
and criticisms have marked a new starting point for theoretical works most of which
have not been surveyed in detail so far. Then, in the third part we will study how
a few theoretical refinements have questioned the importance of the Home Market
Effect, how the spatial configuration can be affected by firms’ heterogeneity and
taste heterogeneity, and how, at a high globalization level, the presence of land rent
and commuting costs largely accounts for dispersion. Lastly the fourth part will
focus on public policy and more particularly on welfare, tax competition and R&D
investment. In more, we present the current research frontier, but we also pose new
questions and try to cross this frontier by demonstrating new results. In particular
we would like to understand how trade and growth can interact on an urban scale,
whether tax competition and trade liberalization generate a race to the top followed
by a race to the bottom, and if agglomeration is always desirable for entrepreneurs.

2 From dispersion to agglomeration, the initial theoretical considerations

This first part presents models where entrepreneurs are: 1) immobile between sectors
but mobile geographically 2) mobile between sectors but immobile between regions
3) mobile between regions and at the source of endogenous growth.

2.1 Inter-regional mobility

2.1.1 Consumers’ behavior

All individuals in the Krugman (1991a) model share the Dixit and Stiglitz (1977)
constant-elasticity-of-substitution utility function:

$$ U = M^\mu A^{1-\mu} \quad \text{with} \quad M = \left[ \int_0^N \frac{e_i^{\sigma-1}}{c_i^{\sigma}} \, di \right]^{\frac{\sigma}{\sigma-1}} $$

(1)

where $M$ is the consumption of the manufactures aggregate, $A$ that of the agricul-
tural product, $N$ is the large number of potential varieties and where $\sigma > 1$ is
the elasticity of substitution among these varieties. A share $\mu$ of nominal income
(denoted $Y$) is spent on manufactures. The budget constraint is then given by
$PM + p_A A = Y$, where $p_A$ is the price of the agricultural product and $P$ the price
index of industrial varieties:

\[ P = \left[ \int_0^N p_i^{1-\sigma} di \right]^{\frac{1}{1-\sigma}} \]  

(2)

which is a decreasing function of the number of varieties \( N \) (because \( 1 - \sigma < 0 \)). \( p_i \) is the price of a typical variety \( i \). The impact of \( N \) on the price index depends on the elasticity of substitution. The more differentiated the product varieties, the greater the reduction in the price index. The maximization problem yields the following uncompensated demand for agriculture and manufactures:

\[ M = \mu \frac{Y}{P}, \quad A = (1 - \mu) \frac{Y}{p_a} \]  

(3)

\[ c_i = \mu \frac{Y}{P_1^{1-\sigma}} p_i^{-\sigma} \]  

(4)

We can now turn to the firms’ behavior.

2.1.2 Firms’ behavior

Concerning the cost function many types of modelling have been proposed. Firstly in its Core-Periphery model (CP for short), Krugman (1991a) considers that there are specific factors: entrepreneurs are the only input necessary to the production of a typical variety with a technology that involves a fixed amount \( \alpha \) of input and a marginal requirement of \( \beta \). The cost function is thus given by \( TC_{CP} = \alpha w + \beta wq \). Secondly in their Footloose Entrepreneurs model (FE for short), Forslid and Ottaviano (2003) consider that the fixed cost and the marginal cost are associated with different factors: the fixed cost involves \( w \) units of entrepreneurs while the variable cost requires \( a \) units of workers. Thus the total cost of producing \( q \) units of a typical manufactured variety is \( TC_{FE} = \alpha w + \beta w_a q \). To sum up we can write these total costs in the following general form:

\[ TC = \alpha f + \beta mq \]  

(5)

with \( (f, m) = (w, w) \) in the CP, \( (f, m) = (w, w_a) \) in the FE. Because each firm produces a distinct variety, the number of firms is also the number of varieties consumed. Thus each firm is a monopolist in the production of its variety and faces the demand function (4). But a key feature of the Dixit-Stiglitz monopolistic competition is that firms ignore the effects of their action on income \( Y \) and on price index \( P \). Hence the demand curve as perceived by a typical firm is not (4), but rather \( q = bp^{-\sigma} \) where \( b = \mu Y / P_1^{1-\sigma} \) is considered as a constant by each firm. According to this behavior, when maximizing its profit, a typical firm sets the following price:

\[ p = \beta m \sigma / (\sigma - 1) \]  

(6)

Because there is free entry profits are always equal to zero, which, using equations (5) and (6), gives the output level:

\[ q = \alpha f (\sigma - 1) / \beta m \]  

(7)
In equilibrium in the CP model a typical firm employs \( \alpha f + \beta mq \) entrepreneurs, so that the total demand is \( n(\alpha f + \beta mq) \), which according to (7) is equal to \( na\sigma \). As entrepreneurs’ labour supply is given by \( S \), the equalization provides the following number of varieties: \( n_{CP} = \frac{S}{\sigma} \). In the FE a typical firm employs \( \alpha \) units of entrepreneurs, so that the total demand is \( na \). As entrepreneurs’ labour supply is also \( S \), the equalization gives the following number of varieties: \( n_{FE} = \frac{S}{\sigma} \). In short:

\[
n = \frac{S}{\varepsilon}
\]

with \( \varepsilon = \alpha \sigma \) in the CP, \( \varepsilon = \alpha \) in the FE. The number of varieties produced is then proportional to the number of workers.

### 2.1.3 Transaction costs

So far, the model has almost been described as a closed economy. The next step is to relax this assumption. Industrial varieties are exchanged between regions under transaction costs which take the form of iceberg costs: if an industrial variety produced in the Northern market is sold at price \( p \) there, then the delivered price (c.i.f) of that variety in the South is going to be \( \tau p \). The assumption of iceberg costs implies that firms charge the same producer price in both regions. The first-order conditions for a typical firm’s sales on its local market and on its export market are \( p = \beta m\sigma/(\sigma - 1) \) and \( p^* = \tau\beta m\sigma/(\sigma - 1) \), but in all models authors assumed that \( \beta = (\sigma - 1)/\sigma \), which gives \( p = m \) and \( p^* = \tau m \). This normalization and the fact that wages in the agricultural sector are taken as the numeraire and normalized to one simplify prices, which only depend on entrepreneurs’ wages in the CP and which are equal to one in the FE. Furthermore the total number of entrepreneurs is also normalized to one: \( S + S^* = 1 \). Then, iceberg transaction costs imply a modification of the price index. Using the above normalization the authors find:

\[
P^{1-\sigma} = \frac{Sm^{1-\sigma} + \phi S^*m^{1-\sigma}}{\varepsilon}, \quad (P^*)^{1-\sigma} = \frac{Sm^{1-\sigma} + \phi S^*m^{1-\sigma}}{\varepsilon}
\]

where \( \phi \) measures the freeness of trade : \( \phi = (\tau)^{1-\sigma} \). This trade freeness increases from \( \phi = 0 \) with infinite trade costs, to \( \phi = 1 \), with zero trade costs. *Ceteris paribus*, at the symmetric equilibrium \( S = 1/2 \), an increase in \( S \) (and so a decrease in \( S^* \)) causes the price index in the South to increase and that of the North to decrease. We now need to integrate transaction costs into the demand function. By inserting the above prices (6) in the demand function (4) and by considering the total demand as the sum of local demand and export demand, we get:

\[
q = \mu\left(\frac{Y}{P^{1-\sigma}p^{-\sigma}} + \phi \frac{Y^*}{P^{*1-\sigma}p^{-\sigma}}\right)
\]

Considering this expression, we have just seen that an increase in the northern population increases \( P^{1-\sigma} \) and decreases \( P^{*1-\sigma} \), and thus fosters a decrease in the total demand \( q \) in the North (if \( \phi < 1 \)). The effect of a change of income location
can be analyzed through the following equations:

\[ Y = Sw + Lw_a \]  \hspace{1cm} (11)

\[ Y^* = S^*w^* + L^*w_a \]  \hspace{1cm} (12)

An increase in the entrepreneurial force \( S \) in the North, and thus a decrease in the South \( S^* \) raises expenditure in this country and lowers it abroad, which causes, as long as there are impediments to trade \( (\phi < 1) \), an increasing demand \( q \). These equations permit us to present the market clearing in a tidy form by equalizing demand (10) and supply given by equation (7):

\[ fm^{\sigma-1} = b\left(\frac{Y}{p_1^{1-\sigma}} + \phi\frac{Y^*}{p^{1-\sigma}}\right) \]  \hspace{1cm} (13)

with \( b = \frac{\phi}{\alpha} \). The most significant difference between the FE and CP models is that in the CP model, nominal wages cannot be obtained analytically since \( Y, P, f \) and \( m \) depend on \( w \) while in the FE only \( f \) and \( Y \) depend on \( w \) and thus the analytical expression of nominal wages is rather straightforward. The detailed description of the FE is postponed to the section concerning public policy; indeed the tractability of this model is going to be an advantage in this section. For the moment we focus on the CP model which is one of the main tools of section 4. Indeed in this section we are going to consider models which drop agricultural workers, and in such a context the FE total cost function is no longer useful and the interest of the CP model is restored.

2.1.4 Market clearing condition and the long run in the CP

In the CP model, \( f = m = w \), which gives the following market clearing thanks to (13):

\[ w^\sigma = b\left(\frac{Y}{p_1^{1-\sigma}} + \phi\frac{Y^*}{p^{1-\sigma}}\right) \]

As we have just noticed, nominal wages cannot be obtained analytically. However it is possible to investigate the relationships between price indices and wages by linearizing the model around the symmetric equilibrium (when \( S = S^* = 1/2 \)). Indeed, at the symmetric equilibrium, a modification of one variable in one region is associated with an equal modification of the corresponding variable in the other region, but with an opposite sign, so by letting \( \frac{dw}{w} = -\frac{dw^*}{w^*} = \hat{w} \) and so on, we get the following expression by way of a log differentiation of price indices and wages:

\[ \hat{P} = \frac{1}{1-\sigma}Z\phi[\hat{S} + (1-\sigma)\hat{w}] \]  \hspace{1cm} (14)

\[ \sigma\hat{w} = Z\phi(\hat{Y} + (\sigma - 1)\hat{P}) \]  \hspace{1cm} (15)

where \( Z\phi = \frac{1-\phi}{1+\phi} \) can be considered as a reversed measure of trade openness: there is autarky when \( Z\phi = 1 \), and free trade for \( Z\phi = 0 \). Concerning the first expression (14), and since entrepreneurs’ supply of labor is perfectly elastic, \( \hat{w} = 0 \), the more entrepreneurs in the North, \( \hat{S} \), the lower the price index in this country. This effect
is known as the local competition effect (Fujita al. (1999)) or as the market crowding effect (Baldwin et al. (2003)) Regarding the second expression (15), we can eliminate \( b \) by using (14); we then divide the two sides of this equation by \( b \), which gives:

\[
\frac{\tilde{S}}{\tilde{Y}} = \frac{1}{Z_\phi} - \left( \frac{\sigma}{Z_\phi^2} + 1 - \sigma \right) \frac{\tilde{w}}{\tilde{Y}}
\]

(16)

Thus by considering once again that entrepreneurs’ labour supply is perfectly elastic, \( \tilde{w} = 0 \), we get the famous home market effect (Krugman (1980)), also known as the market access effect (Baldwin et al. (2003)): \( \frac{\tilde{S}}{\tilde{Y}} = \frac{1}{Z_\phi} \), then one percent change in the northern demand for manufactures, \( \tilde{Y} \), increases entrepreneurs’ employment, \( \tilde{S} \), by \( \frac{1}{Z_\phi} (> 1) \) percent in the North. This effect highlights that increasing returns activities tend to concentrate their production near the largest market and export to the smallest. When trade is liberalized \( (Z_\phi \text{ tends to zero}) \) this effect increases, which can be explained by the fact that the large market serves as a base for exports, therefore lower export costs render agglomeration more attractive. We are going to discuss the pervasiveness of this HME in the next section. Furthermore we can observe that the higher \( S \) the higher \( Y \), (see (11)). Then the larger manufacturing sector has a larger home market, and since we have just seen that the larger home market has a more than proportionally larger manufacturing sector, these equations provide what Krugman (1991a) called the backward linkage in honor of Hirschman (1958) and was considered as a ‘cryptic moniker’ by Baldwin et al. (2003) who qualify this mechanism as the demand-linked circular causality.

To sum up, two opposite forces drive these relative nominal wages. On the one hand, an increase in the number of entrepreneurs in one region exacerbates local competition among firms. Thus new entry triggers a slump in the price index, and thereby in operating profits, too. So that in order to stay in the market, firms need to remunerate their workers less (local competition effect). But on the other hand, as the income generated by the new entrepreneur is spent locally, sales and operating profits increase and under the ‘zero profit condition’ this implies a higher nominal wage (market access effect). However entrepreneurs do not consider the relative nominal wage when they decide to migrate but the relative real wage. Hence in the long run migration stops when real wages are equalized in case of symmetry \( (S = \frac{1}{2}) \), or when agglomeration in one city generates a higher relative real wage. We denote this relative real wage \( \Omega \) and we define it as:

\[
\Omega = \frac{V}{V^*} = \frac{w}{w^*} \left( \frac{P^*}{P} \right)^{-\mu}
\]

(17)

where \( V \) is the location’s total real income in the North. We will have a stable total agglomeration in the North if \( \Omega > 1 \), and a stable dispersed equilibrium if \( \frac{dH_l}{ds} < 0 \). Let us notice that in the long run (17) one additional force appears: the term \( P^*/P \) which is the cost of living effect, is known as an agglomerative force. Indeed, we already know that goods are cheaper in a central place because imports are lower and thus the burden of transaction costs too. Hence, entrepreneurs’ purchasing power is higher in this location, which attracts other entrepreneurs. This is the cost-linked circular causality also known as forward linkage.
2.1.5 The tomahawk diagram

In order to see how trade liberalization globally affects the size of regions, authors use numerical simulations and obtain Figure 2, called the 'tomahawk diagram'.

In Figure 2, the vertical axis measures the north’s share of the world endowment of entrepreneurs, while the horizontal axis plots the trade liberalization level. Two important results can be drawn:

**Proposition 1** Trade liberalization leads to agglomeration.

Indeed before a critical point of trade liberalization $\phi_s$, the negative effects of high transaction costs are too strong to be overtaken by the positive effects of agglomeration, and thus the two regions remain identical. But after this point, agglomeration can occur catastrophically in one of the two regions.

**Proposition 2** Depending on the trade liberalization level, two or three different location configurations are stable equilibria

Indeed, after a critical point of trade liberalization $\phi_s$, and before the critical point $\phi_b$, dispersion of activities in the two regions, or agglomeration in the South, or agglomeration in the North are all stable equilibria (in Figure 2 stable equilibria are represented by a solid line, as opposed to dashed lines which represent unstable equilibria). After $\phi_b$ agglomeration in the North is equivalent to agglomeration in the South.

These two propositions are strong and deserve to be compared to other models.
2.2 Sectorial factor mobility

If entrepreneurs’ mobility is a central determinant of agglomeration on a regional scale, one may wonder what will happen if this mobility is restricted by immigration laws or by cultural differences. This question was raised by Krugman and Venables (1995) in a model called the Core Periphery Vertical Linkages (CPVL for short). In this model agglomeration or more exactly specialization is driven by the interest of firms that produce and use intermediate goods. Indeed in a big market these firms find bigger outlets (backward linkage), but also intermediate inputs at a lower price (forward linkage). This model was simplified by Ottaviano (2002), actually known as the Footloose Entrepreneurs Vertical Linkages (FEVL for short) and a unified version, which we propose to survey here, was proposed by Ottaviano and Robert-Nicoud (2006). In this last version the authors consider the fixed cost as a Cobb-Douglas function of workers and manufacturing goods which are used respectively in the proportion $1 - \mu$ and $\mu$, while the variable cost is also a Cobb-Douglas function of workers and manufacturing goods but where these two inputs are used in the proportion $1 - \gamma$ and $\gamma$; in such a case the total cost is given by $TC = \alpha w_u^{1-\mu}P^\mu + \beta w_u^{1-\gamma}P^\gamma q$ and as before we may write this total cost in the following general form:

$$TC = \alpha f + \beta mq$$

(18)

with $(f, m) = (w_u^{1-\mu}P^\mu, w_u^{1-\gamma}P^\gamma)$. The CPVL is based on the assumption that the share of workers used as a fixed cost and as a variable cost is the same: $\mu = \gamma$, while in the FEVL, the authors assume that firms only use agricultural workers as a variable cost: $\mu \neq \gamma = 0$. Agricultural wages are taken as the numéraire and normalized to one. Then in the CPVL we get $(f, m) = (P^\gamma, P^\gamma)$, while in the FEVL we have $(f, m) = (P^\mu, 1)$. The pure profit of the firm, $\Pi$, is thus composed of the operating profit minus the fixed cost:

$$\Pi = \pi - \alpha f \quad \text{with} \quad \pi = pq - \beta mq$$

(19)

By maximizing this pure profit a typical firm sets the following price:

$$p = \beta m\sigma/(\sigma - 1)$$

(20)

By assuming that $\beta = \frac{\pi}{\sigma}$ and by integrating this price in operating profit (19) we get:

$$q = \frac{\pi\sigma}{m}$$

(21)

By using (20) and iceberg costs in the price index yields:

$$P_1^{1-\sigma} = nm^{1-\sigma} + \phi n^* m^{1-\sigma}, \quad (P^*)^{1-\sigma} = nm^{1-\sigma} + \phi n^* m^{1-\sigma}$$

This implies that in the CPVL price indices are defined in a recursive way since $m = P^\gamma$, while in the FEVL this is not the case because $m = 1$.

When it comes to sales two kinds of agents need to be taken into consideration: on the one hand workers and entrepreneurs consume manufactured goods as final goods, their demand is denoted $d_1$, and on the other hand, firms use these goods
as intermediate inputs; their input demand is denoted $d_2$. Then the total northern demand is $c = d_1 + nd_2 + \tau(d_1^* + nd_2^*)$ where $d_1$ and $d_2$ are found by applying respectively the Roy identity to the indirect utility, and the Shephard lemma to the total cost function. The market clearing condition is obtained by equalizing (21) to the total demand $c$, which gives:

$$\pi = mb\left(\frac{E}{P^{1-\sigma}} + \phi\frac{E^*}{P^{1-\sigma}}\right)$$

with $b = \frac{\mu}{\sigma}$

with $E$ the total expenditure in the North:

$$E = L + n(1 + (\sigma - 1)\frac{\bar{Y}}{\mu})\pi$$

The expression of this short run equilibrium is very close to that displayed by the classic CP model. Indeed two opposite forces drive this operating profit. On the one hand an increase in the number of firms in the industrial sector exacerbates local competition between them. Thus new entry triggers a slump in the price index, and thereby in operating profits, too. But on the other hand as the revenue generated by the new firm is spent on intermediate inputs, sales and operating profits increase more than proportionally in the presence of iceberg costs (market access effect). However firms do not consider operating profit but pure profit when they decide about entry or exit. Hence in the long run, there is neither entry nor exit when pure profit is equal to zero or by equivalence when $q = \frac{\pi}{\sigma f}$ is equal to the unity. Firms creation or destruction stop in the North and in the South respectively when $\dot{n} = 0$ and $\dot{n}^* = 0$, where $\dot{n}$ and $\dot{n}^*$ are defined as:

$$\dot{n} = n(q - 1), \quad \dot{n}^* = n^*(q^* - 1)$$

with $q = \frac{\pi}{\sigma f}$

In the two models $f$ is a function of the price index. Then a new agglomeration force occurs in the long term. In the CP/FE model this force was called the cost of living effect, here we had better call it the production-cost-linked agglomeration (Baldwin et al. (2003)). Indeed we already know that goods are cheaper in a central place because imports are lower and thus the burden of transaction costs, too. Hence, the firms’ profitability is higher in this location, which attracts further entry. As in the CP/FE model, in these two vertical linkage models, the dispersive force, which is strongest when trade is restricted, erodes faster than agglomeration forces with respect to $\phi$. Thus the dispersive equilibrium where the two countries produce manufacturing goods turns into an agglomerative one where only the North or the South produces these items. But the similarity between the two classes of model does not stop here: indeed, the configurations of equilibrium are exactly the same as those described by the Tomahawk diagram (See Figure 2), even the break point and sustain point are identical. This result is surprising since these models are
quite different. In fact Robert-Nicoud (2005) has demonstrated that these models are all isomorphic: "[these models] can entirely be characterized by the same set of equations in the appropriate state-cum-parameter space". Then Proposition 1 and 2 seem to be strong results, and we may wonder if the introduction of forward-looking behavior and endogenous growth à la Grossman-Helpman-Romer might change these results.

2.3 Regional growth

Relying on capital mobility, Martin and Ottaviano (1999) and Baldwin et al. (2001) introduced endogenous growth in the NEG framework. This introduction has generated a flourishing literature which has been surveyed thoroughly by Baldwin and Martin (2004). But surprisingly models with entrepreneurial mobility take a small place in this literature: only two models, those of Baldwin and Forslid (2000) and Fujita and Thisse (2002), focus on the issue. Those models mainly differ from the FE/CP model owing to the introduction of an R&D sector, where entrepreneurs produce a patent which represents industrial firms’ fixed costs. Then the number of patents produced depends on the capital of past ideas and on the number of entrepreneurs. The presence of this new sector thus generates spillovers for the industrial sector (by increasing its productivity level) that can be local (limited to one region) or global. The first interesting result is that the growth rate follows a U-shape with respect to the number of entrepreneurs. This means that agglomeration leads to a higher growth than dispersion. Moreover whatever entrepreneurs’ spatial distribution, any policies that make local spillovers global raise the growth rate. By considering global and local spillovers in turn, Fujita and Thisse conclude that in the former case, the R&D sector is always agglomerated in one country, for instance in the North, while the industrial sector moves from a partial agglomeration in the North to a total agglomeration in this country when trade is liberalized; dispersion is then never sustainable. On the contrary, with local spillovers, the liberalization process leads both sectors from dispersion to agglomeration. We consider the latter assumption of local spillovers as central and thus decide to study it. The necessary condition to prove that northern agglomeration is a stable steady state growth path (ss-growth path for short) is given when the northern ratio of indirect utility is greater than one, while this ratio has to be equal to one when $S \in (0,1)$. However this condition is not sufficient here: indeed, the North can have a higher nominal wage while the South can have a lower price index (and vice versa) and then entrepreneurs can increase their intertemporal utility by changing their location during their life cycle. As the authors remark, entrepreneurs “convexify their location choices” since their forward looking behavior and their mobility allows them to average their consumption in their lifetime. As a result, the authors show that in the interval $[\phi_s, \phi_b]$ for a given and constant number of entrepreneurs in one region, cross-migration may occur. Moreover authors show that when trade is liberalized unexpectedly, then in case of high trade costs ($\phi < \phi_s$), the dispersive ss-growth path is stable while after $\phi_b$ the agglomerative ss-growth path is the only stable one. Then once again we may point out that proposition 1 and 2 still hold. Indeed trade liberalization leads to agglomeration (Proposition 1) and even if between the
break and sustain point the dispersive equilibrium is not a stable one as in the FE/CP/VL models, two opposite configurations of location (agglomeration in the North or in the South) remain stable (Proposition 2). These results are strong and thus deserve empirical verifications.

3 Empirical observations

3.1 Questioning Proposition 1

3.1.1 The Japanese case

In order to analyze how the Japanese population has changed locations over the past 8,000 years, Davis and Weinstein (2002) calculated different measures of the Variation in Regional Density (VRD for short). Regarding this VRD, three central points are remarkable. At first in the early stages of economic development (from -6000 to 300) the VRD was relatively high, in that case the first nature (climate, endowment of resources) explains the agglomeration of the population; secondly from 1721 to 1872, all indicators decrease and this period of dispersion seems to correspond to a period of closure to trade in Japan. Lastly the shift to free trade and the start of Japan’s industrial revolution around the year 1872 were marked by a rise in the VRD, which revealed an increase in the concentration of activities. These intuitions were confirmed by Fujita et al. (2004) who pointed out that from 1920 to 2000, the share of the peripheral population (non-metropolitan areas) decreased from 36.4% to 17.9% while the share of Tokyo’s population increased from 11.8% to 25.1%. However these observations are insufficient. The close study of the net migration of the three largest metropolitan areas (Tokyo, Osaka, Nagoya) and of Theil’s measure of the interregional per capita real income differential allow Fujita et al. (2004) to reach interesting conclusions. For the authors three urbanization cycles can be defined. In the first one, within the years 1955-1975, they consider that the intuitions of the CP model were perfectly verified between 1955 and 1962. Indeed, this period was characterized by an increasing share of manufacturing industries such as electrical or material industries, which have strong increasing returns and technological linkage, but also by decreasing primary activities and internal transportation costs. However on account of higher wages and land prices in these metropolitan areas, the growth rate of net migration has decreased around the year 1965. Land rent is thus a central variable that is lacking in the CP model and that we are going to consider in the next section. During the second cycle (1975-1985), the globalization of the world economy and the appreciation of the yen increased competition between Japan, the USA and the newly industrialized countries of East Asia. This exacerbation of competition led to a significant structural change in the Japanese economy. Indeed workers in the periphery, specialized in traditional commodities, endured a decrease in their real wages while the only city which enjoyed at the time an international airport and a major concentration of human capital i.e. Tokyo could set about specializing in knowledge-intensive activities such as the R&D of high-tech products. Finally the decrease of net migration during the period 1985-95 was in part due to the recession that started in 1990 after the breakdown of the land markets bubbles.
3.1.2 The U.S case

At global level  Kim (1995) was one of the first authors to analyze the distribution of economic activities in the U.S on long series, his study over the period 1860-1987 revealed that the regional specialization in manufacturing first increased until 1947 and then declined by 32.2 \%^{8}. By using the Gini location coefficient, based on the Hoover Location Quotient (LQ for short)^{9}, this author moreover showed that this specialization and de-specialization seemed to be followed by a phenomenon of concentration (between 1860-1927) and dispersion of activities (between 1947-1987). More recently, Holmes and Stevens (2004) confirmed this point of view: indeed, by calculating the same location quotient for the years 1947 and 1999, they observed that "the term manufacturing belt is no longer appropriate for this area" since the LQ strongly decreased over the period. On the contrary, the LQ of the residual region rose from 0.45 to 0.9, which may indicate that a dispersion phenomenon grew in a few US areas. By analyzing the effects of the NAFTA on American border states, Hanson (2001) found clear evidence of a positive impact on employment in American border states. For him this indicates "that the manufacturing activities that are expanding in U.S. border cities are activities that previously took place in interior U.S. sites, NAFTA is likely to have a larger than anticipated impact on industry location". By focusing on US counties, Hanson (2005) performed a structural test of one of Krugman's (1991.a) competitive models, the Helpman (1998) model, which displays a dispersion of activities when trade is liberalized (this model is going to be discussed in the next section), and found that the estimated parameters verify the theoretical hypothesis.

At city level  This dispersion phenomenon is perhaps more obvious on an urban scale: Glaeser and Kahn (2004) showed that the spreading out of the metropolitan area, commonly known as sprawl^{10}, is an important feature of the American landscape. For those authors the automobile mainly accounts for this evolution: before this means of transport big cities were located around ports and railroad hubs, but the introduction of cars made agglomeration in the core of these urban giants less vital. With Margo (1992), the authors also conceded that rising American incomes had generated a higher demand for land, thus the concomitance of a mode of transport that reduces distance and the ability to obtain a larger lot tends to lower agglomeration density. The major concern is to find out whether this sprawl is caused by an efficient market or by distortions. For instance, the lack of marginal cost pricing in transportation can distort the price of road travel, then the use of the automobile becomes artificially too cheap, which can generate excessive urban sprawl (Walter (1961)). Thus, fighting against urban sprawl can be a good policy in this case. In an important work, Burchfield et al. (2005) analyzed the causes of urban sprawl by using satellite photography data and through this "portrait from space" they observed that urbanized land increased from 1.30\% to 1.92\% over the period 1976-1992, which can appear relatively little. However Anas and Rhee (2006) emphasized that the annual growth rate, equal to 2.48\%, represents 2.5 times the annualized growth rate of the population. We may therefore wonder about the causes of this urban sprawl. Burchfield et al. (2005) identified a large panel of causes which are i)
the specialization of cities in sectors where jobs are not located near the city center
ii) the provision of infrastructures that are more appropriate for cars than for public
transports iii) the slow population growth and the uncertainty about the future of
this growth iv) geographical elements such as absence of mountains or rugged land
in the fringe, but also a temperate climate\textsuperscript{11} v) the availability of ground water\textsuperscript{12}
vi) the presence of land beyond municipal boundaries since it allows the developer
to deviate from the city’s regulation\textsuperscript{13}.

3.1.3 The EU case

At global level Forslid et al. (2002) were the first to simulate a NEG model at
a very global level. Their model based on Haaland and Norman (1992) considered
three kinds of factors: physical capital, human capital (entrepreneurs) and labour,
which move within the 14 sectors taken into account but not within the 10 chosen
regions. Among the 14 sectors, energy and agriculture are in pure and perfect com-
petition and use sector-specific natural resources which lead to decreasing returns
to scale. Other sectors are increasing-returns-to-scale activities which evolve under
monopolistic competition. The authors analyze the impact of trade liberalization on
four of the ten regions, which are northern Europe (Finland, Iceland, Norway, Swed-
en), the southern EU (Greece, Italy, Portugal, Spain), the western EU (Benelux,
Ireland, France, the UK) and the central EU (Austria, Denmark, Germany, Switzer-
land). The most significant change in the location of production, concerns textile,
leather products, and foodstuffs. In those three sectors comparative advantages seem
to account for agglomeration. Indeed, textile and leather industries move toward the
South which has a labour comparative advantage, while food industries which have
become more and more intensive in capital, have left the southern EU for the North
and the West. The most important increasing-returns-to-scale activities which are
metals, chemicals, transport equipments and machinery, all seem to follow a non-
monotonous relationship between trade liberalization and agglomeration. This last
result is confirmed by the calculation of a concentration index which is very clearly
bell-shaped with respect to trade openness. The main criticism that can be levelled
at this model lies in the fact that parameters used in simulations are not estimated
but chosen in an ad-hoc way which is why these results should be considered with
great care.

At the national level

« Modelling the desirability of variety has been thought to be difficult, and
several indirect approaches have been adopted. The Hotelling spatial model,
Lancaster’s product characteristics approach, and the mean-variance portfolio
selection model have all been put to use. These lead to results involving
transport costs or correlations among commodities or securities, and are hard
to interpret in general terms. We therefore take a direct route » Dixit and
Stiglitz (1977, p297)

Dixit and Stiglitz (1977) proposed a ‘direct route’ to model diversity; now this
direct route is the most travelled, but the Dixit-Stiglitz monopolistic competition
has some drawbacks: firms have a myopic behaviour when they neglect the impact of their individual decisions on the industry's price indices (this drawback however vanishes when there is a continuum of firms as in Dixit-Stiglitz (1975)). Nor do they take into account the impact of their own pricing behaviour on income (what d'Aspremont et al. (1996) called the "Ford effect"). Already pointed out by Bain (1967) and confirmed by Head and Mayer (2004) for whom « oligopolistic markets seem empirically more important than those that combine atomism with product differentiation », this lack of interaction among firms does not appear totally satisfying. According to this, Combes's work (1997) is of interest since it conciliates economic geography with strategic behaviour among firms in an oligopolistic model. Moreover his goal to integrate interaction leads this author far away from the modelling presented before, indeed the preference for diversity is given up and workers are not mobile; unemployment is assumed and the equilibrium is restored after a shock by an adjustment of the number of workers employed in each market (wages are rigid). Then agglomeration causes employment and then demand to increase. In such a setting the author shows that in the long run, the conclusion of the CP model holds, agglomeration occurs endogenously when increasing returns are high and when trade is liberalized. Combes and Lafourcade (2007) enhanced this model by integrating vertical linkages between firms, accordingly the agglomeration incentive becomes even stronger since the need for intermediate inputs increases demand, and since the Cournot competition lowers the price of these intermediate inputs and thus production costs. By extracting from this theoretical framework a specification that is structurally estimable the authors were able to test their model (which is not rejected) and next to use the results of these estimations in order to make numerical simulations. This model is based on 64 French sectors over the period 1978-1993. Simulations indicate that the firms' mark-up is higher in Paris and in other agglomerated locations than in the Periphery. Yet, the sales levels are even higher in these central places, then the authors show that profits decrease monotonically with distance from the French capital. Lastly the authors simulated the impact of a 30% reduction in transaction costs and showed that after such a shock profits in Paris but also in Lyon reached a peak. The last result and the variation after the 30% shock of a concentration indicator (the Glaeser and Ellison (1997) coefficient, calculated at regional and national levels) seem to indicate that the agglomeration process decreases on a national scale but increases on a regional scale. Moreover a similar analysis was achieved by Teixeira (2006) on Portuguese data, and the author drew a very similar conclusion. As a complement to their work, Crozet (2004) showed that agglomeration forces were limited in space. In the case of Germany and Italy, the activities that the Lombardy and Bavarian regions can attract are located within a radius of approximately 100 to 150 km. Lastly Redding and Sturm (2005) worked on German cities' data over the period 1919-2002 and found that during the forty years of German division the cumulative differential of growth between border cities and other West German cities approximated 30%. All that may indicate that the Core-Periphery conclusion of agglomeration in case of lower trade costs is well adapted to the regional scale but perhaps not to the national scale.
At city level  Lastly we can remark that the urban development of some European capitals seems to follow a reversed U-curve. For instance Brakman et al. (2005) showed that the population of Madrid and Dublin increased respectively between 1900-1970 and 1840-1970, but since the seventies these two cities had seen their populations move to the Periphery. Moreover Marchand (1993) noted that the first/closer Parisian suburb (called Petite Couronne) and the second suburb (Grande Couronne) increased to the detriment of the center over the period 1920-1990.

3.2 Questioning Proposition 2

The previous studies seem to conclude that the agglomeration of activities has followed a reversed U-curve since the industrial revolution, therefore the question remains as to whether multiple equilibria are theoretical curiosities or if a region hosting an agglomeration of activities can become peripheral after a significant shock. In order to test this, Davis and Weinstein (2002) calculated the Rank Correlation (RC) between the regional density in a given year and the regional density in 1998. If we drew this RC over the period, we could notice that the shape of this curve is remarkably flat. Actually, in the year 725 this RC averaged 0.70 and attained monotonically 1 in 1998, which means that after significant structural changes which brought the Japanese economy from the feudal regime to the modern era, the rank of regions remained approximately the same. This result suggests that radical changes are not so frequent, agglomeration in one location seems to be a stable equilibrium. Moreover, in order to isolate one temporary shock that can lead to a multiplicity of equilibria, Davis and Weinstein (2002) analyzed the effects of the U.S. bombing campaign against Japanese cities. They strictly rejected the hypothesis that the growth of cities’ size follows a random path, in other words a temporary shock does not have a permanent effect. Indeed cities recovered their initial population 15 years after the shock\textsuperscript{14}. David and Weinstein (2004) went beyond this result by analyzing a more detailed database: they found that cities not only recovered their population but also their specialization. Those results need to be taken with great care because Japan is a small country where the mountains and the sea restrict people’s location. However, they are an indication that agglomeration is a stable equilibrium\textsuperscript{15}. And that is all the more so as Brakman et al. (2004) drew a similar conclusion concerning the impact of the shock of the allies’ bombing on the growth of German cities. However, by making a distinction between West and East Germany they showed that the effect of the shock was only temporary in West Germany and had a permanent effect in East Germany. This result suggests that public policies can have a significant impact on the existence of multiple equilibria. Indeed the policies of the two governments were very different, while the Federal Republic of Germany (FRG) wanted to promote a reconstruction of the country, the German Democratic Republic (GDR) focused on the creation of new attractive areas like Eisenhüttenstadt or Neuhoyerswerda. However, the authors did not confirm this suggestion: the introduction of the Federal Republic of Germany’s aid to reconstruction into the assessment tends to show that FRG policies have hindered the adjustment process by returning to a relatively pre-war city size. The reason for such a result is that the most heavily hit cities were not those which benefited from the most significant
aids. More investigations seem necessary; the authors did not include the GDR’s support in their estimation and thus they did not explain why the shock had been permanent in East Germany.

4 From agglomeration to dispersion, theoretical refinements

Empirical investigations seem to disqualify some of the conclusions of the FE model. More precisely dispersion seems to occur after agglomeration. We then propose to examine a few extensions or criticisms such as the weakness of the Home Market Effect, the introduction of heterogeneity, and the integration of land rent and commuting costs.

4.1 The weakness of the home market effect

As we have seen, the Home Market Effect (HME) is one of the pillars of the NEG, thus an important issue is to check the robustness of this effect. This test was initiated by Head et al. (2002), who found that the HME is pervasive in the sense that this effect exists in a model without product differentiation but with strategic competition among firms (the Brander (1981) model) and remains verified when iceberg transaction costs and Dixit-Stiglitz preferences are relaxed (Ottaviano et al. (2002)). However, by using the Markusen and Venables (1988) model these authors demonstrated that the introduction of the Armington assumption in the increasing returns activities was not innocuous concerning the HME. The reason for such a result is that in a world with no firm mobility, national product differentiation protects one from competition, but entrepreneurial mobility increases local competition. Indeed, when an entrepreneur migrates from the South to the North his status switches from being an imperfect competitor to being a perfect one for other northern entrepreneurs. This tends to decrease wages and thus the agglomeration incentive. More recently Head and Mayer (2004) asked several other relevant questions concerning the home market effect such as: « How do we construct demand measures in the presence of more than two countries? Indeed how does one even formulate the home market effect hypothesis? The ratios and shares of the theoretical formulations neglect third country effects ». These questions were answered in the theoretical paper of Behrens et al. (2004). The authors extended the Krugman (1980) model to a multicountry set-up, and demonstrated that the home market effect exists in such a framework only when the actual production and trade data have been corrected for the impact of asymmetrical access to world markets. This analysis raises the question of the real existence of the HME. Until now empirical works, such as those of Davis and Weinstein (1999, 2003), had approximated the HME by an index which was far from the theoretical framework, which leads Behrens et al. (2004) to write that "this regression coefficient is hard to interpret clearly". In other words, more than 25 years after its theoretical discovery, the HME only begins to be tested seriously. In the following we will survey how the weakness of the HME in the framework of the NEG changes the conclusions.
4.1.1 Imperfect elasticity of substitution of the labor supply

The HME is easily understandable: because of transport costs it is more profitable to produce in the country offering a larger market, and to export to the other. What is more puzzling is the fact that this HME increases with respect to trade liberalization. Indeed equation (16) of the HME, reported below for convenience (22), clearly shows that when the elasticity of the labor supply is equal to zero ($\frac{b_w}{b_Y} = 0$), then the HME becomes equal to $1 + \frac{1}{1+\phi}$ and thus increases in $\phi$. This paradoxical result, emphasized by Baldwin (2000), is known as the magnification effect of globalization. This author remarks that this "effect emphasizes the tendency for a general lowering of natural and manmade barriers to trade to make the remaining barriers and discrimination more important, not less important."

$$\frac{\hat{h}}{Y} = 1 - \frac{\sigma}{Z^2} + 1 - \sigma \frac{\hat{w}}{Y}$$

with $Z_\phi = \frac{1-\phi}{1+\phi}$. Notice however that this effect is thwarted by a positive elasticity of the labor supply ($\frac{b_w}{b_Y} > 0$). Indeed equation (22) shows that an increase in trade liberalization (a decrease in $Z_\phi$) makes the impact of demand ($\hat{Y}$) stronger on nominal wage ($\hat{w}$) which tends to reduce the HME. In Figure 3, we follow Head and Mayer (2004) by plotting the numerical simulations of equation (22). The horizontal axis measures $\phi$ while the vertical axis displays $\frac{\hat{h}}{Y}$, i.e. the HME. Then we can verify that the monotonically increasing HME with respect to trade liberalization (the magnification effect) only holds when the elasticity of the labor supply tends to zero. When that elasticity is high (0.04), the HME vanishes.

![Figure 3 - The Home Market Effect](image)

In the CP model this elasticity is equal to zero since it is assumed that labour is specific to each sector. By introducing arable land into the production of agricultural goods, and by assuming that entrepreneurs do not change regions but sectors, Puga
introduced such an imperfect elasticity of substitution, which decreases the HME and which exacerbates competition among entrepreneurs since firms can now substitute entrepreneurs’ labor with land when wages are relatively too high. With this modelling the author shows that the dispersion-agglomeration configuration of the CP is enhanced by a dispersion of activities when trade liberalization is high enough. Moreover, unlike the CP and FE, agglomeration (and de-agglomeration) gradually appears.

4.1.2 Other transaction costs

In the FE model, one basic assumption is that there is no trade costs in the agricultural sector. So it is important to know if this simplification is central to the conclusion of the CP. We already know from Davis (1998), which is based on Helpman and Krugman (1985) that if trade costs are introduced in the constant returns to scale activity and are high enough to impede international trade in the agricultural good, then the HME disappears. In that case Fujita Krugman and Venables (1999) showed that the dispersive equilibrium can be stable at all levels of industrial trade costs. Another interesting extension made by these authors is to integrate a specialization of each country in the production of homogeneous goods (the Armington (1969) assumption). Indeed this extension gives less categorical results since the liberalization of agricultural transaction costs generates a threefold scenario: in case of high industrial trade costs dispersion prevails, it appears in case of intermediate costs and vanishes when industrial transaction costs are low.

4.2 Land rent and commuting costs

The introduction of land rent and commuting costs was brought mainly by Krugman and Livas (1996), Tabuchi (1998), Helpman (1998), and Murata and Thissen (2005). On the one hand, Tabuchi (1998) added commuting costs and land rent to the seminal model of the NEG and then built a tale between Krugman (1991a)’s Core-Periphery and Alonso (1964)’s monocentric city. On the other hand, Helpman (1998) dropped the agricultural sector and introduced an exogenous housing stock while in Livas and Krugman (1996) the dispersive force of immobile farmers was replaced by entrepreneurs’ costly commuting need. In the last model, the authors analyzed two cities’ trade relations and their relations with the rest of the world. Murata and Thissen (2005) simplified this setting by removing the rest of the world. Here, we propose to survey the last analysis. There are two monocentric cities in this model, the North and the South, and only one sector, the Increasing Returns to Scale activity (IRS). Entrepreneurs work in this IRS sector and move from one city to the next. Those agents who own one land unit are spread along a line, and because their business is located in the middle of this line (called the Central Business District (CBD)) they need to commute. These commuting costs have a direct impact on their labour force. As each of them owns one labour unit, the total amount supplied by an entrepreneur who lives on the fringe of the CBD (i.e. at location \( x \), the CBD being at location 0 by convention) is:

\[
s(x) = 1 - 2\theta \mid x \mid
\]
where $\theta$ (with $\theta < 1$) is the commuting cost level. $|x|$ measures distance from the CBD. Furthermore, as the number of entrepreneurs is $h$, entrepreneurs’ maximal distance from the CBD is $\frac{h}{2}$, thus total labour supply net of commuting costs in one city are equal to:

$$S = \int_{-h/2}^{h/2} s(x)dx = h(1 - \theta h/2)$$  \hfill (24)

In comparison with the CP, replacing $S$ by $h(1 - \theta h/2)$ has modified the force at work, the home market effect given by (16) is now equal to

$$\frac{h}{Y} = \frac{1}{Z_h Z_\theta} - \left( \frac{\sigma}{Z_h Z_\theta} + \frac{1-\sigma}{Z_\theta} \right) \frac{\bar{w}}{Y}$$

where $Z_h = \frac{1-\theta/2}{1-\theta^2/4}$ is an inverse measure of commuting costs (there is no commuting cost when $Z_h = 1$ and very high commuting costs when $Z_h = \frac{1}{3}$ $(Z_\theta \in [\frac{1}{3}, 1])$). So by considering once again that entrepreneurs’ labour supply is perfectly elastic, $\bar{w} = 0$, we get:

$$\frac{h}{Y} = \frac{1}{Z_h Z_\theta}$$  \hfill (25)

Hence a one percent change in the northern demand for manufactures, $\frac{\hat{Y}}{Y}$, increases entrepreneurs’ employment, $\hat{h}$, by $\frac{1}{Z_h Z_\theta} (> 1)$ percent in the North. The novelty is that here the home market effect is reduced by commuting costs. Moreover the income earned by entrepreneurs is different from the CP since entrepreneurs earn a land rent and pay commuting costs. As the land rent on both edges of the segment is normalized to zero, with $w$ the entrepreneurs’ wage near the CBD, the wage net of commuting costs earned at both edges is:

$$s(\frac{h}{2})w = s(-\frac{h}{2})w = (1 - \theta h)w$$  \hfill (26)

Because consumers are identical in terms of preferences and income, at equilibrium they must reach the same utility level. Thus entrepreneurs who live in the fringe of the segment only receive a net wage of $(1 - \theta h)w$, but pay no land rent. On the contrary, workers who live near the CBD do not pay high commuting costs, but the price of the services yielded by land is higher in this location. Accordingly, the increase in real wage near central places offsets the land rent. A move from the suburb to the CBD reduces commuting and therefore increases the net wage, but also equally raises the land rent, which equalizes utility among individuals. In other words, the condition $s(x)w - R(x) = (1 - \theta h)w$ must be verified, where $s(x)$ is the total amount supplied by a worker who lives on the fringe of the CBD, $R(x)$ is the land rent prevailing at $x$, while the RHS (right-hand side) represents the wage net of commuting costs earned at both edges given by (26). By integrating expression (23) into this system we find the land rent $R(x) = \theta(h - 2 | x |)w$ with $x \in (-h/2, h/2)$. By doing so, we can find the Aggregate Land Rent (ALR) :

$$ARL = \int_{-h/2}^{h/2} R(x)dx = \theta h^2 w/2.$$  \hfill (27)

At this stage we can notice that, while on the one hand, Tabuchi (1998) assumed that there were absentee landlords, and on the other, Helpman (1998) assumed that the aggregate land rent was owned at global level, Murata and Thisse (2005) (but also Livas and Krugman (1996)) assumed that
Each entrepreneur owns an equal share of the ALR where they reside. Thus their non-salaried income is:

\[ ALR/h = \theta hw/2 \]  

In comparison with the CP, each individual’s income now incorporates the land rent and commuting costs \( Y = h(1 - \frac{n}{2}h)w \), where \( (1 - \frac{n}{2}h)w \) comes from the income of land ownership (27) and from the wage net of commuting costs \( ((1 - \theta h)w) \). Thence we can ensure that a positive variation of \( h \) increases \( Y \). Indeed we know by log differentiating this income equation that we get \( \beta Y = \beta h + \beta w \), then one percent change in entrepreneurs’ employment in the North, \( \beta Y \), increases the northern demand by \( \frac{1}{\beta w} \) percent (with \( \beta w = 0 \)). Therefore the larger manufacturing sector has a more than proportionally larger home market. Moreover from HME equation (25) we know that an increase in the size of the home market causes a more than proportional increase in the size of the manufacturing sector, this is the demand-linked circular causality already emphasized previously. However here this backward linkage decreases with commuting costs. The last difference with the CP model concerns the migration equation which becomes equal to \( \Omega = \frac{1}{w} \frac{1 - \theta h/2}{1 - \theta h/2} \left( \frac{P^*}{P} \right) \). Hence the novelty comes from the term \( (1 - \theta h/2) \) which enters multiplicatively in the indirect utility, and creates a dispersive force independently of transaction costs, which is the land market crowding effect.

In comparison with the CP, the conclusion of the model is reversed, agglomeration appears when trade costs are high, but when trade is liberalized, entrepreneurs prefer to be dispersed as this spatial configuration allows them to alleviate the burden of commuting costs.

### 4.3 Taste heterogeneity

Murata (2003) and Tabuchi and Thisse (2002), improved the NEG’s conclusions by introducing taste heterogeneity in the choice of location. This taste heterogeneity is central since the common assumption that all individuals share the same preferences is highly implausible. As the authors pointed out, mobile people are attached to their regions, not only for financial reasons but also for non-market attributes (local or social amenities: climate, culture, family etc.), thus such heterogeneity can weaken agglomeration forces. More precisely, by using a probabilistic migration dynamics borrowed from the discrete choice theory, the authors show that a gradual or partial agglomeration arises from the dispersion of activities and then gives rise to a gradual re-dispersion when trade is liberalized. Another interesting result suggested by this model is that non-market interests can be higher when individuals reach a certain level of real income, which indicates that people’s mobility may decrease with the development level.

### 4.4 Firms’ heterogeneity

So far we have only analyzed entrepreneurs’ mobility. Yet, in many cases entrepreneurs are not mobile, they simply invest in another country and the reward of their
capital is repatriated. Martin and Rogers (1995) were the first to analyze such a situation. Their model, where entrepreneurs’ mobility is replaced by capital mobility, is called the Footloose Capital (FC for short). The FC is very close to the FE. One significant difference is that entrepreneurs do not own labour but capital, thus in all the equations describing the FE model in the first section the label $w$ needs to be replaced by $\pi$, which is the reward to capital. Then, capital mobility is ruled by the following equation:

$$\Omega = \frac{\pi}{\pi^*}$$

where price indices have disappeared since entrepreneurs consume where they reside and not where they invest. Unlike the FE, the fact that the capital is repatriated cuts out demand-linked circularity as the shift in production cannot lead to a shift in expenditure, and also rules out cost-linked circularity since the cost of living effect does not impact on the equation of capital mobility. In terms of conclusion, the FC model does not verify either Proposition 1 (departing from dispersion trade liberalization never generates any agglomeration of the population) or Proposition 2 (there are no multiple locally stable equilibria except for free trade where each location defined by $S \in [0,1]$ is a stable equilibrium). However, by assuming that countries’ size are asymmetric, Baldwin (1999) showed that agglomeration appeared gradually with respect to $\phi$. This is illustrated by Figure 4 where the dashed line represents the result of this asymmetric FC model. In such a case, the agglomeration rent ($\Omega > 1$) also exists after $\phi^P$. However things are different if we integrate firms’ heterogeneity. Such work initiated by Melitz (2003) was used in the FC model by Baldwin and Okubo (2005). Heterogeneity was introduced via marginal production costs. Then each firm produced a variety with a specific labour input coefficient, denoted $\beta_i$ for a typical firm $i$. The innovation technology that permits to obtain a particular $\beta_i$ is stochastic in the sense that this marginal cost is randomly distributed to a firm. More precisely the distribution of $\beta_i$ is assumed to be a Pareto probability distribution. Unlike the classical Dixit-Stiglitz model, selling a new variety requires firms to pay fixed costs to enter local and export markets. These ‘beachhead’ costs can be advertising costs, rules of origins (see Demidova et al. (2005)), etc. As a result of such modelling, three kinds of firms are distinguished: the X-type which enjoys low marginal costs ranked between 0 and $\beta_X$ and which pays a beachhead cost $F$; this type of firm sells on both the local and foreign markets. The D-type which produces under intermediate marginal costs ranked between $\beta_X$ and $\beta_D$, and also pays a beachhead cost $F$. This type of firms is not competitive enough to sell abroad, so they restrict their objective to the domestic market. Lastly the N-type has to produce a variety with high marginal costs ranked between $\beta_D$ and $\beta_0$; these marginal costs are even too high to sell varieties on the local market. These beachhead costs which are sunk require firms to take into account their future operating profits, the authors assume that the discount rate is equal to zero and that the firm’s death process follows a Poisson law. With this setting, Baldwin and Okubo consider two polar cases, the case where delocation is very slow and free/entry is fast, which is a simplified version of the Melitz model, and the case where delocation is high and free/entry slow, which is a modified version of the FC model with heterogeneity and beachhead costs. Here we are going to focus on the
last case. Authors assume that the North is initially larger than the South and show that when trade gets freer, the South’s most efficient firms start moving gradually toward the North. Then after a certain level of trade liberalization, denoted \( \phi^P \), all the X-type firms have moved, and the move of D-type firms begins. If the market entry cost is low enough, then a total agglomeration can happen in the North. Conversely with a high entry cost only a partial agglomeration occurs. Diagram 4 illustrates these findings. The next challenge for the NEG literature is perhaps to link this firms’ heterogeneity with the process of international fragmentation of the supply chain (see Baldwin (2006) for a survey on this subject).

![Figure 4 – The tomahawk diagram](image)

### 5 Public Policy

#### 5.1 Individual welfare

Up to now we have only analyzed entrepreneurs’ relative welfare in order to know where they wish to reside. Charlot et al. (2006) propose a deeper analysis by studying the individual welfare of the four interest groups (S northern entrepreneurs, \( S^* \) southern ones, L northern workers and \( L^* \) southern ones) which are given by:

\[
\begin{align*}
V_S(h, \phi) &= wP^{-\mu}, & V_S^*(h, \phi) &= w^*P^{*-\mu} \\
V_L(h, \phi) &= P^{-\mu}, & V_L^*(h, \phi) &= P^{*-\mu}
\end{align*}
\]

Notice that Charlot et al. (2006) work with the CP model, while here we use the FE model for convenience. Thus \( P \) and \( P^* \) are given by (9) with \( m = 1, \varepsilon = \alpha = 1 \) and \( w, w^* \) are found by resolving the market clearing condition (13) with \( m = 1, \varepsilon = \alpha = 1, f = 1 \). The objective is to analyze individual welfare under the opposite equilibria of agglomeration and dispersion, and then to compare these equilibria in
order to determine which is the better social outcome. When all entrepreneurs are located in the North $S = 1 - S^* = 1$, this group’s welfare is given by:

$$V_S(1, \phi) = \frac{2bL}{1-b}$$  \hfill (28)

According to this expression, entrepreneurs do not take care of transaction costs and this is easily understood as they have nothing to import. Immobile workers in the North share the same indifference to transaction costs, while in the South these costs have a real importance, their decrease, being welfare-enhancing:

$$V_L(1, \phi) = 1, \quad V^*_L(1, \phi) = \phi^a$$  \hfill (29)

with $a = \mu/(\sigma - 1)$  \hfill (30)

Under dispersion entrepreneurs’ and workers’ welfare is given by:

$$V_S\left(\frac{1}{2}, \phi\right) = V^*_S\left(\frac{1}{2}, \phi\right) = \frac{2bL(1 + \phi)/2}{1-b}$$  \hfill (31)

$$V_L\left(\frac{1}{2}, \phi\right) = V^*_L\left(\frac{1}{2}, \phi\right) = \left(1 + \frac{\phi}{2}\right)^a$$  \hfill (32)

An entrepreneur prefers agglomeration to dispersion when $V_S(1, \phi) > V_S\left(\frac{1}{2}, \phi\right)$, then by using (28) and (31) we can verify what happens if $\phi < 1$, which means that:

**Proposition 3** Whatever the level of transport costs, entrepreneurs prefer agglomeration to dispersion.

Notice however that through their FEVL where entrepreneurs are intersectorially mobile but cannot move from one nation to the other Ottaviano and Robert-Nicoud (2006) showed that this proposition is found in the Core, but not in the Periphery where trade costs have to be low enough to verify this result.

With regard to workers, those in the North prefer agglomeration to dispersion on condition that $V_L(1, \phi) > V_L\left(\frac{1}{2}, \phi\right)$. Such a result is obtained through (29) and (32) when $\phi < 1$. Southern workers prefer agglomeration to dispersion on condition that $V^*_L(1, \phi) > V^*_L\left(\frac{1}{2}, \phi\right)$. Such a result is obtained through (29) and (32) when $\phi > 1$, which is impossible. This implies that whatever the value of trade costs, peripheral workers always prefer dispersion.

This provides the following proposition in the FE:

**Proposition 4** Regional conflict of interests: whatever the value of trade costs, an agglomeration’s inhabitants (entrepreneurs and workers) prefer agglomeration while peripheral workers prefer a dispersive equilibrium. Then no agreement is possible for trade liberalization which is a non Pareto improving policy.

Since a unanimous preference either for agglomeration or dispersion is impossible, the authors propose to follow Kaldor (1939), Hicks (1940) and Scitovsky (1941) in their cost-benefit analysis. For Kaldor a situation A (Agglomeration) is preferred to a situation D (Dispersion), if the gainers of the reform that lead from D to A
(trade liberalization) can compensate for the losers of this reform. Ex-ante the Hicks criterion considers that a situation A (Agglomeration) is better than a situation D (Dispersion), if the losers of the change (D → A) cannot bring the gainers to give up this reform. As a result, the Kaldor criterion concludes that the agglomerative equilibrium is socially desirable only when trade costs are low enough, while according to the Hicks criterion agglomeration is always desirable. This means that the former criterion considers dispersion as the best outcome when trade costs are high while the latter argues the reverse. According to the Scitovsky criterion, a situation A is better than a situation D only if the two previous criteria agree with that proposition, then here we are typically in a situation of indetermination. Concerning the FEVL, Ottaviano and Robert-Nicoud are faced with the same indetermination problem. Then the cost-benefit analysis does not shed much light on the new economic geography model. In order to tackle the problem, Charlot et al. turns to Rawls and Bentham. In particular they use the Social Welfare Function proposed by Atkinson (1970), which allows one to distinguish between the Utilitarian and the Rawlsian point of view:

\[
W = \frac{1}{1-\eta} \left[ L(V_L)^{1-\eta} + L^*(V_L^*)^{1-\eta} + S(V_S)^{1-\eta} + S^*(V_S^*)^{1-\eta} \right]
\]

where \( \eta \) is the aversion to inequality: with \( \eta = 0 \) the government is utilitarian and maximizes the sum of individual indirect utilities while with \( \eta = +\infty \) it is Rawlsian, and then maximizes the utility of the poorest. Authors find that a utilitarian government always prefers agglomeration when varieties are sufficiently differentiated. However, when these varieties are good substitutes dispersion becomes the best equilibrium under a critical value of trade liberalization. Since this critical value is lower than the critical value obtained by Kaldor, a utilitarian government is going to sustain the agglomeration for a wider range of \( \phi \) than the government advised by Kaldor. Accordingly the authors consider that the utilitarian behavior is biased regardag agglomeration and "cautions against the use of utilitarian welfare functions as a foundation for regional policy recommendations". Ottaviano and Robert-Nicoud find the same result when vertical linkage and trade costs are low in the FEVL. Furthermore in the two models the Rawlsian government is always in favour of dispersion. In other words, the evaluation of the market outcome depends hugely on societal values. Lastly Fujita and Thisse (2002) showed that this proposition of regional conflict of interests can disappear if the growth gained from agglomeration is significant enough. However if the boost given by agglomeration is not strong enough then the same remark concerning societal value holds.

5.2 Tax competition

5.2.1 Theoretical consideration

The Basic Tax Competition model (BTCM) emphasizes that a race to the bottom in tax rates can occur when capital mobility increases. However, the New Economic Geography literature has put this result into perspective: beside the mobility of factors, the mobility of goods also matters. In a world of increasing returns, and
in order to be close to demand and to minimize transaction costs firms have strong incentives to be agglomerated in central places. Such agglomeration generates a rent for mobile factors, which allows the government that hosts this agglomeration to increase its taxes without capital flight. By means of the FE model when $h = 1$ this rent is given by the following expression:

$$\Omega(1, \phi) = \frac{2\phi^{1-a}}{1 + \phi^2 - b(1 - \phi^2)}$$

To see how this expression varies with transactions, a log differentiation of it gives:

$$\frac{d\Omega}{d\phi} = (1 - a) - \frac{\phi^2(1 + b)}{1 - (1 - \phi^2)(1 + b)/2}$$

this expression increases until $\phi^{\max} = \sqrt{\phi_b}$ and decreases later. This is summarized by the following result:

**Proposition 5** *The agglomeration rent is bell-shaped with respect to trade liberalization.*

This agglomeration rent, which is linked to trade liberalization in a non monotonic way, has generated a flourishing literature concerning the trystic trade integration / imperfect competition / agglomeration (Andersson and Forslid (2003), Baldwin and Krugman (2004), Kind et al. (2000), Ludema and Wooton (2000), Ottaviano and Van Ypersele (2002)). Indeed because this rent successively increases and decreases with the freeness of trade, a race to the top can be followed by a race to the bottom. In order to understand this result we propose to survey Baldwin and Krugman (2004)’s work, since this model is the only one that has been tested empirically. In order to take into account the difference of size between the Core and the Periphery, Baldwin and Krugman (2004) analyzed the governments’ tax policies in case of total agglomeration of activities in the North, so they henceforth limited themselves to $\phi > \phi_b$. Preferences are given by $U = M^\mu A^{1-\mu}G$, where $G$ is the supply of local public goods. With Andersson and Forslid (2003) it is assumed that public goods are produced via private goods. Thus the composition of demand and all the variables that have been analyzed so far (wages and prices) are not affected by tax. We assume that each jurisdiction supplies the same amount of public goods but the financing of these goods $f$ and $f^*$ can differ from one jurisdiction to the next (implicitly each government’s efficiency can vary):

$$f = tY, \quad f^* = t^*Y^*$$

Because the supply of public goods is the same everywhere, migration stops when post-tax reward is higher in the Core. The location equilibrium condition is therefore given by:

$$\frac{V}{V^*} \frac{1 - t}{1 - t^*} \geq 1$$

We then assume that governments maximize the following objective function:

$$W = W(f, t), \quad W_f > 0$$
As has been pointed out by Baldwin and Krugman (2004), denoted BK for short, the common point between the Leviathan and a benevolent government is that their objective functions rise with the collected revenue and decline with the tax rate. Nevertheless, because the tax rate also has an impact on the revenue, a shift of revenue first has a positive effect on the objective and then a negative one. Thus, the objective function is necessarily a bell-shaped curve when the tax rate increases. Furthermore, the local governments play a Stackelberg game whose sequence is the following: i) The North sets a tax rate \( t \) ii) The South reacts through a tax rate \( t^* \) iii) Migration occurs. This game is resolved by backward induction. The third stage is already known by the previous analysis of the model. The second stage is summarized by Figure 5. The vertical axis represents the objective function of the South, while the horizontal axis plots the taxation level in this country. There are two possible situations: if the South does not succeed in stealing the Core, then its revenue is small because it can tax only immobile workers (thus equal to \( t^* L w_a \)), on the contrary, if this country succeeds in attracting all the activities, then its revenue is going to be much higher (\( t^*(w^* + L^* w_a) \)). Total agglomeration in the North depends therefore on the potential taxation levied by the South. The Northern government can indeed be upset by a break-point tax rate \( t^*_b = 1 - \Omega(1, \phi)/(1 - t) \), which enables the South to break the Core equilibrium.

Thus, if we start from a situation where the Core/North sets a high tax rate, say \( t' \), then the break-point tax rate, denoted \( t^*_{b'} \) on the diagram is also high, and thus the South can steal the Core by setting \( t^* \) equal to \( t^*_{b'} \). If the North decides to choose a lower tax level, say \( t' \), then in such a case the South can steal the Core by lowering \( t^* \) to \( t^*_b \). However, in that case \( t^*_b \) is equivalent to the tax rate that maximizes the objective function of the South without the Core, namely \( t^*_u \), so the South is indifferent either to the Core or the Periphery. Therefore, in the first stage, if the North wants to keep the Core, it needs to set a tax below or equal to \( t' \), such that the South does not want to deviate from \( t^*_u \). This limit tax rate, denoted \( t_e \), is thus equal to \( t' \) on the diagram and analytically found by:

\[
t_e = 1 - \frac{1 - t^*_b}{\Omega(1, \phi)}
\]

where \( t^*_b \) s.t \( W^*[t^*_b Y^*(0, \phi), t^*_u] = W^*[t^*_u Y^*(1, \phi), t^*_u] = W^*_e \)

We now need to verify that the North always prefers to have the Core. This is ensured by the fact that the North wins \( W_e \) with the Core and nothing without it, accordingly the North will always "limit tax" the South. By using a specific objective function, \( W(f, t) = G - t^2/2 \), and by applying the above described game, the authors determine the differential \( t_e - t^*_u \) which represents the tax competition level:

\[
t_e - t^*_u = 1 - L - \frac{(1 + \sqrt{b}) - (1 - \sqrt{b})L}{(1 + \sqrt{b})\Omega(1, \phi)}
\]

This gives the following proposition:

**Proposition 6** Trade liberalization leads to an increase in the tax gap between the Core and the Peripheral regions at first and then to a race to the bottom in terms of taxation.
Proof. Indeed we get \( \frac{\partial(t_t - t_t^*)}{\partial \phi} = C \frac{\partial \Omega(1, \phi)}{\partial \phi} \) with \( C = \frac{(1+\sqrt{b})-(1-\sqrt{b})L}{(1+\sqrt{b})\Omega(1, \phi)^2} > 0 \) and we know from proposition 5 that \( \frac{\partial \Omega(1, \phi)}{\partial \phi} \) is at first positive and then negative when \( \phi \) increases. ■

\[ W, W^\ast \]
\[ W^\ast \text{ (core in South)} \]
\[ W \text{ (core in North)} \]
\[ W_0 \]
\[ W_a \]
\[ W_e \]
\[ t^* \]
\[ t \]
\[ t^* \]
\[ t_b \]
\[ t_a \]
\[ t_e \]
\[ t_t = 1 + \frac{(1 - t_t^*)}{\Omega(1, \phi)} \]

Figure 5 – The tax game

5.2.2 Empirical verification

The previous proposition is a strong result which goes against the conclusion of the basic tax competition model but it is also a result that can disappear in a model à la Krugman and Livas (1996) where the agglomeration rent is strictly decreasing in \( \phi \). A formal proof of this can be found in Candau (2006). Then it is necessary to test this proposition empirically. Such work has been undertaken by Gilbert et al. (2005), who propose to test the following equations:

\[
t_t - t_t^* = \alpha_1 \phi_t + \alpha_2 \phi_t^2 + F + \epsilon_t \tag{35}
\]

\[
t_t - t_t^* = \alpha_1 \phi_t + \alpha_2 \phi_t^2 + \alpha_3 (Y_t - Y_t^*) + F + \epsilon_{ijt} \tag{36}
\]

\[
t_t - t_t^* = \alpha_1 \phi_t (Y_t - Y_t^*) + \alpha_2 \phi_t^2 (Y_t - Y_t^*) + F + \epsilon_{ijt} \tag{37}
\]

\[
t_t - t_t^* = \alpha_1 \phi_t + \alpha_2 \phi_t^2 + \alpha_3 (Y_t - Y_t^*) + \alpha_4 \phi_t (Y_t - Y_t^*) + \alpha_5 \phi_t^2 (Y_t - Y_t^*) + F + \epsilon_{ijt}
\]
where \( t_t - t^*_t \) and \( Y_t - Y^*_t \) represent respectively the tax gap and the Gross Domestic Product (GDP) gap between two countries at the period \( t \). This estimation is based on the period 1982-2001 and tested on the EU15 and EU25. For the EU15 the authors use successively the Statutory Tax Rate, the ex-post tax rate, the Effective Average Tax rate and the Effective Marginal Tax rate, while only the former two are available for the EU25. \( F \) combines fixed effects associated with countries and time: \( F = \nu + \nu^* + \nu_t \). For all these estimations, the following constraint is set:

\[
Y_t > Y^*_t
\]

This constraint allows one to distinguish between the Core and the Periphery. Then in the first equation the authors simply test if the tax gap is bell-shaped with respect to \( \phi \). In the second one they include the GDP gap in order to understand if this gap plays an independent role on \( \phi \) in the shape of the tax gap. In the third equation the authors want to test if the GDP gap and \( \phi \) interact on the evolution of the tax gap. Lastly the fourth equation tests simultaneously all the assumptions presented above. Concerning the measure of \( \phi \), the authors adopt the method developed by Head and Mayer (2004), and then use the following measure \( \phi = \sqrt{c_{ji} c_{ij}} \). The interest of this method is that \( c_{ji} \), \( c_{ij} \) are bilateral trade data between country \( i \) and country \( j \), which are easily available, and \( c_{ii} \) and \( c_{jj} \) are approximated by the national production minus the total exportation of the country. The results of equations (35) and (36) reveal that the tax gap measured by the effective average tax rate and the effective marginal tax rate follows a bellshape with respect to \( \phi \) for the EU15. When it comes to statutory tax rate and ex-post rate, the relationship is less robust for the EU15 and EU25. However equation (37), which takes into account interaction between the tax gap and the GDP gap is significant for the two samples. This means that the bell-shaped tax gap tends to be flatter when there is little difference between countries, or in other words, the more countries’ sizes differ, the higher the bell-shaped tax gap with respect to \( \phi \). This result is interesting and may deserve to be studied on a theoretical angle. Indeed Baldwin and Krugman (2004) only considered the case of total agglomeration in one country and thus implicitly studied the evolution of the tax gap between countries that are diametrically different. Then we can imagine that if these countries were less different, then the bell-shape would be less pronounced. But actually this is not the case: indeed the sole model that analyzes tax competition under partial agglomeration, i.e. the Borck and Pfüger (2006) model, found that the tax gap strictly increases when trade is liberalized. Moreover, we wished that the empirical study of Gilbert et al. were based on aggregate data. As a matter of fact the governments’ choice of taxes is certainly made by considering the most mobile firms, and then a definition of \( \phi \) at industry level (as in Head and Ries (2001), Head and Mayer (2004) etc) would improve the analysis. Besides, Gilbert et al. use their estimation results in order to calculate the integration level that allows one to obtain the maximal tax gap between two countries. They found that in 2000 the new members of the EU25 were far from this maximum level of \( \phi \) beyond which a race to the bottom is predicted.
5.3 Some extensions and propositions

5.3.1 Tax competition

In this section, we propose two extensions of the literature on tax competition. The Baldwin and Krugman (2004) model analyzed the effect of identical and reciprocal trade liberalization on tax competition. However, in Europe, the enlargement starts with preferential and non-reciprocal trade liberalization. For instance, after the collapse of communism the CEECs were integrated into the General System of Preference, and soon afterwards they signed the European Association Agreements and more recently the Free Trade Agreements. Then we are going to show that such liberalization changes the issue concerning tax competition. Secondly, the EU fosters economic integration by liberalizing trade between countries, but also by investing in internal infrastructure inside peripheral regions. One may wonder to what extent such a policy can affect tax competition.

Unilateral trade policy Instead of analyzing reciprocal trade, we consider what happens when market access to the Core is given by $\phi$, while market access to the South is given by $\phi^*$. With such trade costs, price indices and nominal wages are now given by:

$$P_1^{1-\sigma} = h + \phi^* h^*,$$
$$w = b\left(\frac{Y}{P_1^{1-\sigma}} + \phi^* \frac{Y^*}{P^{1-\sigma}}\right), \quad w^* = b\left(\phi^* \frac{Y}{P_1^{1-\sigma}} + \frac{Y^*}{P^{1-\sigma}}\right)$$

In that case the agglomeration rent is:

$$\Omega = \frac{2\phi^{1-a}}{1 + \phi^{a} + b(\phi^{a} - 1)}$$

This expression allows one to observe that when the Core liberalizes its market to the Periphery then the agglomeration rent decreases. Indeed by differentiating $\Omega$ with respect to $\phi$ we obtain:

$$\frac{\partial \Omega}{\partial \phi} = -\frac{2(1 + b)\phi^{2-a}}{(1 + \phi^{a} + b(\phi^{a} - 1))^2} < 0$$

Concerning tax competition, the tax gap is still given by equation (34) (but with $\Omega$ given by (40)), which gives the following result:

Proposition 7 Non-reciprocal trade liberalization of the Core leads countries to launch a race to the bottom.

Proof. Indeed we get $\frac{\partial(t_u - t^*_u)}{\partial \phi} = C\frac{\partial \Omega}{\partial \phi}$ with $C = \frac{(1 + \sqrt{\beta} - 1 - \sqrt{\beta})L}{(1 + \sqrt{\beta})\Omega^2} > 0$ and we know from (41) that $\frac{\partial(1, \phi, \phi^*)}{\partial \phi} < 0$ thus $\frac{\partial(t_u - t^*_u)}{\partial \phi} < 0$. A corollary to this proposition is that trade liberalization by peripheral countries foster a race to the top. This is especially important in the European debate because the Core countries have liberalized most of their trade in manufactures, whereas "accession countries" are only catching-up. Therefore, Core countries could be under less pressure to low those tax rates.
Improvement of internal infrastructure in the Periphery We are going to consider that trade between regions is reciprocal and given by as in Baldwin and Krugman, but that trade inside nations implies internal costs $\phi_I$ in the North and $\phi_I^*$ in the South. Thanks to that, price indices and nominal wages are given by:

$$
P^{1-\sigma} = \phi_I h + \phi h^*, \quad (P^*)^{1-\sigma} = \phi h^* + \phi_I^* h^*$$

$$
w = b(\phi_I^* \frac{Y}{P^{1-\sigma}} + \phi^* \frac{Y}{P^{1-\sigma}}), \quad w^* = b(\phi^* \frac{Y}{P^{1-\sigma}} + \phi_I^* \frac{Y}{P^{1-\sigma}})$$

In such a case the agglomeration rent is the following:

$$
\Omega = \frac{2\phi_I^{1+a} \phi^{1-a}}{(1+b)\phi^2 - (b-1)\phi_I^* \phi_I^*} \quad (42)
$$

From this expression we are able to show that when a supranational policy improves the internal infrastructure of the Periphery, then the agglomeration rent decreases. Indeed, by differentiating $\Omega$ with respect to $\phi_I^*$ we obtain:

$$
\frac{\partial \Omega}{\partial \phi_I^*} = \frac{2(b-1)\phi_I^{2+a} \phi^{1-a}}{((1+b)\phi^2 - (b-1)\phi_I^* \phi_I^*)^2} < 0 \quad (43)
$$

which is negative since $b = \frac{\mu}{\sigma} < 1$.

With respect to tax competition, the tax gap is still given by equation (34) but with $\Omega$ given by (42), which yields the following result.

**Proposition 8** A supra national policy that improves the internal infrastructure of the Periphery leads countries to launch a race to the bottom.

**Proof.** We have $\frac{\partial (t^* - t^*)}{\partial \phi_I^*} = C \frac{\partial \Omega}{\partial \phi_I^*}$ with $C = \frac{(1+\sqrt{\beta}-(1-\sqrt{\beta})L}{(1+\sqrt{\beta})\Omega^2} > 0$ and we know from (43) that $\frac{\partial (t^* - t^*)}{\partial \phi_I^*} < 0$, thus $\frac{\partial (t^* - t^*)}{\partial \phi_I^*} < 0$.

Ireland’s past policy in terms of taxation can perhaps be understood in this light. Then it is understandable that countries like France are trying to restrain tax competition from countries that benefit from the EU regional policy. However, the magnitude of this effect at cross-country level in the EU is bound to be negligible, given the amount of the EU budget devoted to regional policies.

5.3.2 Regional infrastructure and Growth

So far we have talked about governments’ aim to attract as many activities as possible. Such a policy is often aimed at promoting national growth. Surprisingly, the FE model with endogenous growth built by Fujita and Thisse (2002) has never been used in order to illustrate how policies can affect the growth rate. The sole result that is available concerns the mobility of the physical capital. For instance by using the FC model, Martin (1998, 1999) and Baldwin et al. (2003) demonstrated that public policies that aimed to improve internal infrastructure in the Periphery were detrimental to growth while policies that succeeded in spreading the technological knowledge of the North to the Periphery fostered global growth. But what happens
on a urban scale? How is economic growth affected when entrepreneurs flee the Core because, as we have seen in the Murata and Thisse model, the cost of living is higher in the agglomerative area? Dealing with these kind of questions constitutes one of the main theoretical and empirical challenges that the NEG will surely have to face in the future.

5.3.3 Welfare

As has been shown, Charlot et al. (2006) demonstrated that agglomeration and dispersion could not be Pareto ranked while Ottaviano and Robert-Nicoud (2006) showed that when the Core was diversified while the Periphery was specialized then agglomeration Pareto-dominates dispersion when trade costs are low enough. This Pareto domination was also checked by Robert-Nicoud (2006) in a model where entrepreneurs intersectoral mobility had been replaced by physical capital mobility. The sole article that does not reach consensus on the desirability of agglomeration is Gaigné (2006) who used the FEVL model (on condition that the Core be fully specialized in the manufacturing sector while the periphery be specialized in the traditional one), to demonstrate that when trade costs are low enough, dispersion Pareto-dominates full agglomeration. That result is important because it proves that the New Economic Geography cannot be used as a tool that exclusively encourages a polarization policy in a globalized world. We want to confirm this result by performing a welfare analysis in Murata and Thisse (2005)’s model. This analysis which is missing in the seminal paper, can easily be carried out since this model, based on the CP one, has already been presented previously. Consequently, we find that:

**Proposition 9** Whatever the level of regional transaction costs, dispersion Pareto dominates agglomeration.

**Proof.** At the dispersed equilibrium \( S = S^* \), price index are given by \( P^{1-\sigma} = \frac{S w^{1-\sigma} (1+\phi)}{\sigma} \) thus by using (24) we get \( P = \left( \frac{h(1-h/2)}{\sigma} \right)^{1/(1-\sigma)} w \). Then the indirect utility in the North \( \mathcal{V}(h, \phi) = \left( \frac{1-h/2}{\sigma} \right)^{1/(1-\sigma)} (1 - h/2)^{1+1/(\sigma-1)} \). Conversely, when all entrepreneurs are agglomerated in the North, the price index in this location is \( P^{1-\sigma} = \frac{S w^{1-\sigma}}{\sigma} \) so we get \( P = \left( \frac{h(1-h/2)}{\sigma} \right)^{1/(1-\sigma)} w \) and with \( h = 1 \) entrepreneurs’ indirect utility is \( \mathcal{V}(1, \phi) = \left( \frac{1}{\sigma} \right)^{(1/(\sigma-1))} (1 - \phi)^{1+1/(\sigma-1)} \). Then entrepreneurs prefer dispersion if \( \mathcal{V}(1, \phi) > \mathcal{V}(1, \phi) \) which gives \( \left( \frac{1}{\sigma} \right)^{(1/(\sigma-1))} \left( 1 - \frac{\phi}{2} \right)^{1+1/(\sigma-1)} > \left( \frac{1}{2} \right)^{(1/(\sigma-1))} (1 - \frac{\phi}{2})^{1+1/(\sigma-1)} \). This is equivalent to \( \phi > Z_\sigma^0 - 1 \) and as \( Z_\sigma \in [\frac{1}{2}, 1] \), this inequality is always verified. Consequently entrepreneurs prefer dispersion for all \( \phi \). \( \blacksquare \)

It is worth recalling that this result is obtained under the assumption that land rents accrue not to absentee landowners, but fully to residents. Therefore when the market forces lead to an agglomeration of activities, we are in a situation akin to a Prisoner’s dilemma. Then, that result warns us that the desirability of agglomeration merits closer consideration.
6 Concluding remarks

Propositions that were initially put forward by the NEG have been questioned by empirical works which seem to correct the idea that agglomeration occurs when trade liberalization tends towards free trade and are also sceptical as to the hypothesis of multiple equilibria. These investigations and criticisms have marked a new starting point for theoretical works, which now integrate land rent, commuting costs, endogenous growth and have investigated the public policies that the NEG may recommend. We have thus reached a point where empirical works need to test these recommendations in particular on the ground of tax competition. Besides this conclusion, it is also worth noting, that theoretical literature still has to take up many challenges. For instance we have seen that an NEG model that can build a tale between Alonso’s model and Grossman-Helpman-Romer’s model, may be a first interesting stage to understand how migration, cities’ economic growth and trade interact. Concerning welfare the literature presents agglomeration as a process that is unambiguously good for entrepreneurs, and we have seen that such a point of view is questionable. For instance, in Japan agglomeration in some cities like Tokyo, Osaka and Nagoya has been considered as excessive by some entrepreneurs who have fled those cities around the year 1965 on account of increasing wages and land prices. Then by using the Murata and Thisse (2005) model we have shown that urban costs are not inoffensive concerning conclusions on welfare. We have in particular demonstrated that agglomeration is not always a Pareto efficient outcome. That result which contrasts with the literature leads us to question the desirability of agglomeration. Besides such a statement is motivated by the fact that on the one hand the literature works with a Benthamine government (Pflüger and Südekum (2006), Baldwin et al. (2003)) whereas Charlot et al. "cautions against the use of utilitarian welfare functions as a foundation for regional policy recommendations". On the other hand, Charlot et al. make a cost-benefit analysis à la Kaldor (1939), Hicks (1940) and Scitovsky (1941) while many specialists in the field of welfare economics (such as Arrow (1951) and Sen (1979)) consider that this kind of analysis must be avoided due to their irrelevance and unfair conclusion\(^1\). We believe that a welfare analysis in the spirit of Pazner and Schmeidler (1978) and Samuelson (1977) via the concept of egalitarian-equivalence\(^2\) can shed more light on the desirability of agglomeration, which is one of the main questions that decision-makers wish to address.

Acknowledgements: The author thanks Marc Fleurbaey, Carl Gaigné, Frédéric Robert-Nicoud, Jacques Thisse, Jacques Le Cacheux and Lionel Fontagné for helpful comments.

Notes

\(^1\)For a survey of these three determinants see Fujita and Thisse (1996) and Ottaviano and Thisse (2005)
\(^2\)Models that deal with capital mobility have been widely surveyed in Baldwin et al. (2003). Since the publication of this monograph no significant improvements have been made except one concerning the firms’ heterogeneity, which is the subject of a large digression.
LOCATION AND PUBLIC POLICIES: A SURVEY OF THE NEG

$\phi_s$ is given implicitly by $1 = \phi_s^{\alpha}(\phi_s + 1 + \frac{1 - \mu}{1 - \alpha})$ with $\alpha = \mu/(\sigma - 1)$.

Concerning forward-looking behavior, Krugman (1991b) was the first to emphasize the importance of expectations versus history, however its model is a pre-economic geographical model. Indeed in place of pecuniary externalities, technological externalities are used as a central force of agglomeration. Ottaviano (1999) and Ottaviano et al. (2002) take one step beyond by linking pecuniary externalities with forward-looking behavior in the linear version of the C-P model while Baldwin (2001) argued that the richness of the C-P model (overlap between break point and sustain point) was a good basis to investigate through numerical simulations the role of expectations in the decision of location. Between those two approaches, Ottaviano (1999) proposed a model which was close to the FE model that we have studied so far, and which emulates the C-P results. From those papers an important conclusion emerges: forward-looking expectations are equivalent to myopic behavior in case of high migration costs, workers' impatience, and weak agglomeration forces; conversely, a specific (agglomerate or dispersive) equilibrium could be locally stable (between $\phi_s$ and $\phi_a$) but globally unstable owing to the mere shift in expectation.

Authors confirm the conclusion of the Footloose Capital model with endogenous growth: geography and growth are linked, an increase in the growth rate fosters agglomeration and agglomeration improves the growth rate.

Since entrepreneurs are forward looking the Tomawak diagram has to be described carefully. We may consider that trade policy changes unexpectedly or as the authors argued "another way may be to assume that the transport parameter changes very slowly in comparison with the actual working of market processes, and hence the actual configuration of the economy at each time can be approximated by that of the ss-growth or quasi ss-growth equilibrium under the corresponding value of trade costs".

This result is obtained at the three-digit level, at the two-digit level the decreasing rate approximates 47% over the period. This indicates a bias in the aggregation.

Hoover's location quotient is defined by:

$$LQ_{ij} = \frac{E_{ij}}{E_{ius}} / \frac{E_j}{E_{us}}$$

where $E_{ij}$ is employment in industry $i$ for region $j$, $E_j$ is total employment in region $j$, $E_{ius}$ is total employment in industry $i$, $E_{us}$ is total employment in the USA. Then the Lorenz curve of location disparities is obtained i) by calculating the LQ of one industry $i$ for all regions $j$ ii) by ranking the regions according to their LQs in descending order iii) by calculating the cumulative percentage of employment in industry $i$ over the regions iv) by calculating the cumulative percentage of employment in total manufacturing over the region. If industry $i$ is perfectly dispersed over space, then the location quotient is equal to one, and the location curve follows a 45-degree line; on the contrary if industry $i$ is agglomerated in some regions, then the location curve becomes concave. The Gini coefficient is defined as the area between the 45-degree line and the location curve divided by the triangular area. This Gini coefficient has many drawbacks, and needs to be considered carefully. Gill and Glaeser (1997) and Maurel and Sédillot (1999) highlighted the problem of such classical indices and proposed a better coefficient based on a "dartboard approach" (see Combes et al. 2006 for a survey of these coefficients).

The concept of sprawl can be considered in two distinct forms: the first one links high job decentralization with high population density. A metropolitan area can be composed of several employment subcenters where each subcenter is dense, the second one, the most common, reckons that high decentralization is linked with low density. In other words subcenters have a lower density than the initial center. Glaser and Kahn (2004) showed that this second version of sprawl tended to prevail.

In order to get two indices of extreme climate, the authors use the cooling degree days, which reflect demand for air conditioning, and heating degree days, which represent demand for fuel. They show that their sprawl index can be reduced by respectively 6.512 and 4.986 points by a one standard deviation increase in mean cooling days and heating days.

This factor is important since it reflects the indivisible public facilities that are present in the agglomeration. For instance, the authors report that in San Antonio, the Water System charges developers of the periphery fees that can attain $24,000. In common language these areas are on a "bad line", which means that they are far from all connection points, then the presence of ground...
water availability can permit to reduce costs by sinking and thus avoid connection to public supply. As a result they find that the sprawl index increases by 1.222 points with water ground availability

Concerning the causes of agglomeration Kim (1997) pursued his study of concentration and specialisation in cities’ data, and demonstrated that a city’s manufacturing structure reflects that of the region to which it belongs. For this author this fact indicates that concentration is ruled by regional comparative advantage rather than by increasing returns or scale economies. The importance of natural comparative advantage is also emphasized by Ellison and Glaeser (1999), who believe that this advantage can account for 50% of the geographic concentration of activities

This result was confirmed by Dumais et al (2002) on US county data over the period 1972-1992. An interesting observation made by these authors and confirmed by Barrios et al (2005) on Irish and Portuguese data was that although regional agglomeration is stable, firms’ mobility between regions is high, a possible explanation was that dead firms are replaced by firms which come from other regions.

We may wonder whether this study actually succeeds in proving that there are no multiple equilibria for two reasons: first the fact that agglomeration is permanent perhaps only means that the shock was not high enough (in order to verify this, we could estimate the shock and then simulate it). Secondly Davis and Weinstein wanted to know if agglomeration and dispersion were multiple equilibria after the second war, but in the theoretical model this happened only for a small range of \( \phi (\phi \in [\phi_a, \phi_b]) \) and perhaps the Japanese trade liberalization was not in this interval during the period analysed.

Crozet and Trionfetti (2007) pursued this work by asking the following question: "Does the HME survive and what shape does it take when trade costs in the CRS-PC good are not high enough to impede trade in this good?" (CRS-PC : constant returns to scale (CRS) and perfect competition (PC)). They found that the HME is reduced and becomes non-linear. In other words, this effect seems more significant for very large and very small countries than for medium size countries.

There is in fact an overlap of industrial transaction costs between which agglomeration and dispersion are stable, but unlike the initial model there is no break point.

See also Bernard et al. (2003), Helpman et al. (2004), Melitz and Ottaviano (2003), Yeaple (2005)

See also Candau (2007) who explicitly assumes that governments embezzle public funds. Another explanation of these equations and of the ‘no delocation condition’ was proposed by Baldwin and Krugman (2004) who assumed that public goods supply can differ from one jurisdiction to the next but that entrepreneurs do not take this difference into account when they decide to move.

According to Sen [1979]: "Problems of consistency of such judgments are quite serious, and in terms of welfare relevance they are exposed to the danger of being either unacceptable, or redundant."

More on the analytical, historical and philosophical underpinnings of this approach can be found in Fleurbaey and Hammond (2004), Fleurbaey and Mongin (2005), Fleurbaey (2007).
References


