GROWTH AND AGGLOMERATION*

BY PHILIPPE MARTIN AND GIANMARCO I. P. OTTAVIANO

CERAS-ENPC, Université de lille-1, France and CEPR,
United Kingdom and Università di Bologna, Italy and CEPR

This article presents a model in which growth and geographic agglomeration of economic activities are mutually self-reinforcing processes. Economic agglomeration in one region spurs growth because it reduces the cost of innovation in that region through a pecuniary externality due to transaction costs. Growth fosters agglomeration because, as the sector at the origin of innovation expands, new firms tend to locate close to this sector. Agglomeration implies that all innovation and most production activities take place in the core region. However, as new firms are continuously created in the core, some relocate their production to the periphery.

1. INTRODUCTION

Spatial agglomeration of economic activities on the one hand and economic growth on the other hand are parallel processes. Indeed, the emergence and dominance of spatial concentration of economic activities is one of the facts that Kuznets (1966) associated with modern economic growth. This strong positive correlation between growth and geographic agglomeration of economic activities has been documented by economic historians (Hohenberg and Lees, 1985, for example), in particular in relation to the industrial revolution in Europe during the 19th century. In this case, as the growth rate in Europe as a whole sharply increased, agglomeration materialized itself in an increase of the urbanization rate but also in the formation of industrial clusters in the core of Europe that have been by and large sustained until now. Another example is China where the sharp increase in the growth rate of the country during the past 20 years has been accompanied by an increase in the disparity between coastal and inland regions. The strong correlation between aggregate growth and spatial agglomeration therefore exists at different levels of agglomeration: the city and,
at the other end of the spectrum, the industrial clusters which can be transregional or transnational.

In this article, we construct a model in which aggregate growth and spatial agglomeration are mutually self-reinforcing processes so that the correlation between the two comes as a natural result of the economic mechanisms we introduce. Growth, through innovation, spurs spatial agglomeration of economic activities which in turn leads to a lower cost of innovation and higher growth so that a circular causation between growth and the geographic concentration of economic activities sets in. More precisely, if economic activities at the origin of innovation and growth use goods from imperfectly competitive industries as inputs, then these industries will be drawn toward the location where the innovation takes place: this gives us a "forward linkage." Because of the presence of transaction costs, this in turn lowers the cost of innovation and increases the incentive to innovate and the growth rate which corresponds to the idea of a "backward linkage".

The logic of the model borrows some features from the "new economic geography" which focuses on the presence of circular causation mechanisms to explain the spatial concentration of economic activities (Krugman, 1991a, b; Venables, 1996). In these models, the centripetal forces come from preference for variety on the consumption side (Krugman's models) or diversity in intermediate goods on the production side (Venables's model). The centrifugal forces arise from the pressure posed by geographic concentration on local factor markets that bids up prices and dispersed demand. When transport costs are low enough, a process of agglomeration sets in. In Krugman's models an increasing returns sector uses a specific input so that, in order to relax the pressure on its local input market, mobility between locations is allowed. In Venables's model there is no specificity in the input used by the increasing returns sector. The pressure can thus be alleviated without interlocational mobility. All that is required is intersectoral mobility in the presence of intrasectoral vertical linkages in the increasing returns sector.

Our article shows a third way along which agglomeration can occur. On purpose, we isolate the model from the occurrence of both Krugman's and Venables's agglomeration processes. We allow neither interlocational mobility of labor nor intrasectoral vertical linkages in the increasing returns sector. Under these circumstances, in the usual static setting of the new geography models no cumulative causation would set in. We show that this is not true anymore if growth is introduced. In order to do so, we build in the model an endogenous process of innovation and growth along the lines of Romer (1990) and Grossman and Helpman (1991). The introduction of growth generates the cumulative process that leads to geographic agglomeration. We start from an initial situation where both locations are identical and show that as long as the economy is growing, the only stable equilibrium is one in which one of the two locations gets all the innovation activity and most of the increasing returns production.

The introduction of growth modifies another result of the new geography. Because the models in this literature look at the spatial distribution of a fixed stock of resources, agglomeration implies that production moves only from the periphery to the core (or from the South to the North). The interpretation is that the static analysis represents the long run steady state of some growth model in which the long run
growth rate does not depend on the geography of economic activities. In our model, in which the two are allowed to interact, because new firms are continuously created in the core, the relocation dynamics are richer and more realistic. Agglomeration takes place in that all innovation activities and most production activities are located in the core, but some firms created in the core relocate to produce in the periphery. This is also consistent with the empirical evidence of Audretsch and Feldman (1996) who show that the geography of innovation and the geography of production tend to mirror each other but not perfectly and that innovation activities are more spatially concentrated than production activities.

We are able not only to identify a new determinant of agglomeration, but also to give to the spatial distribution of economic activities an important role in explaining long term growth. As in Krugman's models, agglomeration generates spatial divergence in income levels. In addition, we show that the more spatially agglomerated an economy is, the faster it grows in aggregate terms. Geography matters for growth.

Our approach is related to a few existing contributions that try to capture the spatial dimension of economic development by merging “new economic geography” and endogenous growth models. Among such contributions, Englmann and Walz (1995) and Walz (1996), building on the centripetal forces highlighted by Krugman (1991a,b) and Venables (1996), show that, with preference for diversity in imperfectly tradable intermediates and interregional factor mobility, linkages between intermediate and final good producers can create a tendency for production and innovation activities to cluster in the region with an initial advantage in the number of intermediates. Martin and Ottaviano (1999) as well as Baldwin and Forslid (1997) show that the same can happen when consumers love variety of imperfectly tradable final goods. However, these articles share with previous works on endogenous growth the property that the location of innovation matters for growth only in the case of local technological spillovers (see, e.g., Grossman and Helpman, 1991; Rivera-Batiz and Romer, 1991; Bertola 1993). On the contrary, we show that in our framework the geography of economic activities matters for growth even in the absence of local technological spillovers. The mechanism that relates growth to geography only goes through market interactions. Therefore, the article highlights also a novel channel, the geography one, through which trade can affect growth.

Section 2 of the article presents the general framework of the model. In Section 3, first, we derive the equilibrium relation that illustrates the demand linkage at work in the model and makes the geography of economic activities a function of the growth rate. Then, we derive the cost linkage that makes the growth rate a function of the geography. Finally, we put these two relations together to show that a situation where both regions are identical is not the only steady state of the model but two additional steady states exist where all the innovation activity and (most of) production activities are agglomerated in one region. In section 4 we assess the stability properties of the steady states. We argue that, whenever growth is positive, the steady state with identical regions is unstable. When such a steady state is perturbed, the two additional steady states described above are the only ones consistent with rational expectations. Section 5 shows how changes in parameters, especially transaction costs, modify the growth/geography joint equilibrium. Section 6 concludes.
2. A TWO-LOCATION MODEL

The model merges a location framework which is a variant of Krugman (1991b) and Venables (1996) and a growth framework similar to Romer (1990) and Grossman and Helpman (1991). There are two regions, called region 1 and region 2, which are initially identical. Variables referring to region 2 are labeled by *. Each region is endowed with a fixed amount of labor $L$ which we assume to be immobile between regions so as to abstract from that particular agglomeration channel. Labor can be used to produce a homogenous good and differentiated goods which are aggregated into a composite good. All goods are final consumption goods. Moreover, the composite good can be used as intermediate input in the innovation sector to create new varieties of the composite good itself. For the production of a variety to be possible, its blueprint has to be invented first. The blueprint is then protected by an infinitely lived patent whose initial property belongs to the region where invention has taken place. After registration, the patent can be sold to any producer located in either region. The innovation and the production processes are therefore conducted by different economic agents and possibly in different regions. We assume that regions own initially equal stocks of patents.

Since the model is symmetric, we concentrate on the specification of region 1. Preferences are instantaneously nested-CES and intertemporally constant elasticity of substitution (CES) with unit elasticity of intertemporal substitution,

$$U = \int_0^\infty \log[D(t)^{\sigma} Y(t)^{1-\sigma}] e^{-\rho t} dt$$

where $Y$ is the consumption of the homogeneous good, $\rho$ is the rate of time preference, and $\alpha \in (0, 1)$ is the share of expenditures devoted to $D$, a composite good which, following Dixit and Stiglitz (1977), consists of a number of different varieties,

$$D(t) = \left[ \sum_{i=1}^{N(t)} D_i(t)^{1-1/\sigma} \right]^{1/(1-1/\sigma)} \quad \sigma > 1$$

where $N$ is the total number of varieties available in the economy. Growth will come from an increase in the variety of goods measured by $N$. $\sigma$ is the elasticity of substitution between varieties as well as the own-price elasticity of demand for each variety.

The value of expenditure $E$ is

$$E = \sum_{i=1}^{n} p_i D_i + \sum_{j=n+1}^{N} \tau p_j D_j + p_Y Y$$

where $p_Y$ is the price of good $Y$, $p_i$ is the price of the $i$th variety, and $n$ is the number of varieties produced in region 1 and $N = n + n^*$. As in Samuelson (1954) and in the economic geography literature, transaction costs in the form of iceberg costs have been introduced. $\tau$ is more than 1 so that only a fraction of the good purchased in the other region is actually consumed. As is common in new geography models, there is no transaction cost for the homogenous good which is introduced to tie down the wage rate.
As to the supply side, the homogenous good is produced using only labor with constant returns to scale in a perfectly competitive sector. Without loss of generality, the input requirement is set to 1 for convenience. Moreover it is assumed that the demand of this good in the whole economy is large enough that it cannot be satisfied by production in one region only. This hypothesis ensures that in equilibrium the homogenous good will be produced in both regions. Hence, because of free trade, the wage rates in the two regions will be equalized to, say, \( w \). In addition, the assumption of unit input requirement in the production of \( Y \) entails \( p_Y = w \) everywhere.

The differentiated goods are produced in a monopolistically competitive sector. The production of each variety exhibits increasing returns to scale. Together with the assumption of costless differentiation this ensures that each firm will produce only its own variety. More precisely, the supply of one unit of each variety requires the use of one patent (the fixed cost at the source of economies of scale) and \( \beta \) units of labor. Under these assumptions, optimal pricing for any variety gives producer prices \( p = p^* = w\beta \sigma/(\sigma - 1) \). The operating profits of a producer using a patent are revenues minus the labor costs,

\[
\pi = px - w\beta x = \frac{w\beta x}{\sigma - 1}
\]

where \( x \) is the optimal output/size of a typical firm in equilibrium.

The invention of new varieties, which is at the source of growth in the model, is performed in the innovation sector. This sector produces ideas that it can patent and then sells these patents in both regions to producers who need a patent to start the production of differentiated goods. Alternatively, we can interpret the model as one where research is performed by firms themselves which then use the invention they have developed to start production. They are free to locate the production process in either region and then repatriate the profits. The value of the patent is in this interpretation the value of the firm and a stock market replaces the market for patents.

No transaction cost hinders the trade in patents or the relocation of firms and the innovation sector is perfectly competitive. The process of innovation requires a composite CES good, made of different varieties of goods, as the only input with an input requirement \( \eta N^\mu \). The second term in this expression implies that the cost of innovation depends on the number of past innovations, \( N \), so that a learning curve exists. We will describe the necessary condition on the form of this learning curve in Section 3.2. The use of intermediates in innovation is closely related to Rivera-Batiz and Romer (1991) and Baldwin and Forsslid (2000). Note that the innovation sector does not employ labor directly but it does indirectly as it uses specialized goods which require labor as the sole primary factor of production. We simplify further by assuming that the same composite good \( D \) is used both for consumption and as an input for innovation and that the same transaction costs hinder trade in goods whether these are consumed by the consumers or used by the innovation sector. This is done for analytical convenience as suggested by Krugman and Venables (1995). The

---

3 This way of introducing economies of scale has been used in Flam and Helpman (1987) in a trade context and Martin and Rogers (1995) in a geography context.
cost for developing a new variety in region 1 is therefore \( F = w\left[ \beta \sigma / (\sigma - 1) \right] \left[ nz_i + n^* z_j \right] \), where \( z_i \) and \( z_j \) are the demands for the differentiated goods produced in region 1 and region 2, respectively. The problem of an innovator is therefore to minimize this cost subject to the input requirement constraint:

\[
\eta N^\mu = \left[ nz_i^{1-1/\sigma} + n^* z_j^{1-1/\sigma} \right]^{1/(1-\sigma)}
\]

Solving for the demands of each variety by a single innovator in region 1, we find

\[
z_j = \eta N^\mu (n + n^* \delta)^{\sigma/(1-\sigma)}
\]

where \( \delta \equiv \tau^{1-\sigma} \in (0, 1) \). The equilibrium cost of innovation is then

\[
F = \frac{w\beta \sigma}{\sigma - 1} \eta N^\mu N^{1/(1-\sigma)} [\gamma + (1 - \gamma) \delta]^{1/(1-\sigma)}
\]

where \( \gamma = n/N \) is the share of varieties produced in region 1 and is less or equal to 1. \( \gamma \) will be a crucial parameter of the model as it will measure the extent of agglomeration of the differentiated good sector in region 1. Equation (7a) builds a pecuniary externality into the model. The cost of innovation is lower in the region where there are more local firms since, in the presence of transaction costs, the CES aggregate used for innovation costs less. This is an example of what is sometimes called a “vertical linkage” in the new economic geography literature (e.g., Venables, 1996). Moreover, for a given value of \( \mu \), which measures the extent of technological spillovers necessary for a constant growth rate (see Section 3.2), the cost of innovation decreases as new differentiated goods are invented (i.e., as \( N \) increases) because these differentiated goods are used as specialized inputs in the innovation sector.

The equilibrium cost of innovation in region 2 is symmetric:

\[
F^* = \frac{w\beta \sigma}{\sigma - 1} \eta N^\mu N^{1/(1-\sigma)} [1 - \gamma + \gamma \delta]^{1/(1-\sigma)}
\]

As the patents produced by the innovation sector, which is perfectly competitive, are costlessly tradable between the two regions, the price and the cost of innovation have to be the same in equilibrium for both regions to engage in innovation. This immediately implies that innovation is split between the two regions only if \( \gamma = 1/2 \), i.e., if the manufacturing sector is also perfectly split. If \( F < F^* \), then all innovation is conducted in region 1 and vice versa.

Finally, we assume that a safe asset exists that bears an interest rate \( r \) in units of the numeraire.\(^4\) Its market is characterized by free financial movements between the two regions. The intertemporal optimization by consumers then implies that the growth rate of expenditures is equal to the difference between the interest rate and the rate of time preference: \( \hat{E} = \hat{E}^* = r - \rho \).

A steady state of the model is defined as an equilibrium where \( \gamma \), the proportion of firms in region 1, is constant and the number of varieties grows at a constant rate

\(^4\) As in Grossman and Helpman (1991), the exact nature of this asset is immaterial since it is introduced only to clarify the intertemporal arbitrage between consumption and investment in innovation.
g = \bar{N}/N. It will turn out that in steady state E and E* are constant so that r = \rho. Moreover, incomes are determined by the value of \gamma, implying that the steady state of the model will be fully characterized by the equilibrium location parameter \gamma and the growth rate of innovation g. Due to its perfect symmetry, we know that the model will have at least a symmetric steady state in which both regions engage in innovation activities and the production of the two goods is evenly split between them (\gamma = 1/2). Section 3 will show that, if the economy is growing, this is not the only steady state of the model because two other equilibria exist in which a single region engages in innovation activity and is (partially) specialized in the supply of the differentiated goods (\gamma < 1/2 and \gamma > 1/2). Under the same circumstances, Section 4 will show that the symmetric equilibrium is unstable and, therefore, it will not be generally observed in equilibrium.

3. MULTIPLE STEADY STATES

In this section we solve for the steady states of the two-region model. We show their multiplicity and discuss how such multiplicity originates from a process of circular causation between growth and agglomeration.

3.1. Economic Geography as a Function of Growth. We begin by characterizing the equilibrium condition for the location of firms. We consider the two possible configurations: one where the innovation activities are split between the two regions and one where they take place in a single region, say in region 1. By symmetry, mutatis mutandis, all the findings would carry through if region 2 were the only one to perform innovation activities. Several equilibrium conditions characterize the steady state location and growth rates. The first is a market clearing condition for the manufacturing sector: if innovation is concentrated in region 1, this condition implies that the supply of each variety equals demand (inclusive of transaction costs) from consumers in both regions as well as demand from innovation activities in region 1:

\begin{align}
(8a) \quad x &= \frac{\alpha L (\sigma - 1)}{w\beta \sigma} \left( \frac{E}{N[\gamma + 1(1 - \gamma)\delta]} + \frac{E^* \delta}{N[\delta \gamma + 1 - \gamma]} \right) \\
&\quad + \frac{\sigma - 1}{w\beta \sigma} \frac{F\bar{N}}{N[\gamma + (1 - \gamma)\delta]}
\end{align}

\begin{align}
(8b) \quad x^* &= \frac{\alpha L (\sigma - 1)}{w\beta \sigma} \left( \frac{E\delta}{N[\gamma + 1(1 - \gamma)\delta]} + \frac{E^*}{N[\delta \gamma + 1 - \gamma]} \right) \\
&\quad + \frac{\sigma - 1}{w\beta \sigma} \frac{F\bar{N}\delta}{N[\gamma + (1 - \gamma)\delta]}
\end{align}

In the above equations, on the right hand side the first term gives the usual demands by consumers derived from utility maximization. The second term is the demand from the innovation activities given in Equation (6) multiplied by the number of new inventions per unit of time (\bar{N}).
If both regions engage in innovation in equilibrium, so that in Equations (8a) and (8b) demands from the innovation sector in region 2 are added, the supply-equal-demand conditions are replaced by

\[
\begin{align*}
(8c) & \quad x = \frac{\alpha L(\sigma - 1)}{w \beta \sigma} \left( \frac{E}{N[\gamma(1 - \gamma)\delta]} + \frac{E^* \delta}{N[(\delta + 1) - \gamma]} \right) \\
& \quad + \frac{\sigma - 1}{2w \beta \sigma} \left[ \frac{FN \delta}{N[\gamma(1 - \gamma)\delta]} + \frac{FN \delta}{N[1 - \gamma + \gamma \delta]} \right]
\end{align*}
\]

\[
(8d) & \quad x^* = \frac{\alpha L(\sigma - 1)}{w \beta \sigma} \left( \frac{E \delta}{N[\gamma(1 - \gamma)\delta]} + \frac{E^*}{N[\delta + (1 - \gamma)]} \right) \\
& \quad + \frac{\sigma - 1}{2w \beta \sigma} \left[ \frac{FN \delta}{N[\gamma(1 - \gamma)\delta]} + \frac{FN \delta}{N[(1 - \gamma) + \gamma \delta]} \right]
\]

In steady state the operating profits and therefore the optimal size of the firms must be the same across regions, so that \( x = x^* \). This equilibrium condition which insures that firms have no incentive to relocate in equilibrium implies that we can solve (8a) and (8b) for \( y \) in the case where innovation activities are concentrated in region 1,

\[
(9a) \quad \gamma = \frac{\alpha L(E + E^*)[(1 + \delta)e - \delta] + gNF}{(1 - \delta)[\alpha L(E + E^*) + gNF]} \quad \text{if} \quad \frac{\alpha L(E + E^*)}{2gNF} > \frac{\delta}{1 - \delta}
\]

and \( \gamma = 1 \) otherwise. \( e \) is region 1 share of total expenditures. Differently, if innovation is split between the two regions, then the location of firms is given by

\[
(9b) \quad \gamma = \frac{\alpha L(E + E^*)[(1 + \delta)e - \delta] + (1 - \delta)gN}{(1 - \delta)[\alpha L(E + E^*) + 2gNF]}
\]

It implies that \( \gamma = 1/2 \) when \( e = 1/2 \) since the two regions are identical in this case. In Equations (9a) and (9b) \( g = \dot{N}/N \) is the growth rate of the economy. Because expenditures grow at the same rate in both region (\( \dot{E} = \dot{E}^* = \gamma - \rho \)) the share of total expenditures in region 1, \( e \), is a constant which is completely determined by the initial geographical distribution of the property rights on patents. Since regions are initially identical, \( e = 1/2 \) which we impose from now on. Equation (9a) illustrates the forward linkage at work in our model implying that geographic agglomeration increases with growth, which manifests itself by an increased activity in the innovation sector. It has the usual interpretation that firms in the increasing returns sector will tend to locate in the region which has the highest expenditure level. Here, the novelty comes from the fact that expenditures on the differentiated goods come not only from consumers but also from the sector at the origin of growth. It can already be seen that, in the case where innovation is agglomerated in one region, a higher growth rate implies a higher demand for the differentiated goods which gives an incentive for firms to move to that region: \( \gamma \) increases with \( g \). In particular, if the intermediate demand from innovation is large enough with respect to the final demand by consumers, firms will be all concentrated in that region.
Whatever the location of innovation activities and of the manufacturing sector, the size of firms is

\[ x = \frac{\sigma - 1}{\omega \beta \sigma} \left[ \alpha L E + E^* \frac{E^*}{N} + Fg \right] \]

(10)

We have found two possible equilibrium locations: one is given by Equation (9a) in the case when region 1 is the only one to have innovation activities; the other one is given by Equation (9b) which tells us that, if innovation activities are perfectly split, then so too will be the production activities.

3.2. Growth as a Function of Economic Geography. We turn now to the intertemporal equilibrium. Calling \( \nu \) the value of a patent, the condition of no arbitrage opportunity between patents and the safe asset implies

\[ r = \frac{\dot{\nu}}{\nu} + \frac{\pi}{\nu} \]

(11)

On an investment on a patent of value \( \nu \), the return is equal to the operating profits plus the change in the value of the patent. This condition can also be derived by stating that the equilibrium value of a patent is the discounted sum of future profits of the firm which buys the patent and has a monopoly forever on the production of the related variety. Because of marginal cost pricing of patents, free entry, and zero profits in the innovation sector, \( \nu = F \) is another equilibrium condition.

For a steady growth path to exist, it must be that the combination of the pecuniary externality and the technological spillovers is such that the cost of innovation decreases at the same rate at which operating profits of firms which buy the patents decrease. This ensures that the incentive to engage in innovation remains constant over time. The requirement that \( F \) is decreasing at rate \( g \) is consistent with \( x \), the size of firms, and therefore profits decreasing at the same rate as with expenditures being constant (see Equations (9) and (10)). The parameter \( \mu \) measures the technological spillovers that are required for the economy to have ongoing growth and well-defined private investment decisions in endogenous growth models.\(^6\) It is picked such that a constant growth rate for this economy exists: \( \mu = (\sigma - 2)/(1 - \sigma) \); i.e., because of technological spillovers, the cost of innovation depends on past innovations.\(^7\) If \( \sigma \) is more than 2, there are positive technological spillovers so that past innovations decrease the cost of innovation. In this case the pecuniary externality will be relatively weak as the differentiated goods are relatively good substitutes. On the contrary if \( \sigma \) is between 1 and 2, there are negative technological spillovers (it becomes harder to find new innovations maybe because there is a tendency to run out of new ideas) but the pecuniary externality effect is then strong enough to compensate the negative

---

\(^5\) It can be easily checked that this is also true for \( \gamma = 1 \), i.e., full agglomeration in location 1.

\(^6\) As shown in Lucas (1988), endogenous growth models generate constant steady state growth rates only under knife-edge assumptions on parameters. Our assumption on \( \mu \) is just an example of this. See Baldwin and Forslid (2000) and Evans et al. (1998) for a similar assumption.

\(^7\) Martin and Ottaviano (1999) draw a sharp distinction between global and local spillovers. Here we do not focus on this issue and only look at the case of global spillovers. The introduction of local spillovers in the innovation sector would not change the qualitative nature of the equilibrium.
technological spillovers. There is a specific case where intertemporal technological spillovers are not required to generate ongoing growth. All that is needed is that the above mentioned pecuniary externality be strong enough. The model will exhibit such a property for \(\sigma = 2\); i.e., the cost saving for innovation due to a new variety is just enough to offset the decrease in revenues due to more competition in the differentiated good market.

If the restriction \(\mu \equiv (\sigma - 2)/(1 - \sigma)\) does not hold, the model will exhibit growth rates that are locally increasing or decreasing over time. The model may still be well behaved but cannot be solved for a constant growth rate. The restriction on \(\mu\) therefore greatly simplifies the solution of the model. It implies that the cost of innovation is

\[
F = w\frac{\beta \sigma}{\sigma - 1} \eta N^{-1} [\gamma + (1 - \gamma) \delta]^{1/(\sigma - 1)}
\]

As \(\gamma\) and \(w\) are constant in steady state, this ensures that \(FN\) is also constant, so that \(\dot{v}/v = \dot{F}/F = -N/N = -g\). Consumers’ expenditures are constant in steady state (see Equation (9a)); the interest rate \(r\) is equal to the rate of time preference \(\rho\). Using these results as well as (4) and (10) in (11), we find

\[
g = \frac{\alpha}{\sigma - 1} \frac{L(E + E^*)}{FN} - \rho \frac{\sigma}{\sigma - 1}
\]

Consider now the market clearing condition in the labor market which implies that labor supply will be employed either in the constant returns sector or in the increasing returns sector,

\[
2L = \frac{(1 - \alpha) L(E + E^*)}{w} + \beta Nx
\]

where we have used \(\rho \gamma = w\). Equation (14a) can be transformed when substituting for \(x\) in Equation (10) into

\[
E + E^* = 2 \frac{w \sigma}{\sigma - \alpha} - \frac{\sigma - 1}{\sigma - \alpha} gNF
\]

When innovation is concentrated in region 1, we can combine (12), (13), and (14b) to find the equilibrium growth rate that can be expressed as a function of \(\gamma\) only:

\[
g = \frac{2\alpha L}{\eta \beta \sigma} [\gamma + \delta (1 - \gamma)]^{1/(\sigma - 1)} - \rho \frac{\sigma - \alpha}{\sigma - 1}
\]

This equation illustrates the backward linkage at work in the model when innovation is entirely located in region 1. An increase in the concentration of economic activities in that region decreases the cost of innovation (because of the existence of transaction costs between the two regions), pushing new researchers to enter the innovation sector until profits in that sector are back to zero. This in turn increases the rate of innovation. In the symmetric equilibrium, growth is still given by Equation (15) where \(\gamma\) takes the value 1/2.
To complete the solution of the model we have to determine total consumers’ expenditures. Equations (13) and (14b) together yield

\[(16) \quad E + E^* = 2w + \rho \frac{FN}{L}\]

The first term on the right hand side of (16) is wage income. The second is the value of the initial stock of patents, which appear in (16) because only the profits accruing to the initial stock of patents are pure rents.\(^8\) Equation (16) can also be derived by noting that \(\rho\) is the propensity to consume out of wealth which is \(FN/L\).

If region 1 gets all innovation activities, its total nominal GDP defined in the usual way as the sum of value added in the three sectors is GDP\(_1\) = \(wL + \gamma w\beta xN/(\sigma - 1)\). This is higher than nominal GDP in region 2, GDP\(_2\) = \(wL + (1 - \gamma)w\beta xN/(\sigma - 1)\), as long as \(\gamma\) is more than 1/2 which will be the case. This is because region 1 produces more differentiated goods and less constant returns to scale goods than region 2.

Expression (16) is well defined up to the choice of a numeraire good. Since the economy lacks a monetary instrument, we are free to set the time path for one nominal variable and to measure prices at every instant against the chosen numeraire. Of course, while the choice of numeraire has no effect on the evolution of real variables, it will be crucial in order to simplify the dynamic analysis. That is why we follow Grossman and Helpman (1991) and, from now on, we normalize prices at every moment so that nominal spending remains constant; i.e., \(E + E^* = 1\).\(^9\)

3.3. Multiple Steady States. We are now ready to fully describe the different equilibria of the model. The situation where both regions remain identical is a steady state equilibrium, \(\gamma = 1/2\) in this case, and the cost of innovation is the same in both regions so that if the economy starts in that equilibrium, there is no incentive to relocate innovation activities. Nor is there any incentive to relocate the production of the increasing returns sector. This is because demands for the differentiated goods from the consumers and from the innovation sectors are the same in both regions so that profits are also identical. When both regions remain identical, the growth rate is given by Equation (15) where \(\gamma\) takes the value 1/2.

However, this is not the only equilibrium. If one region (say, region 1) gets more of the firms producing differentiated goods then the cost of inputs for innovation in that region will be lower due to the presence of transaction costs between the two regions (see Equation (12)). As the innovation sector is perfectly competitive and the good it produces, namely patents, freely tradable between the two regions, then this region will get all the innovation activity and the other region will cease any innovation activity. In this case, we know that the location equilibrium is no more \(\gamma = 1/2\) but it is given by (9a) with \(\gamma > 1/2\). Using (12) and (16), we can rewrite the

\(^8\) Remember that \(FN\) is constant so that the income is fixed by the initial ownership of patents. Total income can also be defined as wage income, plus operating profits net of investment in the innovation sector: \(E = w + (N/2 - NF/2)L\).

\(^9\) Though convenient, this choice of numeraire is a bit unconventional (see, however, Walz (1996) for the same assumption) and sometimes obscures the solution of the model. This is why we have postponed it so far.
equilibrium location Equation (9a) as

\[ g = \frac{(1 - \delta)(2 \gamma - 1)\alpha}{1 - (1 - \delta)\gamma} \left\{ \frac{\sigma - 1}{\eta \beta \sigma} \left[ \gamma + (1 - \gamma)\delta \right]^{\gamma/\sigma - 1} + \frac{p}{2L} \right\} \]

if $\gamma \in (1/2, 1)$ and $\gamma = 1$ otherwise.

Values for $g$ and $\gamma \in (1/2, 1]$ that satisfy both (15) and (17) are steady state solutions of the model. Equation (15) describes how the incentive to engage in innovation and therefore the growth rate $g$ depends on the location of industries. Equation (17) gives the relation between the location of firms and the growth rate. Because of its nonlinearity in $\gamma$, it is written with the growth rate as a function of the location of firms. As it can be checked analytically, while both curves have $g$ as an increasing function of $\gamma$ if $\gamma \in (1/2, 1]$, Equation (17) is always steeper than Equation (15). Therefore, if they cross for $g > 0$ and $\gamma \in (1/2, 1]$, they do so at most once.

To assess the possible configurations of steady states, it is useful to refer to Figure 1 which plots (15) and (17) in the relevant set $\gamma \in (1/2, 1]$ and $g > 0$. The steeper curve represents (17) which becomes vertical at $\gamma = 1$. The flatter curves show three alternative positions of (15) which imply three alternative steady state configurations.

In the first configuration, (15) lies everywhere below (17). In this case—corresponding to (15)' in the figure—growth cannot be positive for any $\gamma$. Thus, the only steady state of the model entails $(\gamma, g) = (1/2, 0)$ as implied by Equations (9a) and (9b) for $g = 0$. In the second configuration, corresponding to (15)'' and (17) and (15) and (17) cross for $\gamma \in (1/2, 1)$: growth is always positive for any $\gamma$ and the model exhibits three steady states, one at $\gamma = 1/2$ and two at $\gamma \in (0, 1/2)$ and $\gamma \in (1/2, 1)$, respectively (partial agglomeration). Due to the symmetry of the model, the last two steady states are symmetric around $\gamma = 1/2$ with the latter shown by point A in the figure. In the final configuration, (15) and (17) cross at $\gamma = 1$ as shown by (15)". In this case, growth is also positive and the model features three steady states, $\gamma = 1/2$, $\gamma = 0$, and $\gamma = 1$ (complete agglomeration).
The occurrence of different configurations depends on the value of parameters. In particular, it can be checked that the first configuration occurs if $\delta \leq \delta_1$, the second if $\delta_1 < \delta < \delta_2$, and the third if $\delta \geq \delta_2$ with the threshold values $\delta_1$ and $\delta_2$ defined as

\begin{equation}
\delta_1 \equiv 2 \left[ \frac{\rho \eta \beta \sigma (\sigma - \alpha)}{2\alpha L(\sigma - 1)} \right]^{\sigma - 1} - 1
\end{equation}

and

\begin{equation}
\delta_2 \equiv \frac{\alpha (\sigma - 1)[2L(\sigma - 1) + \rho \eta \beta \sigma]}{2\alpha L(\sigma^2 - 1) + \rho \eta \beta \sigma [\sigma (1 - \alpha) + \sigma - \alpha]}
\end{equation}

We can also note that, for any level of transaction cost $\tau$, a sufficient condition for the existence of a unique steady state in which regions stay identical and growth is zero is

\begin{equation}
2\alpha L(\sigma - 1) - \rho \eta \beta \sigma (\sigma - \alpha) \leq 0
\end{equation}

because for these parameters, growth as given by (15) can never be positive. This is a standard result (see Grossman and Helpman, 1991). It states that growth is zero if the incentive to innovate is insufficient, which is the case when the labor endowment $(2L)$ as well as the expenditure share of the differentiated good sector ($\alpha$) are small, and if the marginal cost of innovation ($\eta$), the rate of time preference ($\rho$), as well as the elasticity of substitution ($\sigma$) are large.

4. UNSTABLE AND STABLE STEADY STATES

The objective of this section is to show how one region can emerge as the main center for both production and innovation activities even when regions are initially identical and how the emergence of what can be interpreted as an industrial cluster can spur growth in the economy. We therefore want to show that the steady state where both regions remain identical forever is not stable when growth is positive.

The stability properties of the equilibria can be assessed following Grossman and Helpman (1991). In particular, we will argue that (i) the symmetric steady state associated with $\gamma = 1/2$ is unstable whenever the equilibrium growth rate is positive, and (ii) apart from a situation in which the economy starts at $\gamma = 1/2$ and stays there forever, the only paths consistent with rational expectations entail an instantaneous jump to one of the steady states associated with $\gamma \in [0, 1/2)$ or $\gamma \in (1/2, 1]$.

First, the symmetric steady state is not stable because it cannot be approached along any trajectory. As already noted when commenting about (7b), because wages are the same in both regions, only the region that starts with the larger number of firms will ever be active in innovation. Therefore, unless the regions happen to start with the same numbers of firms, an equal wage trajectory must lead to a steady state in which innovation is agglomerated in one region and such a region is also relatively specialized in the production of differentiated good.\(^\text{10}\)

\(^{10}\)A thorough presentation of this line of reasoning can be found in Grossman and Helpman (1991; Chap. 8, Appendix). In a model which exhibits the same qualitative dynamics as the present
Second, we need to show that, when the economy does not start at the symmetric steady state, there are only two alternative equilibria that entail the agglomeration of the innovation activities in one region and the concentration of the differentiated good sector in that same region.

We start by recalling that, with free capital movements, the intertemporal utility maximization requires \( \hat{E} = \hat{E}^* = r - \rho \), which means that expenditures grow at the same rate in the two regions. This implies that total expenditures \( E + E^* \) also grow at the same rate; that is, \( d \ln(E + E^*)/dt = r - \rho \). Thus, our choice of numeraire, which entails that expenditures are constant at \( E + E^* = 1 \), gives \( d \ln(E + E^*)/dt = 0 \) and \( r = \rho \).

Then, we define a new variable \( V \) as the inverse of the nominal value of the existing stock of patents, \( V = 1/\nu N = 1/FN \), so that \( \dot{V} = -\dot{\nu} - g \). This fact, together with \( E + E^* = 1 \) and \( r = \rho \), allows us to use Equations (4) and (10) in order to rewrite the no-arbitrage condition (11) as

\[
\dot{V} = \frac{\alpha}{\sigma} V - \frac{\sigma - 1}{\sigma} g - \rho
\]

Finally, we turn to the resource constraint (14b). Due to the definition of \( V \) and the choice of numeraire, we can use (9a) and (12) to rewrite (14b) as

\[
2L(\sigma - 1) \frac{1 + \delta}{\beta \eta} \left( \frac{1}{2} \right)^{1/(\sigma - 1)} \left( \frac{aV + 2g}{aV + g} \right)^{1/(\sigma - 1)} = (\sigma - \alpha) V + (\sigma - 1)g \quad \text{if } V > \frac{2\delta g}{[\alpha(1 - \delta)]} \\
2L(\sigma - 1) \frac{1}{\beta \eta} = (\sigma - \sigma V) + (\sigma - 1)g \quad \text{if } V \leq \frac{2\delta g}{[\alpha(1 - \delta)]}
\]

which, given (9a), is associated with \( \gamma \in (1/2, 1) \) as long as \( V > \frac{2\delta g}{[\alpha(1 - \delta)]} \) and \( \gamma = 1 \) otherwise. In particular, along (22) \( \gamma \) grows from 1/2 to 1 as we move from the intersection with the vertical axis to the one with the ray \( V = \frac{2\delta g}{[\alpha(1 - \delta)]} \). As expected, growth and agglomeration are positively correlated.

We have thus reduced the system of equilibrium relationships to a single differential equation (21) plus a side condition (22). We proceed now with a diagrammatic analysis of equilibrium dynamics by means of Figure 2. The figure depicts three curves. The downward sloping curve is (22). This equation must be satisfied at every moment in time and represents a constraint on the use of resources and their spatial allocation. The higher is the rate of growth \( g \), the larger is labor employed in innovation and, conversely, the smaller is labor devoted to production. Lower output entails higher prices and wages as well as higher values of patents and, thus, a smaller inverse value of the patents stock \( V \). The solid upward sloping curve depicts, one but allows for wages to differ between regions, they show that the symmetric steady state cannot be approached not only along equal-wage trajectories but also along unequal-wage trajectories. All such paths would involve a violation of the no-arbitrage condition due to the bang-bang behavior of innovation as the system moves from a point in the neighborhood of the symmetric steady state to the steady state itself.
instead the locus of combinations \((g, V)\) such that the economy is in steady state, i.e., \(\dot{V} = \dot{\nu} - g = 0\),

\[
V = \frac{\sigma - 1}{\alpha} g + \frac{\sigma}{\alpha} \rho
\]

Along (23) the rate of decline in the value of a typical patent exactly matches the rate of innovation. The positive slope of (23) can be understood by considering the fact that, as can be readily verified by substituting (4) and (10) into (11), the smaller the value of the aggregate stock of patents \(1/V\) the faster the rate of decline of the value of a typical patent \(\dot{\nu}\). Below (23) the number of varieties grows more rapidly than the value of the typical firm falls so that \(V\) falls. Above (23) the opposite is true. These properties explain the arrows along curve (22) in Figure 2. Finally, the dotted upward sloping curve represents \(V = 2\delta g / [\alpha(1 - \delta)]\). Points on or below such a curve are associated with \(\gamma = 1\).

Figure 2 is drawn for values of the parameters such that \(\delta_1 < \delta < \delta_2\) and, thus, (22) and (23) cross at the point A where \(g > 0\) and \(\gamma \in (1/2, 1)\). The intersection point between (22) and (23) is one where the dynamic forces in the economy effect no change in \(g\) and \(V\). In other words, were the economy to reach this point, innovation would continue at a fixed rate and the division of resources among sectors would remain constant through time as would the geographical distribution of firms. As argued by Grossman and Helpman (1991: Chap. 3), expectations can be fulfilled only if the economy jumps immediately to this steady state. If initial beliefs entailed a point \((g, V)\) different from A, they would remained unfulfilled. In the long run, any initial expectation about the final equilibrium being different than A would drive the economy toward either \(g = 0\) or \(V = 0\). In the first case, \(N\) would eventually stop at
a finite positive value while $V$ would grow without bounds. With $N$ finite, $V$ could grow without bounds only if the value of the typical variety were to approach zero and this could happen if profits were to approach zero too. But profits would never approach zero for finite $N$ and, thus, forecasting $g = 0$ is inconsistent with rational expectations. In the second case, the rate of innovation $g$ would reach its maximum value. New products would be invented continuously only if $v$ remained below the present discounted value of future profits. But this could happen only if $V$ maintained a strictly positive value so that the forecast $V = 0$ would also be inconsistent with rational expectations.\footnote{As in the first case, the argument is based on the assumption that assets are priced at their fundamental level ("no bubbles"); i.e., the value of a patent equals the present discounted value of future profits. On this, see Grossman and Helpman (1991; p. 61).}

We are now ready to merge the stability and steady state analyses. Building on Figure 1, Figure 3 depicts (15) and (17) for $\delta_1 < \delta < \delta_2$. Figure 3 depicts an agglomeration cum growth story.

From an initial situation of the world where both regions are identical, even a small disturbance will lead to both higher growth and higher agglomeration. The reason is that this disturbance increases the share of expenditures in region 1, inducing some firms to move from region 2 to region 1. This triggers all innovation activities to move to region 1 because the cost of innovation is now lower there. This itself induces more firms to move to region 1 and, as this lowers the cost of innovation, the activity of that sector expands, attracting more firms up to the point where there are no profits in the innovation sector and the profits in the manufacturing sector are equalized across regions. As shown by the stability analysis, this process of cumulative causation will be instantaneous: there is no cost to the mobility of firms, so that $\gamma$, the proportion of firms in the North, can move instantaneously and the model has no transitionary dynamics. We can therefore interpret our agglomeration...
mechanism as one of sudden emergence of industrial clusters as centers of innovation and production. Silicon Valley also seems to us a good illustration of our model where agglomeration and technological progress go hand in hand. More generally, it is consistent with experiences of rapid increases in both growth rates and spatial agglomeration.

An important characteristic of the equilibrium is that, even though agglomeration takes place in region 1, the steady state equilibrium describes a process of constant relocation of some of the production facilities toward region 2. This is because new economic activities are continuously created in region 1 and none in region 2. In general, when $\gamma$ is less than 1, some firms, $(1 - \gamma)\tilde{N}$, created in region 1 will decide to go and produce in region 2 where no new competitors are generated locally. Hence, contrary to static geography models where economic activities leave the periphery and locate in the core, our model describes a situation where, even though the core has more economic activities, the relocation dynamics of production activities are from the core to the periphery. This seems to us a more realistic description of reality. This result comes from the inclusion of endogenous capital accumulation in a geography model as well as free capital movements. The absence of capital movements in new geography models is what brings them to produce an extreme core-periphery result. In our model, free mobility of capital (patents or firms themselves) is a stabilizing force as is well known from growth models.

It can also be noted that, while the geographies of production and innovation tend to mirror each other (more firms produce where all innovation takes place), resemblance is not always perfect (production activities are not as fully concentrated as innovation activities as long as $\delta < \delta_z$) which is consistent with empirical evidence (see Audretsch and Feldman, 1996). This comes from the tension between centripetal and centrifugal forces in our model. It should be clear by now that the centripetal forces rely on the vertical linkages between production and innovation: the forward linkage is the demand linkage; the backward linkage is the cost linkage. There is actually another centripetal force which builds on the relation between growth and geography. As growth and industrial concentration increase in the agglomeration process, the value of existing patents diminishes because competition becomes more fierce between the varieties of the differentiated goods. This means that, when geographic concentration and growth increase, the part of incomes that comes from initially existing patents decreases, as can be readily seen by looking at equilibrium expenditures (16). Hence, because of its negative effect on the consumers' level of income and expenditures, an increase in the rate of technological innovation increases the relative differential in market size between the two regions which further reinforces the agglomeration mechanism. The centrifugal force is the localized demand by immobile consumers.

We have abstracted on purpose from the possibility of migration so as to focus on growth as the channel of agglomeration. We can, however, easily guess what would happen if we let some or all agents move from one region to the other. The price index is lower in region 1 than in region 2 because more firms produce in region 1
than in region 2 so that more goods can be bought free of transaction costs\textsuperscript{12} This readily implies in our framework that mobile workers will want to move to region 1 as real income is higher there (nominal incomes are identical due to wage and profit equalization), giving more weight to the agglomeration process because firms will also want to be close to these consumers in region 1. In this case, as more firms move to region 1, so that the cost of innovation decreases, growth increases due to the migration of workers to region 1.

The agglomeration cum growth mechanism can only be put in motion from a situation where growth is zero. When $\delta \leq \delta_1$, the economy is stagnating ($g = 0$), the mechanism behind agglomeration disappears, and the outcome with symmetric regions ($\gamma = 1/2$) is the only possible steady state. However, suppose that, from this situation where the economy is stagnating (zero growth) and regions are identical ($\gamma = 1/2$), one parameter changes so that the economy starts growing. For example, suppose that the market size $L$ increases. In this case, because the operating profits of firms buying patents are higher, the incentive to innovate increases and an innovation sector starts operating. Then, since the equilibrium with $\gamma = 1/2$ is not stable anymore, any small perturbation will cause an agglomeration process to take place, reinforcing the growth process. Such a situation describes an economy that starts growing and at the same time experiences the emergence of an industrial cluster as a center of innovation. The increases in the growth rate and in the geographic concentration of economic activities have come together and have reinforced each other. This example is consistent again with episodes of rapid rises in growth rates coupled with increases in spatial agglomeration.

5. GROWTH, AGGLOMERATION, AND TRADE

More insight on the behavior of the model can be gained by comparative statics. Except for $r$ and $p$ for which the impact on $g$ and $\gamma$ can also be found analytically, numerical simulations are necessary to judge the impact of changes of parameters. Given previous discussions, the results have intuitive appeal.\textsuperscript{13} Growth will be faster and agglomeration stronger, the larger the market size ($L$) as well as the share of differentiated products in consumers’ utility ($\alpha$) and the smaller the labor requirement ($\beta$), the elasticity of substitution ($\sigma$), the transaction cost ($\tau$), the innovation cost ($\eta$), as well as the subjective discount rate ($\rho$).

An interesting result is that the same parameters that increase growth also increase geographic agglomeration so that a positive correlation between growth rates and agglomeration rates across economies comes naturally as a result of the model.

As an example, consider $\sigma$. Lowering the value of $\sigma$ has a double effect. First, since $\sigma$ is the elasticity of substitution between varieties, a lower $\sigma$ increases the rate of innovation for a given location because monopolistic profits of the firms that buy a

\textsuperscript{12} The price index in location 1 is $P^* = [\beta \sigma/(\sigma - 1)]^{\sigma/(1-\sigma)}[\gamma + \delta(1 - \gamma)]^{\sigma/(1-\sigma)}$. It is lower than the symmetric price index calculated for location 2 as long as $\gamma$ is more than $1/2$ and it is decreasing in $\gamma$.

\textsuperscript{13} Given their intuitive appeal, the output of the simulations is not reported. The interested reader may refer to the discussion paper version (Martin and Ottaviano, 1996).
GROWTH AND AGGLOMERATION

patent are higher when differentiation (love of variety for consumers and diversity of intermediates for the innovation sector) is more pronounced. Second, since $\sigma$ is an inverse measure of equilibrium returns to scale, a lower $\sigma$ increases the agglomeration of economic activities for a given growth rate because the incentive to locate in the largest market is then higher. This is not a surprising result from what we know separately about the role of economies of scale from the endogenous growth literature and from the new geography literature. However, here the integration of the two models makes these two relations reinforce each other.

Of particular interest is the parameter of transaction costs. A decrease in transaction costs between the two regions will spur both growth and industrial concentration in the whole economy. In Equation (15), the effect is clear: a decrease in transaction costs decreases the cost of innovation in region 1, because it decreases the cost of the differentiated goods produced in region 2. This will increase the steady state growth rate for a given location of firms. In Equation (17), the location equation, the effect is less obvious because the decrease in transaction costs has three different effects on the location of firms. First, a decrease in transaction costs will give an incentive for more firms to move production to the region with the largest market because of increasing returns: the sensitivity of location to market size differentials increases. This means that some firms move to region 1 ($\gamma$ increases) because this is the region where all the innovation activities are located so that it is also the largest market for the differentiated goods. Second, a decrease in transaction costs also affects the market size differential between the two regions. This is because, as has already been noted in the previous section, a decrease in transaction costs reduces the cost of innovation, increases competition, and decreases the value of existing patents and therefore the consumers' income levels in both regions (see Equation (23)). This increases the relative importance of the innovation sector in the economy and therefore the relative expenditure gap between the two regions which induces firms to move to region 1; i.e., $\gamma$ increases. Third, a decrease in transaction costs reduces (for a given growth rate) the input demand of the innovation sector as less of the goods bought in region 2 melt during the transit. This last effect tends to decrease the demand for differentiated goods in region 1 and therefore tends to reduce the attractiveness of region 1. However, it can be shown analytically that this effect is always smaller than the other two effects so that, for a given growth rate, a decrease in transaction costs increases the concentration of economic activities in the region where the innovation sector is situated.

Both the location and the growth mechanisms go in the same direction. Given the circular causation, this means that the effects of a decrease in transaction costs on growth and agglomeration reinforce each other. The result that a decrease in trade costs spurs growth is similar to the one in Rivera-Batiz and Romer (1991), Grossman and Helpman (1991), and Baldwin and Forslid (2000). However, the channel is dif-

\[ \text{In equilibrium, } \frac{\sigma}{(\sigma - 1)} \text{ is the ratio of average to marginal cost.} \]

\[ \text{Lower trade costs in our model always push toward more agglomeration. In a model, such as Krugman and Venables (1990), where higher wages in the North may push firms to relocate in the South when transaction costs are very low, this result may not hold. This is because nominal wages are equalized across countries in our model.} \]
different as it goes through a direct effect on the cost of innovation and an indirect effect on the spatial distribution of economic activities which further decreases the cost of innovation. Furthermore, the impact is reinforced by the circular causation mechanism at play in our model.

More generally, the model generates an interesting relation between transaction costs and growth. Suppose that we look at a country that, through economic and political reforms, gradually decreases internal transaction costs between two regions. When transaction costs are high, such that \( \delta < \delta_1 \) given by (18), the cost of innovation is too high so that the growth rate of the country is zero and the symmetric geographical distribution of industry is stable. When transaction costs fall below the threshold \( \delta_1 \), growth becomes positive and the innovation sector agglomerates in one of the two regions. As transaction costs continue to decrease (\( \delta \) increases), then the growth rate increases for the reasons given above. Then at the level of transaction costs such that \( \delta = \delta_2 \), the complete agglomeration equilibrium (\( \gamma = 1 \)) becomes stable and any further decrease in transaction costs does not have any more impact on the growth rate which has reached its maximum. The maximum growth rate is given by Equation (15) where \( \gamma \) takes its maximum value 1. This relation is represented in Figure 4 below.

6. CONCLUSION

In this article we have constructed a model where the growth rate and the geography of economic activities (innovation and production) are jointly determined. From a methodological point of view, the contribution is to integrate a model of endogenous growth with a model of the “new geography.” From a theoretical point of view, interesting conclusions are derived both for growth theory and for location theory. Circular causation arises between growth and agglomeration: growth brings agglomeration that fosters growth. For this circular causation mechanism to take place, localized technological spillovers are not required; only market interactions between innovation and the intermediate inputs sector are. Starting from an initial situation
with no growth where two regions are identical (no agglomeration), we show that, when the aggregate economy starts growing, the only steady state outcome is one in which one of the two regions gets all the innovation activity and most of the industrial production.

Introducing growth dynamics in a geography model also generates different location dynamics. Agglomeration takes place in the core in the sense that innovation activities and the majority of production activities are located there but, as new economic activities are continuously created in the core, some firms will prefer to relocate and produce in the periphery where competition is less strong.

Moreover, we have shown that the same factors that spur growth also spur agglomeration and the cumulative process that we have identified reinforces the effect that a change of one factor has on both growth and agglomeration. In particular, we have shown that a decrease in transaction costs between regions of an economy encourages both agglomeration and growth of activities for the whole economy: the growth effect goes through the impact on geography and the agglomeration effect goes through the impact on growth. Hence, the positive correlation between agglomeration and aggregate growth of economic activities comes as a natural consequence of the economic forces at work.

REFERENCES


