Bröcker, Johannes

Working Paper

Agglomeration and Knowledge Diffusion

Economics working paper / Christian-Albrechts-Universität Kiel, Department of Economics, No. 2004,08

Agglomeration and Knowledge Diffusion

by Johannes Bröcker

CAU
Christian-Albrechts-Universität Kiel
Department of Economics

Economics Working Paper
No 2004-08

Agglomeration and Knowledge Diffusion

Johannes Bröcker

Abstract

According to New Growth Theory one can not rely on the convergence mechanisms inherent in traditional neoclassical constant returns to scale models. Convergence as well as divergence is possible, in general, depending on the assumptions about technology, factor mobility and ease of knowledge diffusion. The paper shows by a two-regions endogenous growth model under what conditions divergence, convergence or a stable centre-periphery structure emerge. The model allows for different degrees of knowledge diffusion as well as for different degrees of labor and capital mobility. The paper also evaluates dynamic market equilibria with respect to allocative efficiency. It is shown that the market solution tends to be under-agglomerated, except for parameter constellations generating particularly low agglomeration forces. If agglomeration forces are low enough, no concentration emerges, and this is also socially desirable. For higher agglomeration forces, however, concentration becomes desirable though the market may not bring it about or brings it about to an insufficient degree only.

Keywords: Convergence, divergence, agglomeration, endogenous growth, knowledge diffusion

JEL-Classification: R110, R130, O330, O410, F430

1 Introduction

Convergence versus divergence in a dynamic spatial economy is a classical issue in economics, that gained a lot of attention after the early contributions of Myrdal (1957) and Hirschman (1958), and later by Kaldor (1970). These authors tried to show that diverging tendencies would necessarily dominate converging ones in a growing market economy. Mobility of goods, people

[1] University of Kiel, Department of Economics. Olshausenstr. 40, D-24098 Kiel. Mailto: broecker@economics.uni-kiel.de
Thanks to three anonymous referees for helpful suggestions to extend the paper in several dimensions. I could take account of some but not all of them due to time and space limitations.

and capital would enhance this divergence. No explicit attention was given to mobility of knowledge and ideas, so that these authors left open whether the easing of communication over space would strengthen diverging or converging forces. Solovian growth theory took the counterpart by proving that under constant returns to scale incomes converge, and that convergence is even accelerated by any kind of mobility. Information mobility, though, was not really an issue in these theories either, because equal access to technology in all countries or regions was assumed from the very beginning.

There was little progress in theory until the upswing of New Economic Geography and New Growth Theory, both taking increasing returns and imperfect markets as natural starting points for explaining endogenous growth as well as endogenous agglomeration. Both branches have recently been joined in the work of Walz (1996, 1999), in a series of papers by Baldwin and co-authors (Baldwin and Forslid, 1997, 2000; Baldwin, 2001), by Martin and Ottaviano (2001) and by Fujita and Thisse in their recent monograph (Fujita and Thisse, 2002, chapter 11). In these models growth results from a steadily increasing diversity of goods. Investment comes in the form of ideas allowing to introduce new goods. Ideas are generated by a costly research process, as in the pioneering work of Romer (1990). Different spatial patterns emerge depending on the mobility of goods, capital and knowledge spillovers.

Due to the fact that monopolistic markets with transport costs as well as consumption and investment decisions of forward looking agents have to be handled in these models, they are highly complex. Hence, we try a simpler way by deriving long term growth from Marshallian externalities of capital. Due to these externalities capital has non-decreasing returns on the global level. Knowledge spillovers are hampered by distance, however. Our main question is how this distance deterrence of information flows affects the spatial distribution of economic activity as well as global efficiency. Furthermore, we try to figure out the interplay of factor mobility with the mobility of ideas. Due to the simplicity of the Marshallian externality approach we are able to draw clear-cut conclusions, both positive and normative ones. This simplicity comes at a cost, of course: a microfoundation of the innovation process is lacking in our approach, and we are unable to study also the interplay between factor mobility and mobility of ideas on the one hand, and freeness of trade on the other. This is a major issue of the cited literature.

2 A growth Model with Two Regions

2.1 Firms

We set up a simple growth model with a standard neoclassical production function with three factors of production:

- an immobile factor stock that cannot be accumulated; one may regard it as immobile labor or a combination of immobile labor, land and other natural resources;
- mobile labor;
- capital, which is understood as a combination of real capital and knowledge; knowledge has a local impact on output as well as a global one, that is an extra unit of capital in a region increases output in the region where it is invested, as well as in the other region, though to a lesser extent due to the intervening impediments of communication.

Let us consider two regions (i or $j = 1, 2$) that are completely identical except that one region may cover a larger stock of the immobile factor. One may think of two regions with the same area, but one having a larger density of the immobile part of the population. This asymmetry is exogenously given. We will also consider the completely symmetrical case, where both regions are a priori identical in every respect. The asymmetric case helps to understand the impact of an exogenous starting advantage of a region, everything else being the same in both regions.

Both regions produce the same homogeneous output Y_i by the same Cobb-Douglas technology,

$$Y_i = aL_i^{\lambda} M_i^{\mu} K_i^{\alpha} G_i^{\gamma}, \tag{1}$$

with immobile labor L_i, mobile labor M_i, and capital K_i. Furthermore, output depends on regional access to global knowledge measured by an index G_i. Global knowledge is generated as an externality of capital and knowledge accumulation in both regions. Regional access to this knowledge is assumed to decrease with increasing distance, as convincingly demonstrated empirically by Jaffe et al. (1993). Hence we assume

$$G_1 = K_1 + \beta K_2 \text{ and } G_2 = K_2 + \beta K_1,$$

with parameter $0 < \beta < 1$ measuring the intensity of interregional communication. $\beta = 0$ means infinite impediments to communication; there is no flow of knowledge. $\beta = 1$ means no impediments to communication, access to knowledge externalities is everywhere the same, irrespective of the location where the respective knowledge has been accumulated. The literature has extensively dealt with the limiting cases $\beta = 0$ (pure local knowledge) and $\beta = 1$ (pure global knowledge), which both characterize an extreme and unrealistic world.

The homogeneous good is freely traded between regions without transportation cost. It is either consumed or invested. One unit of the good is transformed one-to-one into one unit of installed capital. The increase of the capital stock per unit of time $\dot{K}_i$ in region i is gross investment I_i minus depreciation δK_i, that is

$$\dot{K}_i = I_i - \delta K_i.$$

Once nailed to the ground in one region, installed capital cannot be relocated (or can be relocated only at a cost that never makes a relocation worthwhile).

The parameters λ, μ, α und γ are the partial production elasticities of the respective factors. They are all assumed to be positive and less then one. Furthermore, we assume constant global returns to capital, which means $\alpha + \gamma = 1$. This is the "knife-edge" assumption usually made in growth modeling. It guarantees convergence to a constant global steady state rate of growth. A deviation from this assumption generates either explosive growth ($\alpha + \gamma > 1$), or stagnation ($\alpha + \gamma < 1$), contrary to observation. It is not nice that one has to build the theory on the assumption of a certain parameter to equal exactly unity without a deeper theoretical justification why the technology should tend to attain this parameter value. This is however a common drawback of the current state of endogenous growth theory, and we must take it as a still unresolved puzzle.

The effect of global knowledge represented by the index G_i is assumed to be entirely external. That means it is available as a public good, firms do not pay for making use of it. To be sure, there are also private knowledge flows in the economy; users in one location pay for knowledge accumulated elsewhere. But we disregard this kind of knowledge flows for the sake of simplicity. The elasticities λ, μ and α may represent internal as well as external effects of the respective factors of production, such that $\lambda = \lambda_p + \lambda_e$, $\mu = \mu_p + \mu_e$ and $\alpha = \alpha_p + \alpha_e$. λ_p and λ_e stand for the elasticities representing the internal and external effect, respectively; similarly for μ and α. We can apply perfect competition theory of pricing if we assume $\lambda_p + \mu_p + \alpha_p = 1$. A special case is where the only externality is that of G_i. In order to allow for perfect competition pricing for this case we must assume $\lambda + \mu + \alpha = 1$.

2.2 Households

Households are either mobile or immobile workers. Mobile workers are perfectly mobile, not facing any relocation costs. Hence they always choose the location offering the highest wage for mobile labor, that is the location with the highest private marginal productivity of mobile labor. All households are assumed to have identical linear-homogeneous preferences, such that they may be represented by a single representative household maximising the isoelastic utility

$$U = \int_0^\infty \frac{\sigma}{\sigma - 1} C(t)^{1-1/\sigma} \exp(-\rho t) \mathrm{d}t, \tag{2}$$

subject to the budget constraint requiring the present value of consumption not to exceed the present value of labor income. $C(t)$ is consumption at time t, ρ is the subjective discount rate, and σ is the elasticity of intertemporal substitution. Households maximise over an infinite horizon; their utility is regarded as the utility of an immortal family.

There is a perfect asset market. Households save by accumulating a risk free asset. As the economy is closed, the asset value equals the value of the stock of capital at any point in time. That means the asset value equals $q_1 K_1 + q_2 K_2$, where q_i denotes the stock price of capital installed in region i.

2.3 Dynamic Equilibrium

We begin with deriving the migration equilibrium for mobile workers. Due to the Cobb-Douglas form of the production function the marginal return of mobile labor in region i is $\mu_p Y_i / M_i$, such that $M_1/M_2 = Y_1/Y_2$ at any time due to perfect mobility. Hence, by (1) we obtain

$$Y_1/Y_2 = \ell^\lambda (Y_1/Y_2)^\mu k^\alpha \left(\frac{k + \beta}{1 + \beta k} \right)^\gamma$$

with $\ell := L_1/L_2$ and $k := K_1/K_2$. Solving for Y_1/Y_2 yields

$$m := M_1/M_2 = Y_1/Y_2 = \left[\ell^\lambda k^\alpha \left(\frac{k + \beta}{1 + \beta k} \right)^\gamma \right]^{1/(1-\mu)}. \tag{3}$$

Given ℓ and the elasticities, the distribution of mobile labor across regions only depends on the distribution of capital across regions, not on the total capital stock.

We normalise the total stock of mobile labor to unity, $M_1 + M_2 = 1$, such that $M_1 = m/(m+1)$ and $M_2 = 1/(m+1)$. Inserting this into (1) and dividing by K_i yields the average capital productivity

$$y_1 := Y_1/K_1 = aL_1^\lambda \left(\frac{m}{m+1}\right)^\mu K_1^{\alpha-1} G_1^\gamma,$$

and noting that $\alpha - 1 = -\gamma$ by assumption this is

$$y_1 = aL_1^\lambda \left(\frac{m}{m+1}\right)^\mu (1 + \beta/k)^\gamma, \tag{4}$$

and similarly

$$y_2 = aL_2^\lambda \left(\frac{1}{m+1}\right)^\mu (1 + \beta k)^\gamma. \tag{5}$$

Note that the private marginal productivity of capital in region i is $\alpha_p y_i$. Hence, these marginal productivities also only depend on the distribution of capital across regions, not on the total capital stock. This is of course the implication of the assumed constant global returns to capital.

Next, we derive households' consumption demand. Maximising (2) subject to the budget constraint yields the well known Keynes-Ramsey rule

$$\hat{C} = \sigma(r - \rho)$$

with interest rate r and consumption growth rate $\hat{C} = \dot{C}/C$. We generally denote the time derivative of a time dependent variable X as $\dot{X}$ and its growth rate as $\hat{X} := \dot{X}/X$. Furthermore, the transversality condition must hold stating that the present value of the household's assets must tend to zero as t tends to infinity (Barro and Sala-i-Martin, 1995, section 2.1).

Next we have to deal with investments and the asset market. As the homogeneous good can be transformed into installed capital one-to-one without adjustment cost, investment demand would be infinite if q_i was larger than one in one of the regions. Private optimising investors are willing to invest in region i, as long as $q_i = 1$. Hence, the complementarity

$$I_i \geq 0, \quad q_i \leq 1, \quad I_i(1 - q_i) = 0$$

must hold in both regions.

There are two kinds of assets that can be held, capital in region 1 and capital in region 2. Both must yield identical private rates of return,

$$rq_i = \dot{q}_i + \alpha_p y_i - \delta q_i.$$

There are two possible cases to be distinguished:

(1) $q_i = 1$ in a time span of positive duration. Then $\dot{q}_i = 0$, and hence $r = \alpha_p y_i - \delta$. In this case $I_i \geq 0$.
(2) $q_i < 1$. Then $I_i = 0$.

Positive investment in both regions is only possible if $r = \alpha_p y_i - \delta$ holds in both regions, that is if $y_1 = y_2$. Note however that $q_1 = 1$ and $q_2 < 1$ does not necessarily imply that capital is more productive in region 1 than in region 2. Even if productivity is bigger in region 2, stock prices $q_1 = 1$ and $q_2 < 1$ can occur, provided that q_2 declines sufficiently fast. We will see below that this situation in fact is observed near unstable steady states.

Finally, the goods market equilibrium condition

$$K_1 y_1 + K_2 y_2 = I_1 + I_2 + C$$

closes the system.

It turns out that there are three types of dynamic equilibria:

(1) Balanced steady state. In this case investment in both regions is positive. This implies $q_1 = q_2 = 1$, and hence $y_1 = y_2$, which in turn implies $\hat{K}_1 = \hat{K}_2$, because k must stay constant. Therefore the interest rate $r = \alpha_p y_1 - \delta = \alpha_p y_2 - \delta$ is constant, and $\hat{C} = \sigma(r - \rho)$ is also constant. Defining $Y := Y_1 + Y_2$ and $K := K_1 + K_2$ we can write

$$Y = AK$$

with $A := y_1 = y_2$. Hence the entire economy becomes a so-called AK-economy (Barro and Sala-i-Martin, 1995, section 4.1). Goods market equilibrium and transversality can only hold if $\hat{Y} = \hat{C}$. Hence we have

$$\hat{Y} = \hat{C} = \hat{K} = \sigma(\alpha_p A - \delta - \rho). \tag{6}$$

The saving rate is

$$s := \frac{Y - C}{Y} = \frac{\dot{K} + \delta K}{K} \frac{K}{Y} = (\hat{K} + \delta)/A.$$

The growth rate is the higher, the higher is capital productivity (which is equalised across regions) and the lower is δ. It is also the higher, the smaller is ρ and the higher is σ, i.e. the more patient consumers are and the less keen on intertemporal smoothing of consumption they are. These are of course standard results of endogenous growth theory.

(2) Concentrated steady state. This is the limiting case for k tending to infinity or to zero, provided that $\mu > \gamma$. If k tends to infinity in the course of time, output and mobile labor are eventually completely concentrated in region 1. Hence, m tends to infinity and eventually we have $y_1 = aL_1^\lambda$ by equation (4). We will see below that y_1/y_2 tends to infinity; hence y_2 tends to zero. Therefore in the end we again have an AK-economy

$$Y_1 = AK_1$$

with $A := y_1 = aL_1^\lambda$ constant, and with steady state growth rate as in equation (6). Similarly for k tending to zero.

(3) Transition with k increasing or decreasing. It is sufficient to deal with "transitions from the left", i.e. with k increasing over time. For this case one can show that $q_1 = 1$ and almost always $q_2 < 1$. That q_1 cannot be less than one is obvious: $q_1 < 1$ would imply $\hat{K}_1 = -\delta$ and — as $\hat{K}_2 \geq -\delta$ — it would imply $\hat{K}_1 - \hat{K}_2 \leq 0$, contradicting $\hat{k} > 0$. For proving that almost always $q_2 < 1$, assume to the contrary that $q_2 = 1$ in a proper time interval. Then $\dot{q}_1 = \dot{q}_2 = 0$, and hence $y_1 = y_2$ during that interval, contradicting non-constancy of k. Therefore in a transition with increasing k there is only investment in region 1 and no investment in region 2, that is K_2 declines with rate δ.

The transition is described by the differential equations

$$\dot{K}_1 = Y_1 + Y_2 - \delta K_1 - C, \tag{7}$$
$$\dot{K}_2 = -\delta K_2, \tag{8}$$
$$\dot{C} = C\sigma(\alpha_p y_1 - \delta - \rho), \tag{9}$$
$$\dot{q}_2 = \alpha_p y_1 q_2 - \alpha_p y_2, \tag{10}$$

which hold as long as $q_2 < 1$. (7) is the goods market equilibrium condition, (8) is due to non-investment in region 2, (9) is the Keynes-Ramsey rule, and (10) is the asset market equilibrium condition.

Equations (7) to (10) can be transformed into a system of two coupled autonomous differential equations in k and $c := C/K_2$. Substituting $\dot{K}_2 k + \dot{k} K_2$ for $\dot{K}_1$ in (7), using $\dot{K}_2$ from (8) and dividing through by K_2 yields

$$\dot{k} = ky_1 + y_2 - c. \tag{11}$$

Furthermore, substituting $\dot{K}_2 c + \dot{c} K_2$ for $\dot{C}$ in (9), using $\dot{K}_2$ from (8) and dividing through by K_2 yields

$$\dot{c} = c[\sigma(\alpha_p y_1 - \delta - \rho) + \delta]. \tag{12}$$

In the following we study transitions to balanced steady states "from the left", i.e. transitions with increasing k, by numerically integrating the system (11), (12) and (10) backwards in time, starting from known steady state values for k, c, and q_2. These steady state values are

- the $\bar{k}$ that solve the equation $y_1 = y_2$ at points where $f(k)$ defined in (14) below has a negative slope (there may be one or two of them),
- $\bar{c} = (1 + \bar{k})(y_1 - \hat{K} - \delta)$, with $\hat{K}$ from (6),
- and $\bar{q}_2 = 1$.

(q_1 stays constant on a path with k increasing over time.) Transitions from the right are obtained by just exchanging indices of regions. As the two-dimensional system (11) and (12) is autonomous, the saddlepath-stability of the dynamic equilibria could be shown by the standard two-dimensional phase space analysis, but this is omitted here.

For simulating a transition to a steady state concentrated in region 1 we integrate a similar system in the variables $h := 1/k$ and $\tilde{c} := 1/K_1$ backwards in time starting at $\bar{h} = 0$, $\bar{q}_2 = 1$, and $\bar{\tilde{c}} = aL_1^\lambda - \hat{K} - \delta$. $\hat{K}$ is given by (6) with $A = aL_1^\lambda$. Note that, unlike balanced steady states, concentrated steady states are only attained in infinite time. (It is therefore necessary to substitute another variable – for example path length – for time in the integration.)

3 Dynamics: Convergence and Divergence

For studying the dynamics of this economy, one has to see how capital productivities in both regions depend on the distribution of capital across the regions, given that mobile labor is distributed such that wages are equalized across regions at any time. Let R denote the ratio of marginal productivities of capital (region 1 over region 2). It is a function of the ratio of capital stocks $k = K_1/K_2$ and the ratio of stocks of immobile labor $\ell = L_1/L_2$:

$$R = \left[\ell^\lambda k^{\mu-\gamma} \left(\frac{k+\beta}{1+\beta k} \right)^\gamma \right]^{1/(1-\mu)}. \tag{13}$$

To see this, divide equation (3) by k and note that $\alpha - (1-\mu) = \mu - \gamma$. R is larger than one, equal to one, or less than one, if and only if the expression

$$f(k) = \frac{\mu - \gamma}{\gamma} \log k + \log \frac{k+\beta}{1+\beta k}. \tag{14}$$

is larger than, equal to, or less than $f^* := -(\lambda/\gamma) \log \ell$. In particular, in the symmetric case $R > 1$ ($R = 1$, $R < 1$), if and only if $f(k) > 0$ ($f(k) = 0$,

$f(k) < 0$). A balanced steady state is attained at point(s) k^* solving $f(k^*) = f^*$.

Intuitively, if f is positively sloped at k^*, an increase of k generates a capital productivity advantage of region 1 making region 1 even more attractive for capital investment. Hence, we call anything making f positively (negatively) sloped a divergence (convergence) force. (Note though, that with perfect foresight the answer as to whether k, starting at some k_0 close to k^*, will increase or decrease is a bit more involved, as will be shown in a moment.)

The first term in (14) shows a divergence (convergence) force, if $\mu > \gamma$ ($\mu < \gamma$). Noting that $\alpha + \gamma = 1$, the condition $\mu > \gamma$ is equivalent to $\mu + \alpha > 1$, that is to increasing local returns of the mobile factors K and M. In other words, the first term shows a divergence (convergence) force, if mobile factors exhibit increasing (decreasing) local returns. The second term in (14) is a divergence force, except for $\beta = 1$ (perfect knowledge mobility), where it vanishes. For $\beta < 1$ the term is strictly increasing in k with lower bound $\log \beta$ and upper bound $-\log \beta$. Obviously, the smaller β, the stronger is the divergence force.

Hence, if mobile factors exhibit increasing local returns, there are only divergence forces, while there is a convergence and a divergence force if mobile factors exhibit decreasing local returns. The convergence force will eventually always dominate, if k gets sufficiently large or sufficiently small, while the dominance at the balance point k^* at or near zero depends on parameters. This gives rise to three possible scenarios to be explained in the the following subsections.

3.1 Divergence

If $\mu > \gamma$ (or equivalently $\mu + \alpha > 1$), $f(k)$ is strictly increasing for all k (see figure 1). Note that this case would be impossible with $\lambda + \mu + \alpha = 1$. As

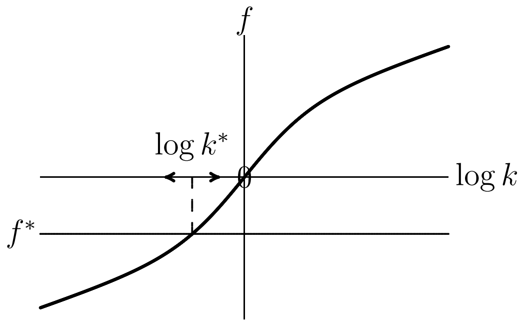

Fig. 1. Divergence

explained before, we can however allow for $\lambda + \mu + \alpha > 1$ and still maintain marginal productivity factor payments if we assume these elasticities to have internal and an external components, with the former ones adding up to unity.

A naive model with myopic agents investing exclusively in the region with the currently higher marginal productivity of capital would reveal k^* in figure 1 to be the watershed between two possible time paths. If k starts from $k_0 > k^*$, capital moves to region 1 with higher capital returns, and mobile labor follows. This even increases the incentive to move, such that eventually capital and mobile labor is completely concentrated in region 1. Similarly one ends up with perfect concentration in region 2 if one starts from $k_0 < k^*$.

With perfect foresight the intuition that k^* is unstable still turns out to be true. If k slightly deviates from k^*, there is no equilibrium path leading back to k^*. For a transition from the left, q_2 would have to approach unity from above; otherwise shares in K_2 would yield higher returns than those in K_1, because left of k^* marginal productivity in region 2 exceeds that in region 1. A similar argument holds for transition from the right.

The dynamics leading the economy to a concentrated equilibrium either in region 1 or in region 2 with perfect foresight are more complicated than in a myopic world. Figure 2 illustrates the symmetrical case by plotting asset prices over k in phase space. The solid (dashed) curve shows q_2 (q_1) for a

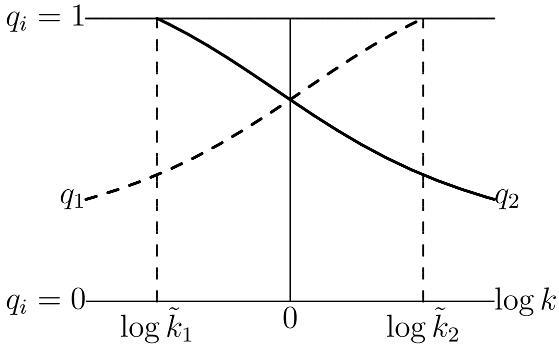

Fig. 2. Asset prices in transition to concentrated equilibria

transition to a concentrated steady state in region 1 (region 2). Let us call the sets $\{k|k \geq \tilde{k}_1\}$ and $\{k|k \leq \tilde{k}_2\}$ the attraction domains of regions 1 and 2, respectively. For any k in the attraction domain of region 1 (region 2) there is an equilibrium path leading to a concentrated steady state in region 1 (region 2). Note that the two domains overlap. Within the overlap the path is non-unique. Which one of both possible paths is chosen depends on self-fulfilling prophecies.

Consider for example a start at $k_0 = 1$. If asset owners are convinced that the economy moves towards a concentrated equilibrium in region 1, then q_2 drops to some price less than one and all investment goes to region 1. In a myopic world agents holding shares in the capital stock of region 2 would want to be compensated by an asset price below one such that $q_2(\alpha_p y_1 - \delta) = \alpha_p y_2 - \delta$. For $k = k^* = 1$ this would just mean $q_1 = q_2 = 1$, because $y_1 = y_2$ at k^*. Hence, in a myopic world attraction domains would not overlap, as already shown above. With perfect foresight, q_2 must be less than one at k^*, because agents predict q_2 to fall.

Figure 3 illustrates the asymmetric case with region 1 having more immobile factors (i.e. $\ell > 1$). The attraction domains shift leftwards, but still overlap, and the unstable balanced steady state is still in the interior of the overlap.

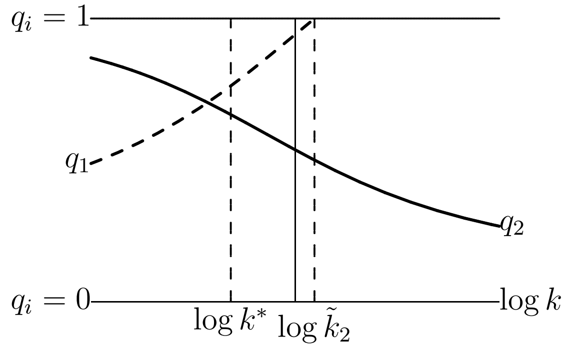

Fig. 3. Asset prices in transition to concentrated equilibria

3.2 Agglomeration

If $1 > \mu/\gamma > 2\beta/(1+\beta)$ (area AA in figure 6 below), mobile factors exhibit decreasing local returns. But if capital is equally distributed across regions ($k = 1$, i.e. $\log k = 0$) the divergence effect of the second term in equation (14) dominates, that is the divergence effect caused by by the global external effect of capital. If capital is concentrated sufficiently in one or the other region, however, the first term eventually dominates, which works into the direction of convergence due to decreasing local returns of mobile factors.

Depending on ℓ (or f^*), there are two possible scenarios:

(1) If regions are sufficiently similar regarding their stocks of immobile factors, i.e. if ℓ is sufficiently close to unity, then there are three balanced steady states, an unstable one in the middle and two stable ones (called agglomerations) with high concentration of mobile factors either in region 1 or in region 2 (see figure 4).
(2) If the regions sufficiently differ with regard to their stocks of immobile factors, then there is only one balanced steady state. It is stable and located in region 1 (region 2), if $\ell > 1$ ($\ell < 1$); see figure 5.

Let us discuss the more difficult first case with three equilibria. In a myopic world, starting from $k_0 > k^*$ would always lead to an agglomeration in region

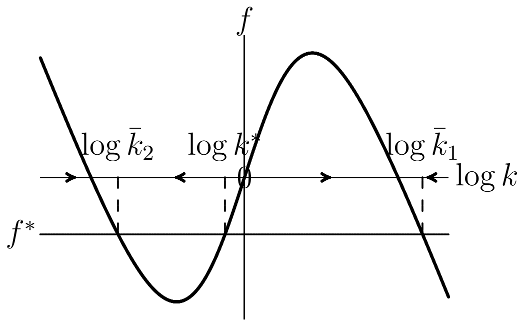

Fig. 4. Agglomeration

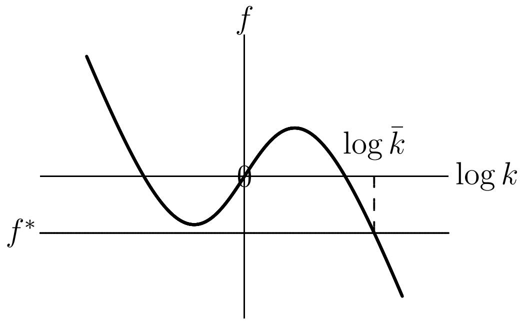

Fig. 5. Region 2 is too small for an agglomeration

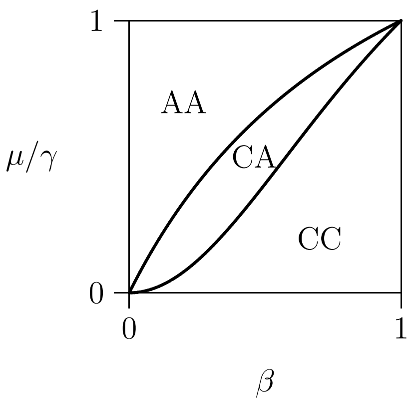

Fig. 6. AA = the market generates agglomeration, efficiency requires a higher degree of agglomeration; CA = the market generates convergence, while efficiency requires agglomeraion; CC = the market generates convergence, which is efficient.

1, because an increasing concentration of capital in region 1 makes it even more attractive to invest in region 1, as long as capital is not too concentrated in region 1 (see figure 4). Later, decreasing local returns dominate and bring the agglomeration process to a halt, once the balanced steady state distribution $\bar{k}_1$ is attained. This happens after a finite transition period. Similarly, the economy converges to an agglomeration in region 2, if it starts from $k_0 < k^*$.

With perfect foresight we again observe overlapping attraction domains of the two balanced steady states, and the unstable balanced steady state k^* is in the interior of the overlap. Figure 7 illustrates the symmetrical case. The solid (dashed) curves show transitions to an agglomerated steady state in region

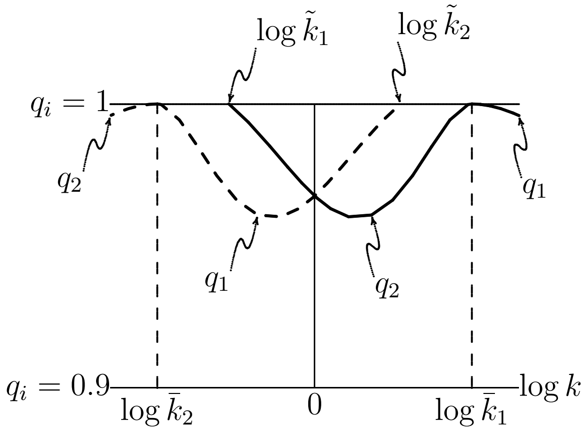

Fig. 7. Asset prices in transition to agglomerated steady states

1 (region 2). Curve segments left of the steady states (transitions from the left) show q_2, while q_1 is unity. Curve segments right of the steady states (transitions from the right) show q_1, while q_2 is unity.

Figure 8 illustrates the asymmetric case. Increasing ℓ shrinks the attraction

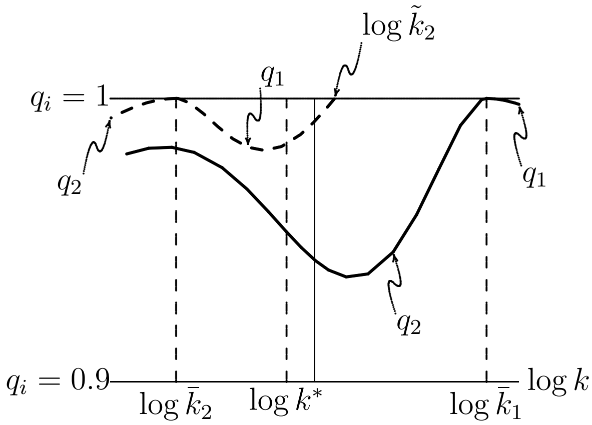

Fig. 8. Asset prices in transition to agglomeration steady states

domain of region 2 right of $\bar{k}_2$ and extends the attraction domain of region 1 left of $\bar{k}_1$. It actually may extend to infinity, if ℓ is large enough (as it is the case in the simulation results shown in the figure).

The dynamic response to decreasing agglomeration forces in case of an agglomeration locked in in the smaller region is worth studying ($\bar{k}_2$ in figure 4 or figure 9). Assume that improved means of communication allow for a better access to to knowledge in distant locations (β increases), then the curve in figure 9 becomes flatter. k increases from $\bar{k}_2$ to $\check{k}_2$ and the attraction domain vanishes. Now the agglomeration in region 2 can no longer survive. At this point in time at the latest, one will observe a catastrophic transition; the agglomeration in region 2 vanishes and a new agglomeration will emerge in region 1. All investments go to region 1. Mobile labor will follow capital to

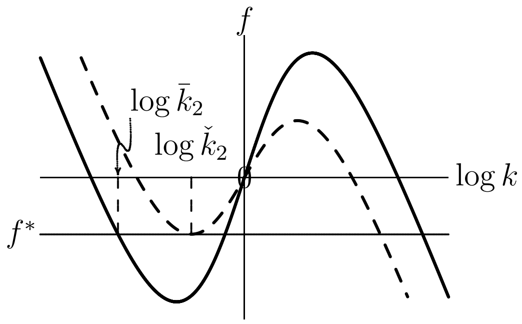

Fig. 9. Catastrophic dissolution of an agglomeration in region 2

region 1. Empirically we would first observe gradual convergence and then a leap-frog: the former poor region would rapidly overtake. Eventually the distribution converges to a new steady state with the larger region being the agglomerated center and the smaller one the poorer periphery.

In a myopic world the dissolution of the agglomeration in region 2 cannot occur before this bifurcation point is attained. Under perfect foresight, however, it can. Take the balanced steady state $\bar{k}_2$ with an agglomeration in region 2 in figure 8 as a case in point. If agents collectively do not believe in an agglomeration in region 2 anymore and notice the transition path to an agglomeration in region 1, q_2 drops down, investment in region 2 stops, all investment goes to region 1 and k goes up until $\bar{k}_1$ is attained.

3.3 Convergence

If $2\beta/(1+\beta) > \mu/\gamma$ (areas CA and CC in figure 6) the economy converges to the unique balanced steady state, which is now stable (see figure 10). Locally

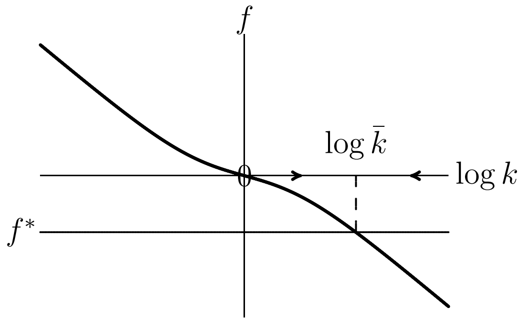

Fig. 10. convergence

decreasing returns dominate even for a balanced distribution of capital. In the symmetrical case ($\ell = 1$) there will be unconditional convergence. From any starting point factor prices are equalised in the course of time, and eventually the economy grows in a balanced steady state with a constant rate. There is conditional convergence, if $\ell \neq 1$. Then the economy converges to a unique steady state as well ($\bar{k}$ in figure 10). But income per capita and the wage rate of immobile labor higher in the larger region than in the smaller one in the steady state.[2] We dispense with plotting asset prices over k. The pattern is obvious for this case: q_2 (q_1) approaches unity from below, if k approaches $\bar{k}$ from the left (right).

4 Efficiency

What about allocative efficiency of the outcome of market forces in this economy? Imagine an omniscient planner aiming at maximising the household's utility subject to the technological constraints of the economy. He solves the problem

$$\max_{C,K_1,K_2} \int_0^\infty \frac{\sigma}{\sigma-1} C(t)^{1-1/\sigma} \exp(-\rho t) \mathrm{d}t$$

subject to

$$\begin{aligned} \dot{K}_i &= I_i - \delta K_i, \quad I_i \geq 0, && (15) \\ I_1 + I_2 + C &= K_1 y_1 + K_2 y_2, && (16) \\ K_i(0) &= K_i^0. \end{aligned}$$

y_1 and y_2 are as given by (4) and (5), with m from (3). Strictly speaking, the planner also has to choose m. He would however make the same choice as the market: he would distribute mobile labor across regions such that $m = Y_1/Y_2$. Though we allowed for a positive externality of mobile labor, it leaves the migration decision undistorted. This is why we can eliminate the choice dimension m, taking for granted that at any point in time mobile labor is distributed optimally across regions.

Letting π_i denote the costate variables associated with (15), and η denote the Lagrangian multiplier associated with (16), we obtain the first order conditions derived from the present value Hamiltonian:

$$\begin{aligned} C^{-1/\sigma} \exp(-\rho t) &= \eta, && (17) \\ \pi_i - \eta \leq 0, &\quad I_i(\pi_i - \eta) = 0, && (18) \\ -\dot{\pi}_i &= -\pi_i + \eta F_{K_i}, && (19) \\ \lim_{t\to\infty} \{\pi_i K_i\} &= 0. && (20) \end{aligned}$$

[2] Let $\ell > 1$, as assumed in the figure. Then $\bar{k} > 1$, hence $G_1 > G_2$ in the steady state. This capital productivity advantage of region 1 must be compensated by a higher capital intensity; this implies a higher per capita income in region 1.

F_{K_i} denotes marginal social productivity of capital, that is

$$F_{K_i} := \frac{\partial F(K_1, K_2)}{\partial K_i},$$

with $F(K_1, K_2) := K_1 y_1 + K_2 y_2$.

Now define the "social interest rate" $v := -\hat{\eta}$ and $\phi_i := \pi_i/\eta$. Then taking logs of (17) and derivating with respect to time yields

$$\hat{C} = \sigma(v - \rho).$$

(18) becomes

$$\phi_i \leq 1, \quad I_i \geq 0, \quad I_i(1 - \phi_i) = 0.$$

Inserting $\dot{\pi}_i = \phi_i \dot{\eta} + \dot{\phi}_i \eta$ into (19) and dividing by η yields

$$v\phi_i = \dot{\phi}_i + F_{K_i} - \delta\phi_i.$$

Finally, (20) becomes

$$\lim_{t\to\infty} \{\eta\phi_i K_i\} = 0,$$

saying that the present value of both types of assets have to vanish as time goes to infinity. "Present value" here means to discount with the social interest rate v rather than the private interest rate r. Putting it differently, the transversality condition now reads

$$\lim_{t\to\infty} \left\{ \phi_i(t) K_i(t) \exp\left(-\int_0^t v(\tau) \mathrm{d}\tau \right) \right\} = 0.$$

These equations coincide with those describing the decentralised market equilibrium, if we substitute r and q_i for v and ϕ_i, except that in the planner's problem the social marginal capital productivity F_{K_i} takes over the role of the private marginal capital productivity $\alpha_p y_i$ in the market solution.

To be sure, the long run steady state of the market economy is inefficient in that the saving rate and, as a consequence, the growth rate are too small as compared to the optimal path. This is due to the fact that the households' saving decision is based on the private rate of return to capital, while the planner bases his decision on the higher social rate of return, which includes the positive capital externality. Just as the market equilibrium path, the optimal path converges to either a balanced or a concentrated steady state. The optimal steady state growth rate is (similar to (6))

$$\hat{Y} = \hat{C} = \hat{K} = \sigma(F_K - \delta - \rho).$$

F_K is the social marginal productivity of capital, which is larger than the private one because of the two externalities, possibly a local one stemming from $\alpha > \alpha_p$, and a global one stemming from the effect of G_i in the production function. Hence, the optimal steady state rate of growth is always larger than the steady state rate of growth in a decentralised market equilibrium. If the steady state is balanced, then $F_K = F_{K_i}$ with k such that $F_{K_1} = F_{K_2}$. We will see in a minute, that this k is always larger than the $\bar{k}$ equalising private rates of capital return.

Our focus is now on the spatial distribution in the steady state: how does the capital distribution in the optimal steady state compare with the market solution? Does the market generate over-agglomeration or under-agglomeration? I confine the analysis to the symmetrical case ($\ell = 1$).

The optimal k is found by derivating $Y = Y_1 + Y_2$ with respect to k, holding $K = K_1 + K_2$ constant. The resulting derivative has the same sign as the function $h(k) = f(k) + g(k)$ with

$$g(k) = \log\left(\alpha + (1-\beta)\gamma\frac{k}{k+\beta}\right) - \log\left(\alpha + (1-\beta)\gamma\frac{1}{1+\beta k}\right).$$

$g(k)$ is strictly monotone increasing, goes through zero for $k = 1$ (i.e. $\log k = 0$) and approaches the lower bound $-\bar{g}$ to the left and the upper bound $\bar{g}$ to the right, with

$$\bar{g} = \log(1 + (1-\beta)\gamma/\alpha) > 0$$

(see figure 11). $g(k)$ is the positive agglomeration externality, which is neglected by private investors when choosing the location of investment. The smaller β (that is the less easily innovation diffuses over space) and the larger γ/α (that is the more important global effects of capital as compared to local effects), the larger is this externality.

Comparing $h(k)$ with $f(k)$ allows to distinguish the following cases:

(1) $\mu/\gamma > 1$: The market leads to a complete concentration of capital in one region, and this is optimal.

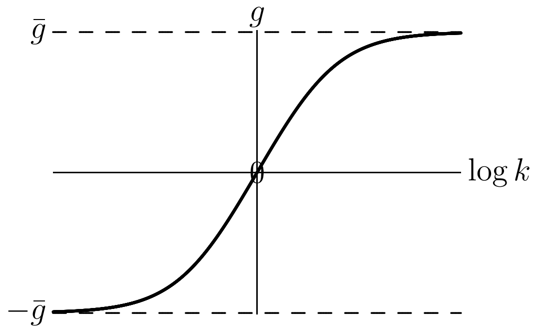

Fig. 11. Function $g(k)$ representing the agglomeration externality

(2) $1 > \mu/\gamma > 2\beta/(1+\beta)$ (Area AA in figure 6): The market leads to an incomplete concentration of capital in one region; a higher concentration would be optimal.

(3) $2\beta/(1+\beta) > \mu/\gamma > c(\beta,\alpha)2\beta/(1+\beta)$ with a positive factor $c(\beta,\alpha) < 1$ depending on β and α (area CA in figure 6)[3]: Capital converges against a symmetrical distribution, but an incomplete concentration would be optimal.

(4) $c(\beta,\alpha)2\beta/(1+\beta) > \mu/\gamma$ (area CC in figure 6): Capital converges against a symmetrical distribution, and this is optimal.

To summarise: the market allocation is either optimal — case 1 with complete concentration, case 4 with equal distribution —, or leads to an insufficient concentration (cases 2 and 3). In case 2 the market induces agglomeration, but not to a sufficient degree. In case 3 the market forces of agglomeration are not strong enough to bring the desirable agglomeration about.

5 Conclusion

This paper showed how factor mobility and knowledge diffusion influenc the spatial distribution of economic activity in a growing economy. Long run growth is due to a positive external effect of knowledge, such that capital, which is understood to consist of real capital and knowledge capital, has globally constant returns to scale. Access to knowledge is however not everywhere the same. Each region only partly participates in the knowledge generated elsewhere.

If mobile factors have jointly increasing local returns to scale, then we always observe a diverging growth path such that in the long run all activity is completely concentrated in one region. The concentration as such is efficient, while the steady state rate of growth is too small as compared to an efficient path,

[3] In figure 6 the boundary between areas CA und CC is drawn for realistic parameter values $\gamma = \alpha = 1/2$.

because the saving decision neglects the positive capital externality. The latter result is standard in endogenous growth models with Marshallian capital externalities, but does not necessarily hold in models where creative destruction of innovation is admitted.

If mobile factors have jointly decreasing local returns, the economy either attains an agglomerated steady state with most but not all activity concentrated in one location, or it converges to a steady state with total equality of the regions or only moderate differences (in case of exogenous size differences of the regions). Agglomeration is the more likely, the worse the access to knowledge generated elsewhere is, and the smaller the production elasticity of global knowledge is in comparison to that of mobile labor. An empirical implication is that we should observe a steady deglomeration process in the course of declining communication cost, increasing access to knowledge and increasing importance of global knowledge as a production factor.

The paper also evaluates dynamic market equilibria with respect to allocative efficiency. It is shown that the market solution tends to be under-agglomerated, except for parameter constellations generating particularly low agglomeration forces. If agglomeration forces are low enough, no agglomeration emerges, and this is also socially desirable. For higher agglomeration forces, however, concentration becomes desirable though the market may not bring it about or brings it about to an insufficient degree only. This conclusion contradicts commonly held wisdom in regional policy supporting equalisation between regions and aiming at a shift of resources from richer to poorer regions. This type of cohesion policy might be justified from a distributional point of view. It favours owners of immobile factors in the periphery and may therefore also support political stability. But, in the light of our theoretical results, it cannot be justified for efficiency reasons.

References

Baldwin, R., 2001. Core-periphery model with forward-looking expectations. Regional science and urban economics 31, 21–49.

Baldwin, R., Forslid, R., 1997. Trade liberalization and endogenous growth: a q-theory approach. Journal of International Economics 50, 497–517.

Baldwin, R., Forslid, R., 2000. The core-periphery model and endogenous growth: stabilizing and destabilizing integration. Economica 67, 307–324.

Barro, R., Sala-i-Martin, X., 1995. Economic Growth. McGraw-Hill, New York.

Fujita, M., Thisse, J.-F., 2002. Economics of Agglomeration: Cities, Industrial Location, and Regional Growth. Cambridge University Press, Cambridge.

Hirschman, A. O., 1958. The Strategy of Development. Yale University Press, New Haven, CN.

Jaffe, A., Trajtenberg, M., Henderson, R., 1993. Geographic localization of knowledge spillovers as evidenced by patent citations. Quarterly Journal of Economics 108, 577–598.

Kaldor, N., 1970. The case of regional policies. Scottish Journal of Political Economy 18, 337–348.

Martin, P., Ottaviano, G., 2001. Growth and agglomeration. International Economic Review 42, 947–968.

Myrdal, G., 1957. Economic Theory and Underdeveloped Regions. Duckworth, London.

Romer, P. M., 1990. Endogenous technological change. Journal of Political Economy 98, S71–S102.

Walz, U., 1996. Transport costs, intermediate goods, and localized growth. Regional Science and Urban Economics 26, 671–695.

Walz, U., 1999. Dynamics of regional integration. Physica, Heidelberg.