9

Industrial Agglomeration under Monopolistic Competition

9.1 INTRODUCTION

The spatial economy is replete with *pecuniary externalities*. For example, when some workers (firms) choose to migrate, they are likely to affect both the labor and product markets in their region of origin as well those in their region of destination. Such pecuniary externalities are especially relevant in the context of imperfectly competitive markets because prices do not perfectly reflect the social values of individual decisions. Pecuniary externalities are also better studied within a general equilibrium framework to account for the interactions between the product and labor markets. Among other things, this allows one to study the dual role of individuals as workers and consumers. At first sight, this seems to be a formidable task. Yet, as shown by Krugman (1991a), several of these various effects can be combined and studied within a relatively simple general equilibrium model of monopolistic competition, which has come to be known as the *core–periphery model*.[1]

Recall that monopolistic competition à la Chamberlin (1933) involves consumers with a preference for variety (*varietas delectat*), whereas firms producing these varieties compete for a limited amount of resources because they face increasing returns. The prototype that has emerged from the industrial organization literature is the model developed by Spence (1976) and Dixit and Stiglitz (1977), which is sometimes called the S–D–S model. These authors assumed that each firm is negligible in the sense that it may ignore its impact on, and hence reactions from, other firms, but retains enough market power for pricing above marginal cost regardless of the total number of firms (like a monopolist). Moreover, the position of a firm's demand depends on the actions taken by all firms in the market (as in perfect competition).

In many applications, the S–D–S model has proven to be a very powerful instrument for studying the aggregate implications of monopoly power and increasing returns – especially so when these elements are the basic ingredients of self-sustaining processes such as those encountered in modern theories of

growth and economic geography (Matsuyama 1995). This is so because of the following reasons:

1. Although each firm is a price-maker, strategic interactions are very weak in this model, thus making the existence of an equilibrium much less problematic than in general equilibrium under imperfect competition (see, e.g., Bonanno 1990).
2. The assumption of free entry and exit leads to zero profit so that a worker's income is just equal to his wage, which is another major simplification.[2]
3. Although it is seldom noticed, the difference between price competition and quantity competition becomes immaterial in monopolistic competitive settings. Indeed, being negligible to the market, each firm behaves as a monopolist on the residual demand, which makes it indifferent between using price or quantity as a strategy.[3]

The setting considered by Krugman, described in Section 9.2, is simple but rich enough to capture the major forces at work in the multiregional economy. It combines the Dixit–Stiglitz model with the iceberg-type transport cost in which, as in Thünen, only a fraction of the good shipped reaches its destination (Samuelson 1954a; 1983). There are two sectors: (1) the agricultural sector producing a homogeneous good under constant returns to scale and (2) the manufacturing sector supplying a differentiated good under increasing returns. As usual, the market equilibrium is the outcome of the interplay between a dispersion force and an agglomeration force. The centrifugal force is simple. It lies in the following two sources: (1) the spatial immobility of farmers whose demands for the manufactured good are to be met and (2) the increasing competition that arises when firms are more agglomerated. The centripetal force is more involved. If a larger number of manufactures are located in one region, the number of varieties locally produced is also larger. Then, because firms do not price discriminate between regions, the equilibrium price index of manufactured goods is lower in this region. This in turn, induces some workers living in the smaller region, to move toward the larger region, where they may enjoy a higher standard of living. The resulting increase in the numbers of workers creates a larger demand for the differentiated good, which, therefore, leads additional firms to locate in this region. This implies the availability of more varieties in the region in question but less in the other because there are scale economies at the firm's level. Consequently, as noticed by Krugman (1991a, 486), *circular causation* à la Myrdal is present because these two effects reinforce each other: "manufactures production will tend to concentrate where there is a large market, but the market will be large where manufactures production is concentrated."

When transportation costs (or, more generally, trade costs) are sufficiently low, Krugman (1991a) has shown that all manufactures are concentrated in a single region that becomes the *core* of the economy, whereas the other region, called the *periphery*, supplies only the agricultural good. Firms are able to

exploit increasing returns by selling more in the larger market without losing much business in the smaller market. For exactly the opposite reason, the economy displays a symmetric regional pattern of production when transportation costs are high. Hence, this model allows for the possibility of divergence between regions, whereas the neoclassical model, based on constant returns and perfect competition in the two sectors, would predict convergence only. As in previous chapters, there would be a monotone decreasing relationship between the degree of agglomeration and the level of transport costs. In addition, Krugman has also shown that there are often multiple equilibria; thus, there is a need for some refinement. As before, instability is used to dismiss some equilibria. Several stable equilibria may exist; consequently, the market outcome is likely to depend on the initial conditions. Or, to put it differently, history matters.[4] The process of circular causation, then, looks like a snowball effect that leads industrial firms to be locked in the same region for long periods: think of the Industrial Belt in the United States, the Pacific Industrial Belt in Japan, or the "Blue Banana" in Europe. This result also represents a neat formalization of an idea put forward by Myrdal (1957, 26–7) a long time ago:

> Within broad limits, the power of attraction of a centre has its origin mainly in the historical accident that something was once started there, and not in a number of other places where it could equally well have been started, and that start met success.

In such a context, it is well known that small initial differences between territories, minor changes in the socioeconomic environment, or both may eventually result in vastly different economic configurations.[5]

Note that the core–periphery model can be used to study the growing geographical concentration of business services in large metropolitan areas (Kolko 1999). In this context, the manufacturing sector is to be reinterpreted as the *service industry* in which firms not only supply consumers and manufacturing firms but also serve each other. This circularity in demand gives the corresponding firms incentives to agglomerate that are very similar to those studied by Krugman. As a result, once the different sectors are properly reinterpreted to account for the new trends emerging in developed economies, the core–periphery model will keep its relevance despite the gradual shift of manufacturers toward smaller cities or even rural areas. This tendency will be strengthened by the growing inclination of business services to work near the headquarters and research laboratories of manufacturing firms, which remain mostly located in large cities.

In Section 9.3, we formalize this idea by considering a large array of more or less tradable intermediate goods (and business services) as a major cause for agglomeration. So far, indeed, agglomeration has been considered as the outcome of a circular causation process fed by the mobility of workers. Yet, in the international marketplace, it is reasonable to expect this mobility to be low, and thus the core–periphery model cannot be used to explain the agglomeration

of industries in the world economy. Consequently, as argued by Venables (1996), dealing with the intermediate sector allows us to explain the possible emergence of a core–periphery structure at the international level. However, instead of following Venables (who assumes that both the intermediate and final sectors operate under increasing returns and monopolistic competition), we use a simpler framework. As in Section 4.2.1, we assume that the intermediate sector produces a differentiated good and exhibits increasing returns; however, the final sector produces a homogeneous good and exhibits constant returns. Finally, it is supposed that workers remain in their region. We then show that both sectors concentrate within the same region provided that the transport costs of the intermediate goods are sufficiently high (typically when they are nontradable). This is so even when the transport cost of the final good is very low. Indeed, the agglomeration of the intermediate sector firms makes it profitable for the final sector firms to agglomerate with them despite the wage gap generated by the immobility of workers.

In summary, the core–periphery structure may emerge owing to the migration of workers and the imperfectly competitive nature of the final sector or to the existence of an imperfectly competitive intermediate sector when workers are immobile. This result is very important for the space-economy, and thus it is crucial to know how it depends on the specificities of the framework employed. First, the use of the CES utility and iceberg cost leads to a convenient setting in which demands have a constant elasticity. However, such a result conflicts with research in spatial pricing theory in which demand elasticity varies with distance. Second, although the iceberg cost is able to capture the fact that shipping is resource-consuming, such a modeling strategy implies that any increase in the mill price is accompanied with a proportional increase in transport cost, which often seems unrealistic.[6] Finally, although models are based on very specific assumptions, they are often beyond the reach of analytical resolution, forcing authors to appeal to numerical investigations.[7] As recognized by Krugman (1998, 164) himself, "To date, the new economic geography has depended heavily on the tricks summarized in Fujita, Krugman, and Venables (1999) with the slogan 'Dixit–Stiglitz, iceberg, evolution and the computer.'"

This state of affairs has led Ottaviano and Thisse (1998) and Ottaviano, Tabuchi, and Thisse (2002) to revisit the core–periphery model using an alternative framework that involves downward-sloping linear demands and a linear transport cost measured in terms of the numéraire. Such a setting, which is very popular in location theory (Beckmann and Thisse 1986; Greenhut, Norman, and Hung 1987), takes us far away from the model used by Krugman and offers the advantage of yielding analytical solutions. Although the conclusions are not exactly the same as those derived by Krugman, this alternative model also yields a core–periphery structure once transportation costs are sufficiently low. Therefore, the core–periphery structure seems to be robust against very different formulations of preferences and transport technologies.

The linear model permits the study of different spatial price policies (Ottaviano and Thisse 1998). Because mill pricing yields the same qualitative results as in Section 9.2, we restrict ourselves to the case of discriminatory pricing in Section 9.4. This framework is very simple to use and is also suitable for studying the welfare properties of the core–periphery structure – an issue that has been untouched in most economic geography models. In Section 9.5, using the model of Section 9.4, we focus on the interplay between history and expectations in the formation of the economic space when migrants maximize the intertemporal value of their utility flows.

The work of Krugman has triggered a plethora of contributions, which have been surveyed by Ottaviano and Puga (1998). As noted earlier, the main result obtained by Krugman is the monotone relationship between the degree of agglomeration and the transportation cost level. In Section 9.6, the generality of such a relationship is discussed through several modifications of the basic model.

9.2 THE CORE–PERIPHERY MODEL

Although we focus on a two-region economy in this chapter, it will prove convenient to have a more general framework for subsequent developments.

9.2.1 The Framework

The economic space is made of R regions. The economy has two sectors: the modern sector ($\mathbb{M}$) and the traditional sector ($\mathbb{T}$). There are two production factors: the high-skilled workers and the low-skilled workers. The $\mathbb{M}$-sector produces a continuum of varieties of a horizontally differentiated product under increasing returns using skilled labor as the only input. The $\mathbb{T}$-sector produces a homogeneous good under constant returns using unskilled labor as the only input.

The economy is endowed with L unskilled workers and with H skilled workers. The skilled workers are perfectly mobile between regions, whereas the unskilled are immobile. As discussed in Section 8.2, this extreme assumption is partially justified because the skilled are more mobile than the unskilled. The share of unskilled workers in region r is fixed and denoted $0 \leq \nu_r \leq 1$ for $r = 1, \ldots, R$. The share of skilled workers in each region r is variable and denoted by $0 \leq \lambda_r \leq 1$ for $r = 1, \ldots, R$.

Although both consumption and production take place in a specific region, it is notationally convenient to describe preferences and technologies without explicitly referring to any particular region.

Preferences are identical across all workers and described by a Cobb–Douglas utility,

$$U = Q^{\mu} T^{1-\mu} / \mu^{\mu} (1-\mu)^{1-\mu} \qquad 0 < \mu < 1, \tag{9.1}$$

where Q stands for an index of the consumption of the modern sector varieties,

and T is the consumption of the output of the traditional sector. When the modern sector provides a continuum of varieties of size M, the index Q is given by

$$Q = \left[\int_0^M q(i)^\rho di \right]^{1/\rho} \qquad 0 < \rho < 1, \tag{9.2}$$

where $q(i)$ represents the consumption of variety $i \in [0, M]$. Hence, each consumer displays a preference for variety. In (9.2), the parameter ρ stands for the inverse of the intensity of desire for variety over the differentiated product. When ρ is close to 1, varieties are close to perfect substitutes; when ρ decreases, the desire to spread consumption over all varieties increases. If we set

$$\sigma \equiv \frac{1}{1-\rho},$$

then σ is the elasticity of substitution between any two varieties, which varies between 1 and ∞. Because there is a continuum of firms, each firm is negligible, and the interactions between any two firms are zero, but aggregate market conditions (e.g., the average price across firms) affect each firm. This provides a setting in which firms are not competitive (in the classic economic sense of having infinite demand elasticity), but at the same time they have no strategic interactions with one another (see (9.4) below).[8]

If Y denotes the consumer income, $p^{\mathbb{T}}$ the price of the traditional good, and $p(i)$ the price of variety i, then the demand functions are

$$T = (1-\mu)Y/p^{\mathbb{T}} \tag{9.3}$$

$$q(i) = \frac{\mu Y}{p(i)} \frac{p(i)^{-(\sigma-1)}}{P^{-(\sigma-1)}} = \mu Y p(i)^{-\sigma} P^{\sigma-1} \qquad i \in [0, M], \tag{9.4}$$

where P is the price index of the differentiated product given by

$$P \equiv \left[\int_0^M p(i)^{-(\sigma-1)} di \right]^{-1/(\sigma-1)}. \tag{9.5}$$

Introducing (9.3) and (9.4) into (9.1) yields the indirect utility function

$$v = Y P^{-\mu} (p^{\mathbb{T}})^{-(1-\mu)}. \tag{9.6}$$

The technology in the $\mathbb{T}$-sector is such that one unit of output requires one unit of L. Each variety of the $\mathbb{M}$-sector is produced according to the same technology such that the production of the quantity $q(i)$ requires $l(i)$ units of skilled labor given by

$$l(i) = f + cq(i), \tag{9.7}$$

where f and c are, respectively, the fixed and marginal labor requirements. Clearly, this technology exhibits scale economies. Without loss of generality, we choose the unit of skilled labor such that $c = 1$. Because preferences exhibit love for diversity and there are increasing returns but no scope economies, each variety is produced by a single firm. Indeed, any firm obtains a higher share of the market by producing a differentiated variety than by replicating an existing one.[9] In turn, this implies that the mass of firms is identical to the mass of varieties and that the output of a firm equals the demand of the corresponding variety.

The output of the $\mathbb{T}$-sector is costlessly traded between any two regions and is chosen as the numéraire, and thus $p^{\mathbb{T}} = 1$. In contrast, the output of the M-sector is shipped at a positive cost according to the "iceberg" technology: when one unit of the differentiated product is moved from region r to region s, only a fraction $1/\Upsilon_{rs}$ arrives at destination, where $\Upsilon_{rs} > 1$ for $r \neq s$ and $\Upsilon_{rr} = 1$. Hence, if variety i is produced in region r and sold at the mill (fob) price $p_r(i)$, the price $p_{rs}(i)$ paid by a consumer located in region s ($\neq r$) is

$$p_{rs}(i) = p_r(i)\Upsilon_{rs}. \tag{9.8}$$

If the distribution of firms is $(M_1, \ldots, M_R)$, using (9.5) and setting $\Upsilon_{rr} = 1$, we obtain the price index P_r in region r from

$$P_r = \left\{ \sum_{s=1}^{R} \Upsilon_{sr}^{-(\sigma-1)} \int_0^{M_s} p_s(i)^{-(\sigma-1)} di \right\}^{-1/(\sigma-1)}. \tag{9.9}$$

Let w_r denote the wage rate of a skilled worker living in region r. Because the price of the traditional good equals 1, the wage of the unskilled workers is also equal to 1 in all regions. Thus, because there is free entry and exit, and therefore zero profit in equilibrium, the income of region r is

$$Y_r = \lambda_r H w_r + \nu_r L. \tag{9.10}$$

From (9.4), the total demand of the firm producing variety i and located in region r is

$$\begin{aligned} q_r(i) &= \sum_{s=1}^{R} \mu Y_s [p_r(i)\Upsilon_{rs}]^{-\sigma} (P_s)^{\sigma-1} \Upsilon_{rs} \\ &= \mu p_r(i)^{-\sigma} \sum_{s=1}^{R} Y_s \Upsilon_{rs}^{-(\sigma-1)} (P_s)^{\sigma-1}. \end{aligned} \tag{9.11}$$

This expression requires some comment. The term $\mu Y_s [p_r(i)\Upsilon_{rs}]^{-\sigma} (P_s)^{\sigma-1} \Upsilon_{rs}$ stands for the quantity shipped from the firm located in r to region s. Here, the regional consumption in s, which is equal to $\mu Y_s [p_r(i)\Upsilon_{rs}]^{-\sigma} (P_s)^{\sigma-1}$, must be multiplied by Υ_{rs} because the firm's output "melts" on the way, thus implying

that the firm must send out a larger quantity of its output for the desired quantity to be delivered.[10]

Because each firm has a negligible impact on the market, it may accurately neglect the impact of a price change over consumers' income (Y_r) and other firms' prices and hence on the regional price indices (P_r).[11] Consequently, (9.11) implies that, regardless of the spatial distribution of consumers, each firm faces an isoelastic downward-sloping demand (the elasticity equals σ). This very convenient property depends crucially on the assumption of an iceberg transport cost, which affects the level of demand but not its elasticity.

The profit function of a firm in r is

$$\pi_r(i) = p_r(i)q_r(i) - w_r[f + q_r(i)] = [p_r(i) - w_r]q_r(i) - w_r f. \quad (9.12)$$

Because varieties are equally weighted in the utility function, the equilibrium price is the same across all firms located in region r. Solving the first-order condition using (9.11) yields the common equilibrium price

$$p_r^* = \frac{w_r}{\rho} \quad r = 1, \ldots, R. \quad (9.13)$$

This means that firms uses a relative markup equal to $1/\rho$, which is independent of the firms' and consumers' distributions. Everything else being equal, more product differentiation leads to a higher markup and, therefore, to a higher equilibrium price. However, the equilibrium price depends on the mass of firms and workers established in region r through the local wage w_r.

Substituting (9.13) into the profit function leads to

$$\pi_r = \frac{w_r}{\sigma - 1} q_r - w_r f = \frac{w_r}{\sigma - 1}[q_r - (\sigma - 1)f]. \quad (9.14)$$

Under free entry, profits are zero, and thus the equilibrium output of a firm is a constant given by

$$q_r^* = (\sigma - 1)f \qquad r = 1, \ldots, R. \quad (9.15)$$

Note that this quantity is independent of the distributions of firms and workers and is the same across regions. As a result, in equilibrium a firm's labor requirement is also unrelated to the firms' distribution:

$$l^* = \sigma f \qquad r = 1, \ldots, R.$$

Thus, the total mass of firms in the M-sector is constant and equal to H/l^*, whereas the corresponding firm distribution

$$M_r = \lambda_r H/l^* = \lambda_r H/\sigma f \qquad r = 1, \ldots, R \quad (9.16)$$

depends only on the distribution of skilled workers. These equalities imply that the core–periphery model allows for the spatial redistribution of the modern

sector but not for its growth, for the total number of firms (or varieties) is constant; this issue is addressed in Chapter 11.

Introducing the equilibrium prices (9.13) and substituting (9.16) for M_r in the regional price index (9.9), we obtain

$$P_r = \left[\sum_{s=1}^{R} \frac{\lambda_s H}{\sigma f}\left(\frac{w_s}{\rho}\Upsilon_{sr}\right)^{-(\sigma-1)}\right]^{-1/(\sigma-1)}$$

$$= \kappa_1\left[\sum_{s=1}^{R} \lambda_s (w_s \Upsilon_{sr})^{-(\sigma-1)}\right]^{-1/(\sigma-1)} \qquad r = 1, \ldots, R, \tag{9.17}$$

where

$$\kappa_1 \equiv \rho^{-1}\left(\frac{H}{\sigma f}\right)^{-1/(\sigma-1)},$$

which clearly depends on the spatial distribution of skilled workers as well as on the values of transport costs.

Finally, we consider the labor market clearing conditions for a given distribution of workers. The wage prevailing in region r is the highest wage that firms located there can pay under the nonnegative profit constraint. For that, we evaluate the demand (9.11) as a function of the wage through the equilibrium price (9.13):

$$q_r(w_r) = \mu\left(\frac{1}{\rho}\right)^{-\sigma} w_r^{-\sigma} \sum_{s=1}^{R} Y_s \Upsilon_{rs}^{-(\sigma-1)} P_s^{\sigma-1}. \tag{9.18}$$

Because this expression is equal to $(\sigma - 1)f$ when profits are zero, we obtain the following implicit expression for the *zero-profit wages*:

$$w_r^* = \kappa_2\left[\sum_{s=1}^{R} Y_s \Upsilon_{rs}^{-(\sigma-1)} P_s^{\sigma-1}\right]^{1/\sigma} \qquad r = 1, \ldots, R, \tag{9.19}$$

where

$$\kappa_2 \equiv \rho[\mu/(\sigma - 1)f]^{1/\sigma}.$$

Clearly, w_r^* is the equilibrium wage prevailing in region r when $\lambda_r > 0$.

Substituting (9.19) for Y and setting $p^{\mathrm{T}} = 1$ in the indirect utility (9.6), we obtain the *real wage* in region r as follows:

$$v_r = \omega_r = \frac{w_r^*}{P_r^{\mu}} \quad r = 1, \ldots, R. \tag{9.20}$$

Hence, the indirect utility is here equivalent to maximizing the real wage.

Finally, the Walras law implies that the traditional sector market is also in equilibrium provided that the equilibrium conditions above are satisfied.

For a given spatial distribution of skilled workers, we now ask whether there is an incentive for them to migrate and, if so, what direction the flow of migrants will take. A *spatial equilibrium* arises when no skilled worker may get a higher utility level in another region: $(\lambda_1^*, \ldots, \lambda_R^*)$ is a spatial equilibrium if there exists a positive constant ω^* such that

$$\omega_r \leq \omega^* \quad \text{for } r = 1, \ldots, R$$
$$\omega_r = \omega^* \quad \text{if } \lambda_r^* > 0.$$

Hence, the zero-profit real wage that local firms could afford to pay in a region containing no skilled workers is lower than (or just equal to) the equilibrium real wage. Because the functions $\omega_r(\lambda_1, \ldots, \lambda_R)$ are continuous in $(\lambda_1, \ldots, \lambda_R)$ over the compact set

$$\Lambda \equiv \left\{ (\lambda_1, \ldots, \lambda_R); \sum_{r=1}^{R} \lambda_r = 1 \quad \text{and} \quad \lambda_r \geq 0 \right\},$$

we can appeal to Proposition 1 of Ginsburgh, Papageorgiou, and Thisse (1985) to guarantee that such an equilibrium always exists.

Following a now well-established tradition in migration modeling, we focus on an adjustment process in which workers are attracted (repulsed) by regions providing high (low) utility levels:

$$\dot{\lambda}_r = \lambda_r(\omega_r - \bar{\omega}) \qquad r = 1, \ldots, R,$$

where $\dot{\lambda}_r$ is the time-derivative of λ_r, ω_r is the equilibrium real wage corresponding to the distribution $(\lambda_1, \ldots, \lambda_R)$, and $\bar{\omega} \equiv \sum \lambda_s \omega_s$ is the average real wage across all regions. In other words, the skilled move from the low-wage regions toward the high-wage ones.

A spatial equilibrium is stable if, for any marginal deviation of the population distribution from the equilibrium, the equation of motion above brings the distribution of skilled workers back to the original one. In doing so, we assume that local labor markets adjust instantaneously when some skilled workers move from one region to another. More precisely, the mass of firms in each region must be such that the labor market clearing conditions (9.16) remain valid for the new distribution of workers. Wages are then adjusted in each region for each firm to earn zero profits in any region having skilled workers because workers move toward high-wage regions.

Observe here one more justification for working with a continuum of agents (workers and firms): this modeling strategy allows one to respect the integer nature of a worker's or firm's location while describing the evolution of the regional share of production by means of differential equations.

9.2.2 The Two-Region Case

Consider two regions A and B. The unskilled workers are equally split between regions ($\nu_A = \nu_B = 1/2$). To keep things as symmetric as possible, we also assume that $\Upsilon_{AB} = \Upsilon_{BA} \equiv \Upsilon$. In this specific context, the basic equations developed in the foregoing are as follows:

$$Y_r = \lambda_r H w_r + L/2 \qquad r = A, B \tag{9.21}$$

$$P_r = \kappa_1 \left[\lambda_r w_r^{-(\sigma-1)} + \lambda_s (w_s \Upsilon)^{-(\sigma-1)}\right]^{-1/(\sigma-1)} \qquad s \neq r \tag{9.22}$$

$$w_r^* = \kappa_2 \left(Y_r P_r^{\sigma-1} + Y_s \Upsilon^{-(\sigma-1)} P_s^{\sigma-1}\right)^{1/\sigma} \qquad s \neq r \tag{9.23}$$

$$\omega_r = w_r^* P_r^{-\mu} \qquad r = A, B. \tag{9.24}$$

Whenever this turns out to be convenient, from now we use $\lambda \equiv \lambda_A$ so that $\lambda_B = 1 - \lambda$. Given a parametric solution to the system (9.21)–(9.24), a *spatial equilibrium* arises at $\lambda \in (0, 1)$ when

$$\Delta\omega(\lambda) \equiv \omega_A(\lambda) - \omega_B(\lambda) = 0$$

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

The stability is studied with respect to the following equation of motion:

$$\dot{\lambda} = \lambda \Delta\omega(\lambda)(1 - \lambda) \tag{9.25}$$

and is defined as in Section 8.4.[12] If $\Delta\omega(\lambda)$ is positive and $\lambda \in (0, 1)$, workers move from B to A; if it is negative, they go in the opposite direction. Clearly, any spatial equilibrium is a steady-state for (9.25).

The system (9.21)–(9.24) of nonlinear equations cannot be solved analytically. As a consequence, deriving a characterization of its solution in terms of λ is not simple. To derive some insight into the nature of the equilibrium, computational experiments have been performed by Krugman (1991a).[13] The results are displayed in Figure 9.1, where the following results appear. For a large value of Υ ($= \Upsilon_1$), there is only one equilibrium corresponding to the full dispersion of the modern sector ($\lambda = 1/2$), which is stable. When Υ takes some intermediate value Υ_2, four more equilibria emerge that are all asymmetric. However, the two interior equilibria are unstable. Hence, three stable equilibria now exist: the symmetric configuration and the core–periphery structure with concentration of the modern sector in region A or region B. Finally, when Υ takes a sufficiently low value ($= \Upsilon_3$), the symmetric equilibrium becomes unstable, and thus the core–periphery structure is the only stable outcome.

These observations will serve as a guide in the rest of the analysis.

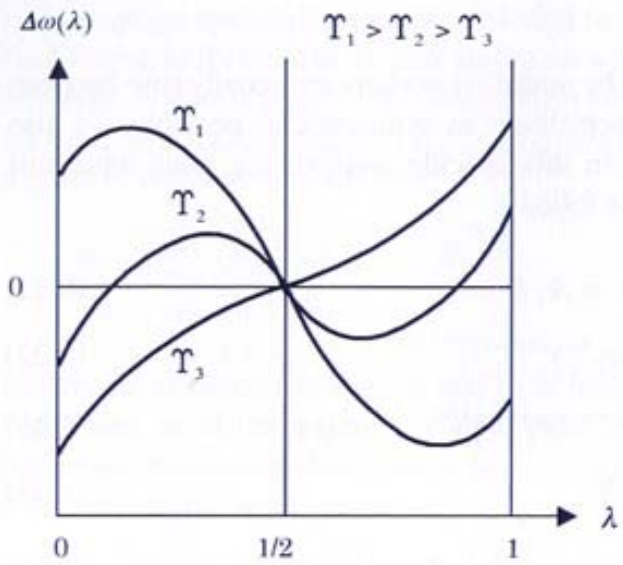


Figure 9.1: Migration dynamics under various values of Υ.

9.2.3 The Core–Periphery Structure

Suppose that the modern sector is concentrated in one region, say region A so that $\lambda = 1$. To check whether this is an equilibrium, we ask whether a skilled worker could be strictly better off in B. More precisely, we wish to determine conditions under which the real wage he may obtain in region B does not exceed the real wage this worker gets in region A. Setting $\lambda = 1$ in (9.21)–(9.24), we get the following equations:

$$\begin{aligned} Y_A &= Hw_A^* + L/2 \quad \text{and} \quad Y_B = L/2 \\ P_A &= \kappa_1 w_A^* \quad \text{and} \quad P_B = \kappa_1 \Upsilon w_A^*. \end{aligned} \tag{9.26}$$

Then, w_A^* is obtained by substituting (9.16) into (9.23) with $r = A$,

$$w_A^* = \kappa_2 \left[Y_A (\kappa_1 w_A^*)^{\sigma-1} + Y_B \Upsilon^{-(\sigma-1)} (\kappa_1 \Upsilon w_A^*)^{\sigma-1} \right]^{1/\sigma}$$

which yields $w_A^* = (\mu/H)(Y_A + Y_B)$, or

$$w_A^* = \frac{\mu}{1-\mu} \frac{L}{H}.$$

From (9.13), it is then possible to determine the common equilibrium price of all varieties in terms of the fundamentals of the economy:

$$p_A^* = \frac{1}{\rho} \frac{\mu}{1-\mu} \frac{L}{H},$$

which shows that the price of the differentiated product within the agglomeration increases with the unskilled–skilled ratio (L/H) as well as with the share of the modern sector (μ).

Finally, when the modern sector is geographically concentrated in A, the regional nominal incomes are as follows:

$$Y_A = \frac{\mu}{1-\mu}L + \frac{L}{2} \quad \text{and} \quad Y_B = \frac{L}{2},$$

and thus the gross domestic product (GDP) of the economy is given by

$$Y_G \equiv Y_A + Y_B = L/(1-\mu).$$

The equilibrium real wage in region A is

$$\omega_A = \kappa_1^{-\mu}(w_A^*)^{1-\mu} = \left(\frac{1}{\rho}\right)^{-\mu}\left(\frac{H}{\sigma f}\right)^{\mu/(\sigma-1)}\left(\frac{\mu}{1-\mu}\frac{L}{H}\right)^{1-\mu},$$

which is independent of Υ.

Agglomeration in region A is an equilibrium if and only if ω_A is larger than or equal to ω_B. Thus, we need to determine ω_B. To find it, we substitute (9.22) for the price index and (9.23) for the nominal wage into the real wage (9.24) and get

$$\begin{aligned}\omega_B &= \kappa_1^{\rho-\mu}\kappa_2(w_A^*)^{\rho-\mu}\Upsilon^{-\mu}\left(Y_A\Upsilon^{-(\sigma-1)} + Y_B\Upsilon^{\sigma-1}\right)^{1/\sigma} \\ &= \kappa_1^{\rho-\mu}\kappa_2(w_A^*)^{\rho-\mu}\Upsilon^{-\mu}\left(Y_A\Upsilon^{-(\sigma-1)} + Y_B\Upsilon^{\sigma-1}\right)^{1/\sigma}.\end{aligned}$$

It can then readily be verified that

$$\frac{\omega_B}{\omega_A} = \left[\frac{1+\mu}{2}\Upsilon^{-\sigma(\mu+\rho)} + \frac{1-\mu}{2}\Upsilon^{-\sigma(\mu-\rho)}\right]^{1/\sigma}. \tag{9.27}$$

When shipping is costless ($\Upsilon = 1$), we always have $\omega_B/\omega_A = 1$: location does not matter. Furthermore, the first term in the right-hand side of (9.27) is always decreasing in Υ. Therefore, because the second term is also decreasing when $\mu \geq \rho$, the ratio ω_B/ω_A always decreases with Υ, thus implying that $\omega_B < \omega_A$ for all $\Upsilon > 1$. This means that the core–periphery structure is a stable equilibrium for all $\Upsilon > 1$. When

$$\mu \geq \rho,$$

which is called the *black hole condition*, varieties are so differentiated that firms' demands are not very sensitive to differences in transportation costs, thus making the agglomeration force very strong. In fact, it is so strong that agglomeration can be viewed as a "black hole" attracting any movable activity.

More interesting is the case in which

$$\mu < \rho, \tag{9.28}$$

that is, varieties are not very differentiated so that the firms' demands are sufficiently elastic and, hence, the agglomeration force is weaker. If (9.28)

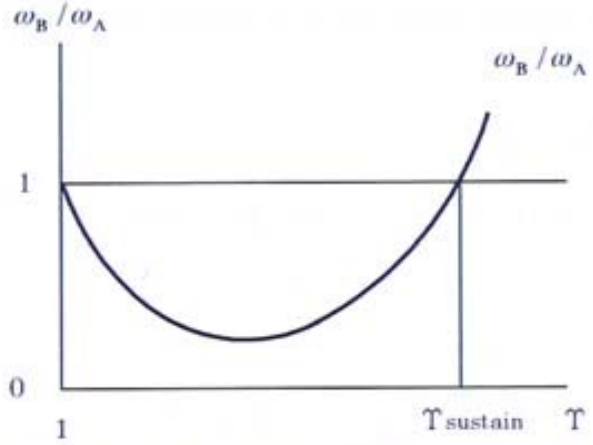


Figure 9.2: The determination of the sustain point.

holds, then the second term in (9.27) goes to infinity when $\Upsilon \to \infty$, and the ratio ω_B/ω_A is as depicted in Figure 9.2.[14]

We see that a single value $\Upsilon_{sustain} > 1$ exists such that $\omega_B/\omega_A = 1$. Hence, the agglomeration is a stable equilibrium for any $\Upsilon \leq \Upsilon_{sustain}$. In other words, once all firms belonging to the modern sector locate together within a region, they stay there as long as carrying their output to the other region is sufficiently cheap.[15] This occurs because firms can enjoy all the benefits of agglomeration without losing much of their business in the other region. Such a point is called the *sustain point* because, once firms are fully agglomerated, they stay so for all smaller values of Υ.[16] On the other hand, when transportation costs are sufficiently high ($\Upsilon > \Upsilon_{sustain}$), firms lose much on their exports so that the core–periphery structure is no longer an equilibrium.

Summarizing these results, we have the following:

Proposition 9.1 *Consider a two-region economy.*

1. *If $\mu \geq \rho$, then the core–periphery structure is always a stable equilibrium.*
2. *If $\mu < \rho$, then a unique solution $\Upsilon_{sustain} > 1$ exists to the equation*

$$\frac{1+\mu}{2}\Upsilon^{-\sigma(\mu+\rho)} + \frac{1-\mu}{2}\Upsilon^{-\sigma(\mu-\rho)} = 1 \tag{9.29}$$

such that the core–periphery structure is a stable equilibrium for any $\Upsilon \leq \Upsilon_{sustain}$.

It is remarkable that $\Upsilon_{sustain}$ depends only on the degree of product differentiation (σ) and the share of the modern sector in consumption (μ).

Interestingly, Proposition 9.1 provides formal support of the claim made by Kaldor (1970, 241) more than 30 years ago:

> When trade is opened up between them, the region with the more developed industry will be able to supply the need of the agricultural area of the other region on more favourable

terms: with the result that the industrial centre of the second region will lose its market and will tend to be eliminated.

The proposition also supports the claim of Giersch (1949, 94), who observed more than half a century ago that

production would tend to be centered in those industrial countries which already provide large domestic markets before the formation of the federal state.

It is worth stressing that the agglomeration is obtained as the aggregate outcome of a handful of individual decisions: the skilled workers do not choose a priori to be (or not to be) together. They are brought together through individual decisions based on current market prices and wages.

9.2.4 The Symmetric Structure

What we have just seen suggests that the modern sector is geographically dispersed when transportation costs are high and when (9.28) holds. To check this conjecture, we consider the symmetric configuration ($\lambda = 1/2$). In this case, there are only four equilibrium conditions,

$$Y_A = Y_B = Y = (H/2)w^* + L/2,$$

where w^* is the common zero-profit wage prevailing at the symmetric configuration given by

$$\begin{aligned} w^* &= \kappa_2\left(Y P^{\sigma-1} + Y\Upsilon^{-(\sigma-1)} P^{\sigma-1}\right)^{1/\sigma} \\ &= \kappa_2 (Y P^{\sigma-1})^{1/\sigma}\left(1 + \Upsilon^{-(\sigma-1)}\right)^{1/\sigma}, \end{aligned}$$

and the common price index is equal to

$$\begin{aligned} P &= \kappa_1\left[\frac{1}{2}(w^*)^{-(\sigma-1)} + \frac{1}{2}(w^*\Upsilon)^{-(\sigma-1)}\right]^{-1/(\sigma-1)} \\ &= \kappa_1 2^{1/(\sigma-1)} w^*\left(1 + \Upsilon^{-(\sigma-1)}\right)^{-1/(\sigma-1)}. \end{aligned}$$

The common real wage is

$$\omega = w^* P^{-\mu}.$$

Because $\omega_A = \omega_B = \omega$, the symmetric structure is a spatial equilibrium for all $\Upsilon > 1$.

For a given $\Upsilon > 1$, the symmetric equilibrium is stable (unstable) if the slope of $\Delta\omega(\lambda)$ is negative (positive) at $\lambda = 1/2$. Checking this condition requires fairly long calculations using all the equilibrium conditions. However, Fujita, Krugman, and Venables (1999, chap. 5) have shown the following results. First, when (9.28) does not hold, the symmetric equilibrium is always unstable. However, when (9.28) holds, this equilibrium is stable (unstable) if Υ

is larger (smaller) than some threshold value Υ_{break} given by

$$\Upsilon_{break} = \left[\frac{(\rho+\mu)(1+\mu)}{(\rho-\mu)(1-\mu)}\right]^{1/(\sigma-1)}, \tag{9.30}$$

which is clearly larger than 1. This is called the *break point* because symmetry between the two regions is no longer a stable equilibrium for lower values of Υ. It is interesting to note that Υ_{break} depends on the same parameters as $\Upsilon_{sustain}$. It is apparent from (9.30) that Υ_{break} is increasing with the share of the modern sector (μ) and with the degree of product differentiation ($1/\rho$).

Figure 9.3 represents all stable (unstable) equilibria by solid (broken) lines. It is shown in the appendix of this chapter that $\Upsilon_{break} < \Upsilon_{sustain}$. Hence, a domain of transport cost values exists over which there is multiplicity of equilibria, namely agglomeration and dispersion. More precisely, for $\Upsilon > \Upsilon_{sustain}$, the economy necessarily involves full dispersion. For $\Upsilon < \Upsilon_{break}$, the core–periphery structure always arises, the winning region depending on the initial conditions: the region with the initially larger share of the modern sector ends up with the whole share. Finally, for $\Upsilon_{break} \leq \Upsilon \leq \Upsilon_{sustain}$, both full agglomeration and full dispersion are stable equilibria. In the corresponding domain, the economy displays some hysteresis because full dispersion still prevails when transport costs fall below the sustain point while staying above the break point. Eventually, as transportation costs decrease sufficiently, sudden agglomeration of the modern sector arises.

9.2.5 A Model of the Modern Sector with Two Factors

It is reasonable to view the factor standing behind the fixed costs as skilled labor, especially when these costs correspond to research and development as well as advertising and sales promotion. Furthermore, Ottaviano (2001) has argued convincingly that a firm belonging to the modern sector could well use

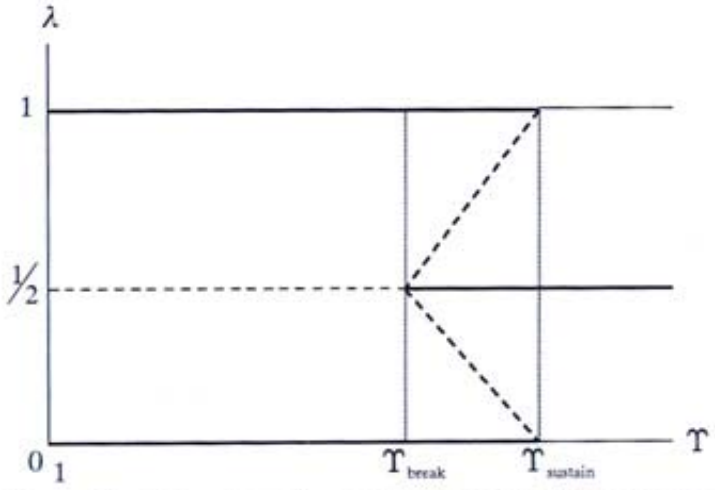


Figure 9.3: Bifurcation diagram for the core–periphery model.

unskilled labor to produce its output. Formally, this means that the production of any variety requires a fixed amount f of skilled labor and a marginal requirement $c = 1$ of unskilled labor. In this case, the marginal labor requirement is measured in terms of the numéraire. This vastly simplifies the analysis and allows for a more detailed treatment of the core–periphery model. Indeed, (9.12) becomes

$$\pi_r(i) = p_r(i)q_r(i) - w_r f - q_r(i) = [p_r(i) - 1]q_r(i) - w_r f \qquad r = A, B.$$

Because demands (9.4) are symmetric and isoelastic and the marginal cost is 1 instead of w_r, the equilibrium price is now the same across firms *and* regions:

$$p_A^* = p_B^* = \frac{1}{\rho}.$$

Thus, the delivered price Υ/ρ is also the same in both regions. Hence, the regional price indices are as follows:

$$P_r = \left[\lambda_r(H/f)(1/\rho)^{1-\sigma} + \lambda_s(H/f)(\Upsilon/\rho)^{1-\sigma}\right]^{1/(1-\sigma)} \qquad s \neq r$$

or

$$P_r = \frac{(H/f)^{1/(1-\sigma)}}{\rho}[\lambda_r + \lambda_s(\Upsilon)^{1-\sigma}]^{1/(1-\sigma)}. \tag{9.31}$$

The zero-profit output of a firm located in region $r = A, B$ is now

$$q_r^* = (\sigma - 1)f w_r,$$

which, unlike (9.15), varies with the region where the firm is located.

From (9.4), the demands in region $s = A, B$ for any variety produced in region $r = A, B$ are given by

$$q_s(r) = \frac{p_{rs}^{-\sigma}}{M_r p_{rs}^{1-\sigma} + M_s p_{ss}^{1-\sigma}} \mu Y_s$$

in which the regional incomes Y_r are

$$Y_r = L/2 + \lambda_r H w_r \qquad r = A, B.$$

Consequently, the market-clearing conditions for the differentiated product are given by

$$q_r^* = q_r(i) + \Upsilon q_s(i) \qquad s \neq r$$

or

$$(\sigma - 1)w_r = \frac{\rho\mu}{\lambda_r H + \lambda_s H \Upsilon^{1-\sigma}}\left(\frac{L}{2} + \lambda_r H w_r\right) + \frac{\rho\mu\Upsilon^{1-\sigma}}{\lambda_r H \Upsilon^{1-\sigma} + \lambda_s H}\left(\frac{L}{2} + \lambda_s H w_s\right),$$

which are linear in the nominal wages w_r and w_s.

Given the solutions to these two equations, (9.20) and (9.31) imply that

$$\Omega(\lambda) \equiv \frac{\omega_A(\lambda)}{\omega_B(\lambda)} = \frac{w_A^*/P_A^\mu}{w_B^*/P_B^\mu} = \left[\frac{\lambda + \Upsilon^{1-\sigma}(1-\lambda)}{(1-\lambda) + \Upsilon^{1-\sigma}\lambda}\right]^{\frac{\mu}{\sigma-1}} \times \frac{2\Upsilon^{1-\sigma}(\sigma-1)\lambda + \left[(\sigma-\mu\rho-1) + \Upsilon^{2(1-\sigma)}(\sigma+\mu\rho-1)\right](1-\lambda)}{2\Upsilon^{1-\sigma}(\sigma-1)(1-\lambda) + \left[(\sigma-\mu\rho-1) + \Upsilon^{2(1-\sigma)}(\sigma+\mu\rho-1)\right]\lambda}.$$

Because

$$\Delta\omega(\lambda) = \omega_B(\lambda)[\Omega(\lambda) - 1]$$

we have

$$\omega(\lambda) \gtreqless 0 \quad \text{if and only if} \quad \Omega(\lambda) \gtreqless 1.$$

Thus, $\Omega(\lambda)$ allows us to obtain the indirect utility differential as an explicit function of the distribution of skilled workers, which is something that cannot be achieved in the original core–periphery model. As usual, because $\Omega(1/2) = 1$, $\lambda = 1/2$ is always a spatial equilibrium.

By studying $\Omega(\lambda)$ with respect to λ, one can show that the equilibrium configurations are similar to those obtained above. First, let us determine Υ_{break}, which is a hard task in the original model. Because $\Omega(1/2) = 1$, the symmetric configuration is stable for some Υ if and only if

$$\Omega'(1/2) < 0$$

for the corresponding value of Υ. Setting $\phi \equiv \Upsilon^{1-\sigma}$, we see that ϕ varies between 0 (prohibitive transportation costs) and 1 (zero transportation costs). It can readily be verified that

$$\Omega'(1/2) = 0$$

has two roots in ϕ:

$$\phi_1 = \frac{(\sigma-\mu-1)(\sigma-\mu)}{(\sigma+\mu-1)(\sigma+\mu)} \quad \text{and} \quad \phi_2 = 1$$

with $\phi_1 < \phi_2$. When $\mu \geq \sigma - 1$, $\phi_1 < 0$ so that the symmetric equilibrium is never stable. This means that this condition is comparable to the black hole condition $\mu \geq \rho \equiv (\sigma-1)/\sigma$ obtained in Section 9.2.3. When $\mu > \sigma - 1$, $\partial\Omega/\partial\lambda$ evaluated at $\lambda = 1/2$ and $\phi = 0$ is negative, and thus the symmetric configuration is stable for any $\phi < \phi_1$. In other words, the symmetric configuration is stable provided that Υ is larger than

$$\Upsilon_{\text{break}} = \left[\frac{(\sigma+\mu-1)(\sigma+\mu)}{(\sigma-\mu-1)(\sigma-\mu)}\right]^{1/(\sigma-1)} > 1, \tag{9.32}$$

which is the counterpart of (9.30). For $\Upsilon < \Upsilon_{break}$, the symmetric equilibrium ceases to be stable. Next, setting $\Omega(1) = 1$ yields

$$\frac{1+\mu/\sigma}{2}\Upsilon^{-(\sigma-1+\mu)} + \frac{1-\mu/\sigma}{2}\Upsilon^{\sigma-1-\mu} = 1, \tag{9.33}$$

which is the counterpart of (9.29). Solving this equation, we obtain the $\Upsilon_{sustain}$ such that the core–periphery structure is a stable equilibrium for any $\Upsilon \leq \Upsilon_{sustain}$.

9.3 STICKY LABOR AND REGIONAL SPECIALIZATION

In the core–periphery model, agglomeration is the outcome of a circular causation process in which more people concentrate within the same region because they love variety. However, if workers are sticky, no agglomeration can arise. Instead, each region specializes in the production of differentiated varieties on the basis of their initial endowments, and intraindustry trade occurs for all values of the transportation costs.

However, the agglomeration of industries is a pervasive phenomenon even when labor is immobile. An alternative way to explain the emergence of agglomeration is to recognize that the modern sector uses an array of differentiated intermediate goods. In this case, *the agglomeration of the modern sector in a particular region can occur because of the concentration of the intermediate industry in that region, and conversely*. In this section, we show how this process works in a context in which we combine elements belonging to the core–periphery model and to the one presented in Section 4.2.1.[17] The key issue in the approach followed here is how workers living in a given region allocate themselves between the different sectors of the regional economy assuming for simplicity that the intersectoral mobility is perfect. For each region, a given allocation of labor generates a certain wage through the labor market clearing condition. At the corresponding wages, firms choose to stay put or to delocate. In equilibrium, no firm of the modern or intermediate sector has an incentive to change location.

9.3.1 The Framework

The economy involves three sectors: the final, intermediate, and traditional ones. Because workers are spatially immobile, love for variety is no longer an agglomeration force, and thus it is convenient to assume that the output of the modern sector is homogeneous. Preferences are identical across all workers and described by the utility (9.1):

$$u = Q^{\mu}T^{1-\mu}/\mu^{\mu}(1-\mu)^{1-\mu} \qquad 0 < \mu < 1,$$

where Q now stands for the consumption of a homogeneous good produced by the modern sector, and T is the consumption of the traditional sector's output.

Workers being immobile, we may consider a single type of labor. Because the output is homogeneous, the M-sector is assumed to operate under constant returns to scale and perfect competition. The M-good is produced according to the production function:

$$X^{\mathrm{M}} = l^{1-\alpha} I^{\alpha} \qquad 0 < \alpha < 1, \tag{9.34}$$

where X^{M} is the output of the M-sector, l the amount of labor, and I represents an index of the consumption of the intermediate varieties defined by

$$I = \left\{ \int_0^M [q(i)]^{\rho} di \right\}^{1/\rho} \qquad 0 < \rho < 1 \tag{9.35}$$

in which $q(i)$ is the quantity of the intermediate good i and M the number of intermediate goods. As usual, a smaller ρ means a more differentiated set of intermediate varieties.

By contrast, the intermediate sector exhibits increasing returns and operates under monopolistic competition (as in Section 4.2.1). Each variety of the I-sector is produced according to the same technology such that the production of the quantity $q(i)$ requires $l(i)$ units of labor, which is again given by

$$l(i) = f + q(i), \tag{9.36}$$

where f is the fixed requirement of labor.

Finally, as in Section 9.2, the technology in the T-sector is such that one unit of output requires one unit of labor.

The demand functions for the two consumption goods are as follows:

$$T = (1 - \mu) Y / p^{\mathrm{T}} \tag{9.37}$$

$$Q = \mu Y / p^{\mathrm{M}}, \tag{9.38}$$

where p^{M} is the price of the M-good.

The price index for the I-sector is as given by (9.5):

$$P \equiv \left[\int_0^M p(i)^{-(\sigma-1)} di \right]^{-1/(\sigma-1)}, \tag{9.39}$$

where $\sigma \equiv 1/(1-\rho)$ and $p(i)$ is the price of the intermediate good i. Given the wage rate w, the unit production cost in the M-sector is

$$c^{\mathrm{M}} = \alpha^{-\alpha} (1-\alpha)^{-(1-\alpha)} w^{1-\alpha} P^{\alpha}, \tag{9.40}$$

whereas the input demands of the M-sector corresponding to output X^{M} are

$$L^{\mathrm{M}} = (1-\alpha) c^{\mathrm{M}} X^{\mathrm{M}} w^{-1} \tag{9.41}$$

$$q(i) = \alpha c^{\mathrm{M}} X^{\mathrm{M}} p(i)^{-\sigma} P^{\sigma-1} \qquad i \in [0, M]. \tag{9.42}$$

Consider an economy with two regions A and B, each endowed with $L > 0$ workers. As in the foregoing section, we assume that the $\mathbb{T}$-good can be shipped costlessly from one region to another; this good is used as the numéraire ($p^{\mathbb{T}} = 1$). The output of the $\mathbb{M}$-sector ($\mathbb{I}$-sector) is shipped from one region to the other at a positive cost according to the iceberg cost $\Upsilon^{\mathbb{M}} > 1$ ($\Upsilon^{\mathbb{T}} > 1$). Let $L_r^{\mathbb{M}}$, $L_r^{\mathbb{I}}$, and $L_r^{\mathbb{T}}$ be the mass of workers living in region r ($= A, B$) and working in the modern, intermediate, and traditional sectors, respectively.

Using the same argument as in Section 9.2.2, we see that the common equilibrium price of the intermediate varieties produced in region $r = A, B$ is

$$p_r^{\mathbb{I}} = \frac{w_r}{\rho},$$

and thus the output of an $\mathbb{I}$-sector firm under zero profit is still given by $(\sigma - 1)f$ and its labor requirement by σf.

The price index for the $\mathbb{I}$-goods in region $r = A, B$ can be shown to be equal to

$$P_r = k_1\left[L_r^{\mathbb{I}}(w_r)^{-(\sigma-1)} + L_s^{\mathbb{I}}(w_s\Upsilon^{\mathbb{I}})^{-(\sigma-1)}\right]^{-1/(\sigma-1)} \qquad s \neq r, \tag{9.43}$$

where

$$k_1 \equiv \rho^{-1}(\sigma f)^{1/(\sigma-1)}.$$

The common output of a $\mathbb{I}$-firm located in r is

$$q_r = \alpha\left(\frac{w_r}{\rho}\right)^{-\sigma}\left[c_r^{\mathbb{M}}X_r^{\mathbb{M}}P_r^{\sigma-1} + c_s^{\mathbb{M}}X_s^{\mathbb{M}}(\Upsilon^{\mathbb{I}})^{-(\sigma-1)}P_s^{\sigma-1}\right] \qquad s \neq r \tag{9.44}$$

where $X_r^{\mathbb{M}}$ is the production of the $\mathbb{M}$-good in region $r = A, B$.

9.3.2 Agglomeration of the Intermediate and Final Sectors

Suppose that both the final and intermediate industries are concentrated in one region, say region A, so that $L_B^{\mathbb{M}} = L_B^{\mathbb{I}} = 0$ and $L_B^{\mathbb{T}} = L$. Hence, region A exports the $\mathbb{M}$-good and region B the $\mathbb{T}$-good. Assume also that workers residing in region A work either for the intermediate or for the modern sector so that $L_A^{\mathbb{T}} = 0$. Then, it follows that $w_A^* \geq w_B^* = 1$.

The corresponding regional price indices are obtained from (9.43) as follows:

$$\begin{aligned} P_A &= k_1\left[L_A^{\mathbb{I}}(w_A^*)^{-(\sigma-1)}\right]^{-1/(\sigma-1)} \\ &= k_1\frac{\mu}{1-\mu}\left(L_A^{\mathbb{I}}\right)^{-1/(\sigma-1)} \end{aligned} \tag{9.45}$$

$$P_B = P_A\Upsilon^{\mathbb{I}}. \tag{9.46}$$

Consider first the equilibrium of the $\mathbb{T}$-good market. Because $Y_A = Lw_A^*$ and $Y_B = L$, regional demands are respectively given by

$$T_A = (1-\mu)Lw_A^* \quad \text{and} \quad T_B = (1-\mu)L.$$

Because $X_A^{\mathbb{T}} = 0$ and $X_B^{\mathbb{T}} = L$, the equality of supply and demand implies that

$$w_A^* = \frac{\mu}{1-\mu}. \tag{9.47}$$

For the traditional sector to be unprofitable in region A, it follows that $w_A^* \geq 1$, which holds if and only if

$$\mu \geq 1/2. \tag{9.48}$$

Because we have

$$p_A^{\mathbb{M}} = c_A^{\mathbb{M}} \quad \text{and} \quad p_B^{\mathbb{M}} = p_A^{\mathbb{M}}\Upsilon^{\mathbb{M}} = c_A^{\mathbb{M}}\Upsilon^{\mathbb{M}},$$

using (9.38), the regional demands for the $\mathbb{M}$-good are given by

$$Q_A = \mu L w_A^*/c_A^{\mathbb{M}} \quad \text{and} \quad Q_B = \mu L/c_A^{\mathbb{M}}\Upsilon^{\mathbb{M}}.$$

Because the $\mathbb{M}$-good is exported to B, the equilibrium of the $\mathbb{M}$-good market implies that the total production $X_A^{\mathbb{M}}$ is such that

$$\begin{aligned} X_A^{\mathbb{M}} &= Q_A + Q_B\Upsilon^{\mathbb{M}} \\ &= \frac{\mu}{1-\mu}\frac{L}{c_A^{\mathbb{M}}} \end{aligned}$$

from (9.47), or

$$X_A^{\mathbb{M}} c_A^{\mathbb{M}} = \frac{\mu}{1-\mu} L. \tag{9.49}$$

Producing the $\mathbb{M}$-good in region B is never profitable if and only if

$$c_B^{\mathbb{M}} \geq p_B^{\mathbb{M}} = c_A^{\mathbb{M}}\Upsilon^{\mathbb{M}},$$

which amounts to $c_B^{\mathbb{M}}/c_A^{\mathbb{M}} \geq \Upsilon^{\mathbb{M}}$. With (9.40), this holds if and only if

$$\Upsilon^{\mathbb{I}} \geq \left(\frac{\mu}{1-\mu}\right)^{(1-\alpha)/\alpha} (\Upsilon^{\mathbb{M}})^{1/\alpha}.$$

It remains to consider the $\mathbb{I}$-sector. Given (9.44), (9.49), and $X_B = 0$, we have

$$\begin{aligned} q_A &= \alpha\left(\frac{w_A}{\rho}\right)^{-\sigma}\left[c_A^{\mathbb{M}} X_A^{\mathbb{M}} P_A^{\sigma-1}\right] \\ &= \alpha\rho^{\sigma}(k_1)^{\sigma-1} L\left(L_A^{\mathbb{I}}\right)^{-1} \end{aligned}$$

$$\begin{aligned} q_B &= \alpha\left(\frac{1}{\rho}\right)^{-\sigma}\left[c_A^{\mathbb{M}} X_A^{\mathbb{M}} (\Upsilon^{\mathbb{I}})^{-(\sigma-1)} P_A^{\sigma-1}\right] \\ &= \rho^{\sigma} k_1^{\sigma-1}\left(\frac{\mu}{1-\mu}\right)^{\sigma} (\Upsilon^{\mathbb{I}})^{-(\sigma-1)}. \end{aligned}$$

The first equilibrium condition is $q_A = q^* = (\sigma - 1)f$, which yields

$$L_A^{\mathbb{I}} = \alpha L.$$

The second equilibrium condition, that is, the nonprofitability of region B for $\mathbb{I}$-firms, means $q_B \leq q^*$, which is equivalent to the condition

$$\Upsilon^{\mathbb{I}} \geq \left(\frac{\mu}{1-\mu}\right)^{1/\rho}.$$

To sum up, we have shown the following:

Proposition 9.2 *Assume $\mu \geq 1/2$. Then, the agglomeration of the intermediate and final sectors into the same region is an equilibrium if and only if the following two conditions are satisfied:*

$$\Upsilon^{\mathbb{I}} \geq \left(\frac{\mu}{1-\mu}\right)^{(1-\alpha)/\alpha} (\Upsilon^{\mathbb{M}})^{1/\alpha} \tag{9.50}$$

$$\Upsilon^{\mathbb{I}} \geq \left(\frac{\mu}{1-\mu}\right)^{1/\rho}. \tag{9.51}$$

Hence, when the transport cost of the intermediate goods is high relative to the transport cost of the final good, there is complete regional specialization because both the final and intermediate sectors are entirely concentrated in one region, whereas the traditional sector operates only in the other region. Condition (9.50) becomes less stringent as the transport cost of the final good declines. In addition, the transport cost of the intermediate goods must also exceed some threshold value (9.51), for $\mu/(1-\mu) \geq 1$. Clearly, this threshold rises when the intermediate goods are more differentiated. The domains of $(\Upsilon^{\mathbb{I}}, \Upsilon^{\mathbb{M}})$ sustaining the core–periphery structure are represented by the shaded areas in Figure 9.4.

Condition (9.50) means that the $\mathbb{M}$-sector does not find it profitable to start operating in region B because importing the intermediate goods from A turns out to be costly owing to the high transport costs of these goods, whereas exporting its output from A to B is less costly because of the relatively low value of $\Upsilon^{\mathbb{M}}$. Condition (9.51) means that no firm of the intermediate sector wants to set up in region B because it has to export all its production to region A at a high transport cost. It should be stressed that both sectors are trapped within the same region even when shipping the final good becomes cheaper and cheaper ($\Upsilon^{\mathbb{M}}$ approaches 1).

To break such a trap, the transport costs of the intermediate goods must fall below some critical value. This is not necessarily easy to implement when the provision of specific intermediate goods requires face-to-face contacts such as for highly differentiated services (in which case $\Upsilon^{\mathbb{I}}$ is high).

As long as (9.50) and (9.51) hold, μ can rise. Because $w_A^* = \mu/(1-\mu)$ and $w_B^* = 1$, this rise generates a widening wage gap between the core region

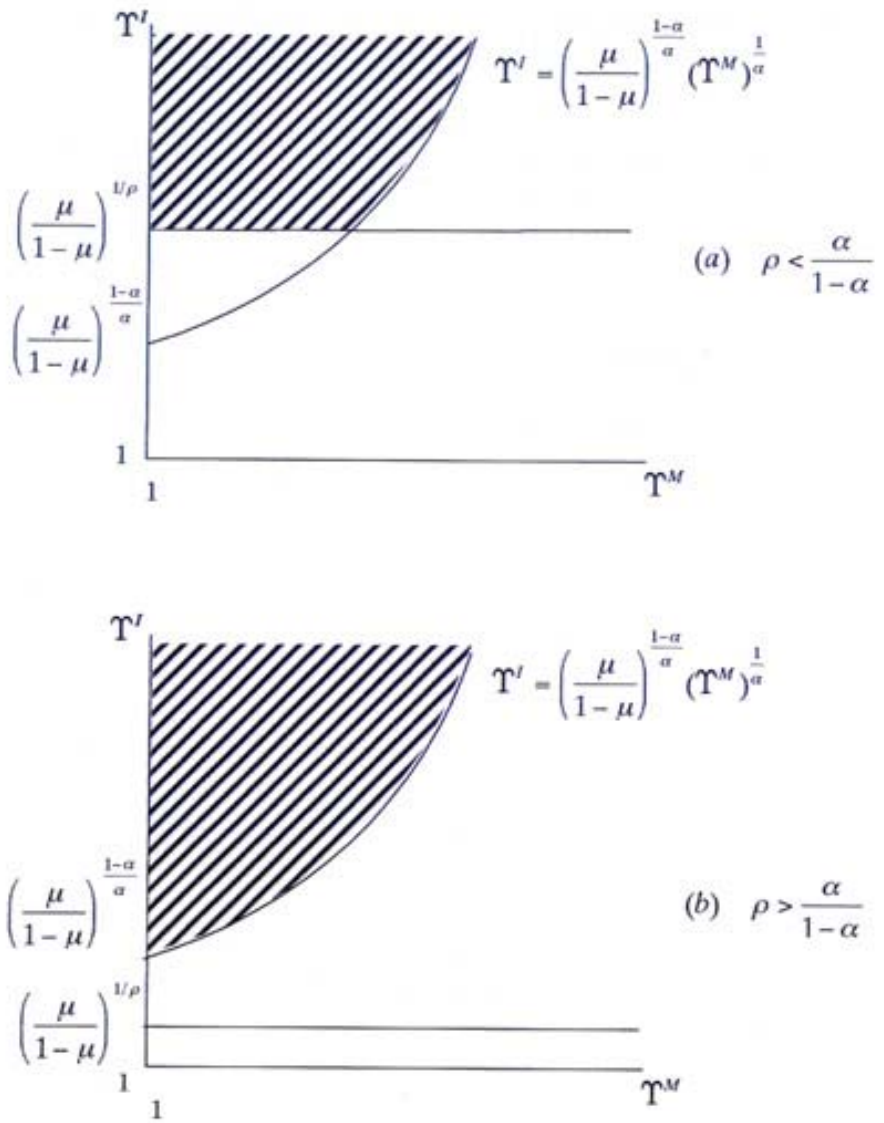


Figure 9.4: Transport cost ranges for sustaining the core–periphery structure.

and the periphery. It can readily be verified that the real wage gap, in turn, becomes even larger. This agrees with the observation that (9.50) becomes less and less stringent as the role of the intermediate goods plays a growing role in the economy (α rises). However, the modern sector will eventually decentralize some of its activities in the periphery as its share in consumption increases.

When transportation costs for the intermediate sector decrease enough for (9.50) not to hold anymore, whereas (9.51) is still valid, one expects the $\mathbb{I}$-firms to remain concentrated in region A, and some share of the final sector now operates in region B. In this case, the intermediate goods needed by the final sector in region B are imported from region A. Finally, when shipping the intermediate goods becomes very inexpensive ($\Upsilon^{\mathbb{I}}$ approaches 1), it is reasonable to expect the symmetric configuration to be the only stable equilibrium. These issues are left for further investigation.

9.4 A LINEAR MODEL OF CORE–PERIPHERY: DISCRIMINATORY PRICING AND WELFARE

The setting considered in this section is very similar to the one used in Section 9.2. However, there are two major differences. First, the output of the M-sector is traded at a cost of t units of the numéraire per unit shipped between regions. This characteristic agrees more with reality as well as with location theory than the iceberg technology does. In particular, as will be demonstrated, this allows us to study and compare various spatial price policies such as discriminatory and mill pricing. Second, preferences are given by a quasi-linear utility encapsulating a quadratic subutility instead of a Cobb–Douglas preference on the homogeneous and differentiated goods with CES subutility.[18] These two specifications correspond to two rather extreme cases: the former assumes an infinite elasticity of substitution between the differentiated product and the numéraire, the latter a unit elasticity. Moreover, firms' demands are linear instead of iso-elastic.

Specifically, we use the model described in Section 8.4, which is now embedded in a general equilibrium framework. In particular, preferences are identical across all workers and described by (8.26) with individual demand $q(i)$ for variety i given by

$$q(i) = a - (b + dM)\, p(i) + d\, P, \tag{9.52}$$

where $p(i)$ is the price of variety $i \in [0, M]$ and P/M the price index in the modern sector. Hence, each worker has the following quasi-linear indirect utility function:

$$v(y; p(i),\ i \in [0, M]) = \frac{a^2 M}{2b} - a \int_0^M p(i)di + \frac{b + dM}{2} \int_0^M [p(i)]^2 di - \frac{d}{2}\left[\int_0^M p(i)di\right]^2 + y + \bar{z}, \tag{9.53}$$

where y is the worker's labor income and $\bar{z}$ his initial endowment in the numéraire. It is assumed that each worker's initial endowment $\bar{z}$ in the numéraire is large enough for his residual consumption of the numéraire to be strictly positive in equilibrium.[19] The technologies are the same as in Section 9.2 but, for simplicity, c is set equal to zero in (9.7).[20]

As stated in the introduction, we focus on *discriminatory pricing*. Recall that, in this case, each firm sets a delivered price specific to the market in which its variety is sold; hence, markets are segmented. In addition, because the iceberg cost enters the profit function multiplicatively in Krugman's model, demands have the same elasticity across locations, and thus both mill and discriminatory pricing policies yield the same equilibrium prices and outputs. Few people noticed the equivalence between the two policies under the iceberg technology

in the spatialized S–D–S model (Hsu 1979; Greenhut et al. 1987, chap. 2). This equivalence no longer holds in the present model.

9.4.1 Equilibrium Prices

Suppose that each firm sets a (delivered) price specific to each region. Hence, the profit function of a representative firm located in region $r = A, B$ is as follows:

$$\pi_r = p_{rr} q_{rr}(p_{rr})(L/2 + \lambda_r H) + (p_{rs} - t) q_{rs}(p_{rs})(L/2 + \lambda_s H) - f w_r \qquad s \neq r, \tag{9.54}$$

where $\lambda \equiv \lambda_A$ and $\lambda_B = 1 - \lambda$.

To illustrate the type of interaction that characterizes this model of monopolistic competition, we describe how the equilibrium prices are determined. Each firm i in region r maximizes its profit π_r, assuming accurately that its price choice has no impact on the regional price indices

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

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

$$M_r p^*_{rr}(P_r) + M_s p^*_{sr}(P_r) = P_r \qquad s \neq r.$$

By solving the first-order conditions for profit maximization with respect to prices, it can readily be verified that the equilibrium prices are as follows:

$$p^*_{rr} = \frac{1}{2}\frac{2a + dt(1 - \lambda_r)M}{2b + dM} \qquad s \neq r \tag{9.55}$$

$$p^*_{rs} = p^*_{ss} + \frac{t}{2} \qquad s \neq r. \tag{9.56}$$

Clearly, these prices depend directly on the firms' (or workers') distribution across regions. More precisely, p^*_{rr} decreases with the mass of firms in region r, $\lambda_r M$, and increases with the degree of product differentiation, which decreases with d. These results agree with what we know from models of product differentiation (see, e.g., Anderson et al. 1992, chap.7). In addition, these prices also show that arbitrage is never profitable because $p^*_{rs} - p^*_{rr} < \tau$.

It is easy to check that the equilibrium operating profits earned in each market by a firm established in $r = A, B$ are as follows:

$$\pi^*_{rr} = (b + dM)(p^*_{rr})^2 (L/2 + \lambda_r H)$$

$$\pi^*_{rs} = (b + dM)(p^*_{rs} - t)^2 (L/2 + \lambda_s H),$$

where π^*_{rk} denotes the profits earned from selling in region $k = r, s$.

Increasing λ_r has two opposite effects on π^*_{rr}. First, as λ_r rises, both the equilibrium price (9.55) and the corresponding quantity of each variety bought by each consumer living in region r, fall. Second, the total population of consumers residing in this region is now larger, and thus the profits made by a firm located in r on local sales may increase. What is at work here is a global demand effect resulting from the increase in the local population that may compensate firms for the adverse price effect as well as for the decrease in each worker's individual demand.

Deducting t from (9.55) and (9.56), we see that firms' prices net of transport costs are positive regardless of the workers' distribution if and only if

$$t < t_{\text{trade}} \equiv \frac{2af}{2bf + dH} \equiv \frac{2\alpha f}{2f + \delta H/(\beta - \delta)}, \tag{9.57}$$

which depends only upon the primitives $(L, H, \alpha, \beta, \delta, f)$ once a, b, d, and M are replaced by their values. More generally, it can readily be verified that

$$\frac{d\tau_{\text{trade}}}{df} > 0 \qquad \frac{d\tau_{\text{trade}}}{d\delta} < 0;$$

thus, trade is more likely the higher are the intensity of increasing returns and the degree of product differentiation. The condition (9.57) must also hold for consumers in s to buy from firms in r, that is, for the individual demands (9.52) in each region evaluated at the equilibrium prices to be positive for all λ. Observe that there is no trade when there are no increasing returns ($f = 0$). In this case, each region supplies all the varieties and is in autarchy – a result that agrees with what we saw in Chapter 2.

The consumer surplus $C_r(\lambda)$ of each worker in region r associated with the equilibrium prices (9.55) and (9.56) is then as follows:

$$C_r(\lambda) = \frac{a^2 H}{2bf} - \frac{aH}{f}(\lambda_r p^*_{rr} + \lambda_s p^*_{sr}) + \frac{(bf + dH)H}{2f^2}\left[\lambda_r (p^*_{rr})^2 + \lambda_s (p^*_{sr})^2\right]$$
$$- \frac{dH^2}{2f^2}(\lambda_r p^*_{rr} + \lambda_s p^*_{sr}).$$

Differentiating this expression twice with respect to λ shows that $C_r(\lambda)$ is concave. Furthermore, (9.57) implies that $C_r(\lambda)$ is always increasing over the interval $[0, 1]$.

Labor market clearing implies that the mass M_r of firms belonging to the M-sector in region r is

$$M_r = \lambda_r H/f. \tag{9.58}$$

Consequently, the total mass of firms in the economy is fixed and equal to $M = M_A + M_B = H/f$. As in Section 9.2, the number of firms is constant. In addition, (9.58) shows that the region with the larger number of workers is also the region accommodating the larger number of firms.

Entry and exit are free; therefore, profits are zero in equilibrium. Hence, (9.58) implies that any change in the population of workers located in one region must be accompanied by a corresponding change in the number of firms. The equilibrium wage rates w_r^* of the skilled are obtained from the zero profit condition evaluated at the equilibrium prices, whereas the wage rate of the unskilled is equal to their marginal product, that is, 1.

More precisely, the equilibrium wage prevailing in region $r = A, B$ may be obtained from (9.54) by computing $w_r^*(\lambda) = (\pi_{rr}^* + \pi_{rs}^*)/f$. This yields the following expression:

$$w_r^*(\lambda) = \frac{bf + dH}{4(2bf + dH)^2 f^2}\left\{(2af + \tau dH\lambda_s)^2\left(\frac{L}{2} + \lambda_r H\right)\right.$$
$$\left. + (2af - 2\tau bf - \tau dH\lambda_s)^2\left(\frac{L}{2} + \lambda_s H\right)\right\}, \tag{9.59}$$

which, after simplification, turns out to be quadratic in λ. Standard but cumbersome investigations reveal that $w_r^*(\lambda)$ is concave and increasing (convex and decreasing) in λ when f is large (small) as well as when t, d, H, and L are small (large). This implies that both $C_r(\lambda)$ and $w_r^*(\lambda)$ increase with λ when t is small, whereas they go in opposite directions when t is large. This provides useful information for the study of the agglomeration process.

9.4.2 Agglomeration versus Dispersion

The indirect utility differential is obtained by plugging the equilibrium prices (9.55) and (9.56) and the equilibrium wages (9.59) into (9.53):

$$\Delta v(\lambda) \equiv v_A(\lambda) - v_B(\lambda) \equiv C_A(\lambda) - C_B(\lambda) + w_A^*(\lambda) - w_B^*(\lambda)$$
$$= C_M t(t^* - t)(\lambda - 1/2), \tag{9.60}$$

where

$$C_M \equiv [2bf(3bf + 3dH + dL) + d^2H(L + H)]\frac{H(bf + dH)}{2f^2(2bf + dH)^2} > 0$$

and

$$t^* \equiv \frac{4af(3bf + 2dH)}{2bf(3bf + 3dH + dL) + d^2H(L + H)}, \tag{9.61}$$

which can also be restated in terms of the primitives of the economy. The stability of an equilibrium is studied with respect to (9.25).

Observe that $\Delta v(\lambda)$ is linear in λ. It follows immediately from (9.60) that $\lambda = 1/2$ is always an equilibrium. Because $C_M > 0$, for $\lambda \neq 1/2$, the indirect utility differential always has the same sign as $\lambda - 1/2$ if and only if $t < t^*$; if $t > t^*$, it has the opposite sign. In particular, when there are no increasing returns in the manufacturing sector ($f = 0$), the coefficient of $(\lambda - 1/2)$ is always negative because $t^* = 0$; thus, symmetry is the only (stable) equilibrium.[21] This shows

once more the importance of increasing returns for the possible emergence of an agglomeration. The same holds for product differentiation because t^* becomes arbitrarily small when varieties become less and less differentiated ($d \to \infty$).[22]

It remains to determine when t^* is lower than t_{trade}, as given by (9.57). This is so if and only if

$$L/H > \frac{6b^2 f^2 + 8bdfH + 3d^2 H^2}{dH(2bf + dH)} > 3, \tag{9.62}$$

where the second inequality holds because $b/d = \beta/\delta - 1 \in (0, \infty)$. The inequality (9.62) means that the population of unskilled is large relative to the population of skilled. When (9.62) does not hold, the coefficient of $(\lambda - 1/2)$ in (9.60) is always positive for all $t < t_{\text{trade}}$. In this case, the advantages of having a large home market always dominate the disadvantages incurred while supplying a distant periphery. The reverse of (9.62) plays a role similar to the black hole condition described in Section 9.2.3.

More interesting is the case when condition (9.62) holds. Although the size of the industrial sector is captured here through the relative population size L/H and not through its share in consumption, the intuition is similar: the ratio L/H must be sufficiently large for the economy to display different types of equilibria according to the value of t. This result does not depend on the expenditure share on the manufacturing sector because general equilibrium income effects are absent: either small or large sectors in terms of expenditure share are agglomerated when t is small enough.

When $t > t^*$, it is straightforward to see that the symmetric configuration is the only stable equilibrium. In contrast, when $t < t^*$ the symmetric equilibrium becomes unstable and workers agglomerate in region A (B) provided that the initial fraction of workers residing in this region exceeds $1/2$. In other words, *agglomeration arises when transportation costs are sufficiently low*, as in Section 9.2.3 and for similar reasons. In addition, because the incentives to move keep rising with the size of the agglomeration, once made, no worker regrets his choice to live in the core region.

Proposition 9.3 *Consider a two-region economy with segmented markets.*

1. *When condition (9.62) does not hold, the core–periphery structure is the only stable equilibrium under trade.*
2. *When (9.62) is satisfied, for any $t > t^*$ the symmetric configuration is the only stable equilibrium with trade; for any $t < t^*$ the core–periphery pattern is the unique stable equilibrium; for $t = t^*$ any configuration is an equilibrium.*

As t decreases from some value larger than t^* but smaller than t_{trade}, we move from a stable dispersed equilibrium with two symmetric regions to a market situation in which any stable equilibrium involves a core–periphery pattern. Note that here the break point and the sustain point are the same and

that history alone matters for the selection of the agglomerated outcome; this follows because (9.60) is linear in λ.

Looking at the threshold value t^*, as given by (9.61), we first observe that it increases when the intensity of preference for the $\mathbb{M}$-good relative to the $\mathbb{T}$-good (as measured by a higher α or a) increases. Second, t^* increases with the degree of product differentiation (δ or d falls) when (9.62) holds.[23] Third, higher fixed costs lead to a smaller mass of firms and varieties. Still, it can readily be verified that t^* also increases when increasing returns become stronger (f rises) when (9.62) holds. In other words, the agglomeration of the modern sector is more likely the stronger are the increasing returns at the firm's level. Last, t^* increases when the mass of unskilled (L) decreases because the dispersion force is weaker.

All these results are similar to those obtained in Section 9.2, which suggests some robustness in the reasons for the core–periphery structure because the two models considered differ significantly.

Because the skilled are mobile, they always reach the same utility level, and thus agglomeration does not entail any redistribution within this group. However, agglomeration gives rise to redistributive effects inside the population of unskilled. Since, as seen above, the surplus C_A increases with λ, any increase from $\lambda = 1/2$ makes the unskilled in A better off and, by symmetry, the unskilled in B worse off.[24] Furthermore, given that C_A is concave in λ, the marginal gain of the former is lower than the marginal loss of the latter: agglomeration hurts the unskilled in the periphery more than it helps those living in the core region.

9.4.3 A Diagrammatic Analysis

It is possible to convey the economic intuition behind Proposition 9.3 through a simple graphical analysis. Figure 9.5 depicts the aggregate inverse demand

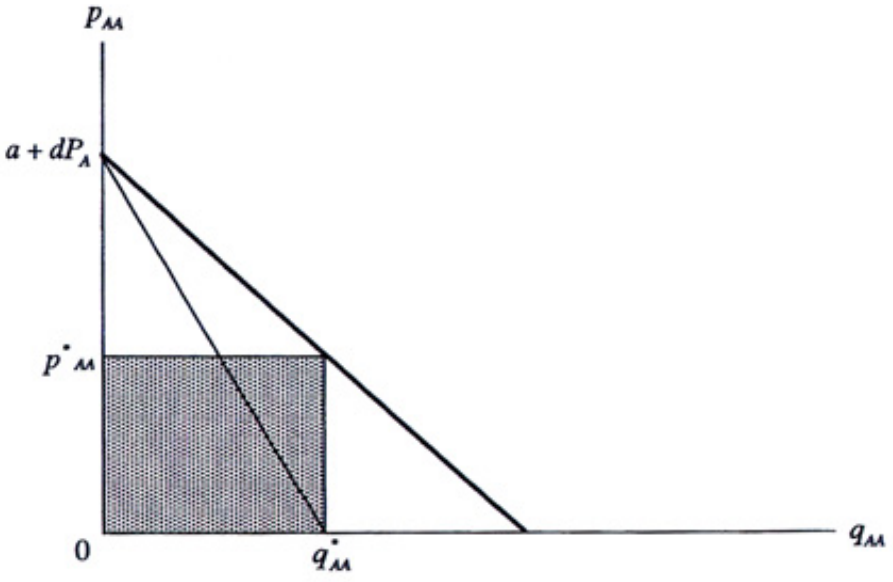


Figure 9.5: The impact of an increase in the number of firms on demand.

in, say, region A for a typical local variety; for simplicity, the unit of the differentiated product is chosen such that $b + dM = 1$:

$$p_{AA} = a + dP_A(M_A, t) - \frac{q_{AA}}{L/2 + fM_A}. \tag{9.63}$$

Because $p_{BA} > p_{AA}$ and the total number of firms is fixed by skilled labor market clearing, the price index P_A is a decreasing function of M_A at a rate that increases with t:

$$\frac{\partial P_A(M_A, t)}{\partial M_A} < 0 \quad \text{and} \quad \left| \frac{\partial^2 P_A(M_A, t)}{\partial M_A \partial t} \right| > 0. \tag{9.64}$$

The horizontal and vertical intercepts of (9.63) are, respectively, $[a + dP_A(M_A, t)]$ times $(L/2 + fM_A)$ and $[a + dP_A(M_A, t)]$. The equilibrium values of q_{AA} and p_{AA} are shown as q^*_{AA} and p^*_{AA}. They are found by setting marginal revenue equal to marginal cost (here zero). The equilibrium operating profits are given by the shaded rectangle and accrue to the skilled workers, whereas, as usual, the triangle above this rectangle represents the consumer surplus enjoyed by both skilled and unskilled workers.

Figure 9.5 is a powerful learning device to understand the forces at work in the model. To see why, start from an initial situation in which regions are identical ($M_A = M_B$). Suppose that some firms move from the foreign to the home region so that M_A rises and M_B falls. For these firms to stay in region A, their operating profits should not decrease. An increase in M_A has two opposite effects on operating profits. First, as new firms enter region A, the price index $P_A(M_A, t)$ decreases. Ceteris paribus, this would shift the inverse demand (9.63) toward the origin of the axes and operating profits would shrink. This effect is due to increased competition in the corresponding market and occurs because fewer firms now face transportation costs when supplying their home market. But this negative competition effect is not the only one. For some firms to move to the home region, some skilled workers have to follow ($M_A = \lambda H/f$). This means that, as M_A increases, λ also goes up so that the market of region A expands. Ceteris paribus, the horizontal intercept of the inverse demand would move away from the origin and profits would expand. This is a positive demand effect induced by the linkage between the locations of firms and skilled workers' expenditures.

Because the two effects oppose each other, the net result is a priori ambiguous. But we can say more than that. In particular, we can assess which effect prevails, depending on parameter values. Start with the competition effect that goes through $[a + dP_A(M_A, t)]$. This effect is strong if d is large, that is, if varieties are close substitutes. It is also strong if $|\partial P_A(M_A, t)/\partial M_A|$ is large. As shown in (9.64), this happens if t is large because, when obstacles to trade are high, competition from the other region is weak and home firms care more about their competitors being close rather than distant. As for the demand effect, it will be strong if f is large, because each new firm brings along many skilled

workers, and if L is small because the skilled have a relatively large impact on the local market size.

We can therefore conclude that the demand effect dominates the competition effect; consequently, agglomeration occurs when goods are very differentiated (d small), increasing returns are strong (f large), and transportation costs are low (t small). Under such circumstances, the entry of new firms in one region raises the operating profits of all firms. Higher profits would attract more firms, generating circular causation among firms and workers. Agglomeration is then sustainable as a spatial equilibrium.

9.4.4 A Welfare Analysis of the Core–Periphery Structure

We now wish to determine whether or not agglomeration is efficient, and if so, why and when (Ottaviano and Thisse 2001). A quasi-linear utility allows us to measure global efficiency by using the sum of individual utilities.

Different distortions and external effects are at work in the core–periphery model that suggest a significant discrepancy between equilibrium and optimum. Besides the standard distortion caused by firms' not pricing at marginal cost, there are several pecuniary externalities and, because our economy is imperfectly competitive, they matter for the level of welfare. In particular, skilled workers impose a pecuniary externality on the workers of the traditional sector. Indeed, their move affects the intensity of local competition in both the product and labor markets. In addition, when skilled workers move from one region to the other, they do not account for the impact of their migration on the other skilled. Their move affects not only the intensity of competition but also the level of demand inside each region, and, therefore, their wages. Finally, recall that there is no over- or under-entry effect. Indeed, the mass of firms is the same in equilibrium and at the optimum because it is determined by the technology and equal to H/f.

Let us first determine the optimum pattern.[25] For that, we assume that the planner is able to assign any mass of skilled workers (or, equivalently, of firms belonging to the modern sector) to a specific region and to use lump-sum transfers from all workers to pay for the loss firms incur while pricing at marginal cost. Since operating profits are just equal to the wage bill, the total welfare W in our two-region economy is given by the consumer surplus augmented by operating profits. The planner then chooses λ to maximize the total welfare W in our two-region economy given by (recall that individual utilities are quasi-linear)

$$W = \frac{L}{2}C_A + \lambda H(C_A + w_A) + \frac{L}{2}C_B + (1-\lambda)H(C_B + w_B) + L + \text{constant}, \tag{9.65}$$

where C_r is the consumer surplus in region $r = A, B$ in which all prices have

been set equal to marginal cost:

$$p_{rr}^0 = 0 \quad \text{and} \quad p_{rs}^0 = t \quad s \neq r.$$

This implies that operating profits are zero; hence, $w_r^0(\lambda) = 0$ for every λ so that firms do not incur any loss. Consequently, (9.65) becomes

$$W(\lambda) = C_0 t(t^0 - t)\lambda(\lambda - 1) + \text{constant}, \tag{9.66}$$

where

$$C_0 \equiv \frac{H^2}{2f^2}[2bf + d(H + L)]$$

and

$$t^0 \equiv \frac{4af}{2bf + d(H + L)},$$

which, again, does not depend on M.

The function (9.66) is strictly concave in λ if $t > t^0$ and strictly convex if $t < t^0$. Furthermore, because the coefficients of λ^2 and of λ are the same (up to their sign), this expression always has an interior extremum at $\lambda = 1/2$. As a result, the optimal choice of the planner is determined by the sign of the coefficient of λ^2, that is, by the value of t with respect to t^0.

Hence we have the following:

Proposition 9.4 *Consider a two-region economy with segmented markets. If $t > t^0$, then the symmetric configuration is the first best optimum; if $t < t^0$, any agglomerated configuration is an optimum; if $t = t^0$, the spatial configuration does not affect the welfare level.*

As expected, it is socially desirable to agglomerate the modern sector into a single region once transportation costs are low, increasing returns in the modern sector are strong enough, the output of the modern sector is differentiated enough, or all three conditions obtain. The higher the degree of increasing returns, the higher is the critical value. In particular, agglomeration is never efficient ($t^0 = 0$) when there are no scale economies in production ($f = 0$) or when varieties are good substitutes ($d \to \infty$). In this case, both the market equilibrium and the optimum involve two self-sufficient regions that each produce all the varieties.

We now assume that lump sum transfers are not available to the planner, who is only able to assign locations to the skilled workers. In such a context, the social welfare function is still given by (9.65) but is now evaluated at the equilibrium prices (9.55) and (9.56) and wages (9.59). This leads to

$$W_S = C_S t(t^S - t)\lambda(\lambda - 1) + \text{constant},$$

where

$$C_S \equiv \frac{H^2(bf+dH)}{8f^2(2bf+dH)^2}[8bf(3bf+2dH+dL)+3d^2H(L+H)]$$

and

$$t^S \equiv \frac{16af(3bf+dH)}{8bf(3bf+2dH+dL)+3d^2H(H+L)}. \tag{9.67}$$

The choice of the planner is similar to that described in the first best case except that the threshold value of t is now given by t^S.

It can readily be verified that $t^* > 0$ (where t^* is given by (9.61)) implies $t^S > 0$. Hence we have the following:

Proposition 9.5 *Consider a two-region economy with segmented markets and assume that (9.62) does not hold. If $t > t^S$, then the symmetric configuration is the second best optimum; if $t < t^S$, any agglomerated configuration is a second best optimum; if $t = t^S$, the spatial configuration does not affect the welfare level.*

Some simple calculations reveal that $t^0 < t^S < t^*$ when (9.62) does not hold. These inequalities reveal several important things. First, $t^0 < t^S$, namely, agglomeration is desirable for higher values of the transport cost in the second best. This is because the individual demand elasticity is much lower in the first best (marginal cost pricing) than in the second best (Nash equilibrium pricing); thus, regional price indices are less sensitive to a decrease in t. The fall in transport costs must be sufficiently large to make the agglomeration of the mobile workers socially desirable.[26]

Second, we also have $t^S < t^*$. This is because skilled workers do not internalize the negative external effects they impose on the unskilled, who always prefer dispersion. Hence, even though the skilled have incentives to move, these incentives do not reflect the social value of their move. This discrepancy is even stronger when we compare the first best outcome and the market equilibrium because we have just seen that marginal cost pricing is more favorable to dispersion. As a result, the market yields the core–periphery structure for a whole range ($t^0 < t < t^*$) of transportation cost values under which it is socially desirable to have a dispersed pattern of activities.

Accordingly, when transport costs are low ($t < t^0$) or high ($t > t^*$), the market yields the optimal location pattern; thus, no regional policy is required from the efficiency point of view, although equity considerations might still justify such a policy when agglomeration arises. On the contrary, for intermediate values of transport costs ($t^0 < t < t^*$), the market provides excessive agglomeration, thus justifying the need for an active regional policy to foster the dispersion of the modern sector from both the efficiency and equity standpoints.[27]

To decipher the nature of the various spatial external effects at work in the core–periphery model, it is useful to consider the problem in which the planner maximizes (9.65) with respect to λ, controlling for the inefficiencies generated by noncompetitive pricing in the modern sector. The first-order condition requires equating the sum of three terms to zero:

$$W_S' = \Delta v H + [(C_A' + w_A')\lambda + (C_B' + w_B')(1-\lambda)]H + (C_A' + C_B')\frac{L}{2}.$$

On the right-hand side, the first term is the indirect utility differential that an independent skilled mover takes into account when migrating. The second and third terms stand for the external effects that the mover imposes on, respectively, the other skilled and unskilled workers. From previous results, we know that these three terms equate zero at $\lambda = 1/2$ so that the dispersed configuration always satisfies the first-order conditions of the planner ($W_S' = 0$), of the skilled as a whole ($\Delta v + [(C_A' + w_A')\lambda + (C_B' + w_B')(1-\lambda)] = 0$), and of the unskilled as a whole ($C_A' + C_B' = 0$), as well as the indifference condition of an individual skilled worker ($\Delta v = 0$). Consequently, any discrepancy between the market outcome and the optimum must arise from the second-order condition:

$$\begin{aligned} W_S'' &= 2[(C_A' + w_A') - (C_B' + w_B')]H \\ &\quad + [(C_A'' + w_A'')\lambda + (C_B'' + w_B'')(1-\lambda)]H + (C_A'' + C_B'')\frac{L}{2}, \end{aligned} \tag{9.68}$$

where we have used the fact that the derivative of Δv with respect to λ is equal to $(C_A' + w_A') - (C_B' + w_B')$. Condition (9.68) can be rearranged to give

$$\begin{aligned} W_S'' &= 2[(C_A' + w_A') - (C_B' + w_B')]H \\ &\quad + (C_A'' + C_B'')\frac{H+L}{2} + (w_A'' + w_B'')\frac{H}{2}, \end{aligned} \tag{9.69}$$

where the first term on the right-hand side is positive whenever agglomeration is the stable outcome, the second term on the right-hand side is always negative, and the third is positive (negative) whenever transportation costs are above (below) the value:

$$t^{\Pi} \equiv \frac{8af}{4bf + d(H+L)},$$

which is larger than t^*.

Therefore, when, for $\lambda \neq 1/2$, a skilled worker finds it individually rational to move to the larger region (i.e., when $(C_A' + w_A') - (C_B' + w_B') > 0$ as implied by $t < t^*$). In doing so, the worker neglects the collective welfare losses caused to the unskilled as expressed by $(C_A'' + C_B'')L/2$. He also neglects the negative effect on the welfare of the skilled expressed by

$$(C_A'' + C_B'')H/2 + (w_A'' + w_B'')H/2 \tag{9.70}$$

in which both terms are negative because $t^* < t^{\Pi}$. Hence, even within the skilled population the gainers cannot compensate the losers. This is because of the following general equilibrium effect: at the margin, fiercer price competition in the larger region depresses operating profits, thus wages, more than it increases them in the smaller region.

By giving weight only to the first positive term of (9.69), the skilled mover always imposes a negative external effect on the economy as a whole. In addition, because the absolute value of (9.70) decreases with t as transportation costs fall, the negative external effects become weak enough to align the agglomerated market equilibrium with the second best outcome. This suggests why a planner would choose agglomeration over a smaller interval of transportation costs than the market does.

9.5 ON THE IMPACT OF FORWARD-LOOKING BEHAVIOR

So far, we have assumed that workers care only about their current utility level, thus implying that only history matters. This is a fairly restrictive assumption to the extent that migration decisions are typically made on the grounds of current and future utility flows and various costs as a result of search, mismatch, and homesickness. In addition, this approach has been criticized because it is not a priori consistent with fully rational, forward-looking behavior. It is therefore important to determine if and how workers' expectations about the evolution of the economy may influence the process of agglomeration. In this section, we want to see how the core–periphery model can be used to shed more light on the interplay between history and expectations in the formation of the economic space when migrants maximize the intertemporal value of their utility flows. In particular, we are interested in identifying the conditions under which, when regions initially host different numbers of workers and firms, the common belief that workers will eventually agglomerate in the initially smaller region can reverse the historically inherited advantage of the larger region.

Somewhat different approaches have been proposed to tackle this problem, but they have yielded similar conclusions (Ottaviano 1999; Baldwin 2001; Ottaviano et al. 2002). In what follows, we use the model developed in Section 9.4.1 because it leads to a simple linear dynamic system that still allows for a detailed analysis of the main issues (Krugman 1991c; Fukao and Bénabou 1993). Formally, we want to determine the parameter domains of this model for which an equilibrium path exists consistent with workers' belief on the assumption that workers have perfect foresights (self-fulfilling prophecy).

Workers infinitely live with a rate of time preference equal to $\gamma > 0$. Because we wish to focus on the sole dynamics of migration, we assume that the consumption of the numéraire is positive for each point in time so that there is no intertemporal trade in the differentiated good.[28] For example, consider the case in which workers believe that region A will eventually attract the modern

sector although region B is initially larger than A. Therefore, we want to test the consistency of the belief that, starting from time $\theta = 0$, all workers will end up being concentrated in A at some future date $\theta = T$, that is, there exists $T > 0$ such that, given $\lambda_0 < 1/2$,

$$\dot{\lambda}(\theta) > 0 \qquad \theta \in [0, T) \tag{9.71}$$
$$\lambda(\theta) = 1 \qquad \theta \geq T.$$

Let $v_r(s)$ be the instantaneous indirect utility at time s in region r and denote by v^C the utility level incurred in region A when $\lambda = 1$. Given workers' expectations, we have

$$v_A(\theta) = v^C \quad \text{for all } \theta \geq T.$$

Because workers have perfect foresights, the easiest way to generate a non-bang-bang migration behavior is to assume that, when moving from one region to the other, workers incur a utility loss that depends on the rate of migration inasmuch as a migrant typically imposes a negative externality on the others (Mussa 1978). Specifically, we assume that the utility loss for a migrant at time θ is equal to $|\dot{\lambda}(\theta)|/\delta$, where $\delta > 0$ is a positive constant whose meaning is given below. Thus, under (9.71), the intertemporal utility at time 0 of a worker who moves from B to A at time $\theta \in [0, T)$ is assumed to be given by

$$\begin{aligned} U(\theta) &\equiv \int_0^\theta e^{-\gamma s} v_B(s)\, ds + \int_\theta^\infty e^{-\gamma s} v_A(s)\, ds - e^{-\gamma\theta}\dot{\lambda}(\theta)/\delta \\ &= \int_0^\theta e^{-\gamma s} v_B(s)\, ds + \int_\theta^T e^{-\gamma s} v_A(s)\, ds - e^{-\gamma T} v^C/\gamma - e^{-\gamma\theta}\dot{\lambda}(\theta)/\delta, \end{aligned}$$

where the first term stands for the utility accumulated in region B before moving to A, the second term for the utility obtained in A after migration, and the last represents the migration cost at time θ.

Because in equilibrium the skilled residing in region B do not want to delay their migration beyond T, it follows that[29]

$$\lim_{\theta \to T} |\dot{\lambda}(\theta)|/\delta = 0.$$

When θ tends toward T, we therefore obtain

$$U(T) = \int_0^T e^{-\gamma s} v_B(s)\, ds + e^{-\gamma T} v^C/\gamma.$$

Let

$$V_r(\theta) \equiv \int_\theta^T e^{-\gamma(s-\theta)} v_r(s)\, ds + e^{-\gamma(T-\theta)} v^C/\gamma \quad \theta \in [0, T) \tag{9.72}$$

be the discounted sum of future utility flows gross of moving costs of a worker

currently (i.e., at time θ) residing in region $r = A, B$. Then, we have

$$\begin{aligned} U(\theta) - U(T) &= \int_{\theta}^{T} e^{-\gamma s}[v_A(s) - v_B(s)]ds - e^{-\gamma\theta}\dot{\lambda}(\theta)/\delta \\ &= e^{-\gamma\theta}[V_A(\theta) - V_B(\theta)] - e^{-\gamma\theta}\dot{\lambda}(\theta)/\delta \\ &= e^{-\gamma\theta}\Delta V(\theta) - e^{-\gamma\theta}\dot{\lambda}(\theta)/\delta, \end{aligned} \tag{9.73}$$

where $\Delta V(\theta) \equiv V_A(\theta) - V_B(\theta)$. Given that workers are free to choose when to migrate, it follows that

$$U(\theta) = U(T) \qquad \theta \in [0, T)$$

along the equilibrium path, and thus (9.73) implies that

$$\dot{\lambda}(\theta) = \delta\Delta V(\theta) \qquad \theta \in [0, T); \tag{9.74}$$

therefore, the private marginal cost of moving equals its private marginal benefit at any time $\theta < T$. In this expression, δ represents the speed of adjustment.

Using (9.72), we obtain the second law of motion by differentiating the utility gap $V_A(\theta) - V_B(\theta)$ with respect to θ

$$\begin{aligned} \dot{\Delta V}(\theta) &= \gamma\Delta V(\theta) - \Delta v(\theta) \qquad \theta \in [0, T) \\ &= \gamma\Delta V(\theta) - C_M t(t^* - t)(\lambda(\theta) - 1/2) \qquad \theta \in [0, T), \end{aligned} \tag{9.75}$$

where $\Delta v(\theta) \equiv v_A(\theta) - v_B(\theta)$ is given by (9.60) for all $\theta \in [0, T)$, whereas $\Delta v(\theta) = 0$ for all $\theta \geq T$ because, as already noted, all workers expect a utility flow $v_A(\theta)$. As argued by Ottaviano (1999), expression (9.75) states that, during the migration process, the "annuity value" of being in A rather than in B (i.e., $\gamma\Delta V$) equals the "dividend" (Δv) plus the "capital gain" ($d\Delta V/d\theta$).

As a result, we obtain a system of two linear differential equations instead of one (see (9.60)), with the terminal conditions $\lambda(T) = 1$ and $\Delta V(T) = 0$. Because $\lambda = 1/2$ implies $\Delta V = 0$, the system (9.74) and (9.75) always has an interior steady state at $(\lambda, \Delta V) = (1/2, 0)$, which corresponds to the symmetric configuration. When $t > t^*$, it is the only steady state that is globally stable; hence, for the assumed belief (9.71) to be consistent with equilibrium, transportation costs must be low ($t < t^*$) – a case on which we concentrate from now on.

To identify the conditions under which expectations matter, we must study the global stability of the system (9.74) and (9.75). Because this system is linear, local and global stability properties coincide, thus allowing us to appeal only to the former. The eigenvalues of the Jacobian matrix system (9.74) and (9.75) are given by

$$\frac{\gamma \pm \sqrt{\gamma^2 - 4\delta C_M t(t^* - t)}}{2}. \tag{9.76}$$

When $t < t^*$ two scenarios may arise. In the first one, $\gamma > \sqrt{C_M \delta} t^*$, which is more likely to occur when moving costs are high, when consumers are impatient, or both. The two eigenvalues are still real but both positive. The steady state $(1/2, 0)$ is an unstable node, and there are two trajectories that steadily go to the endpoints, $(0, 0)$ or $(1, 0)$, depending on the initial spatial distribution of workers. In this case, only history matters: because $\lambda_0 < 1/2$ by assumption, there is a single trajectory that goes to $\lambda = 0$, as in the case in which the dynamics are given by (9.60). This implies that belief (9.71) is inconsistent with the equilibrium path.[30]

Matters turn out to be quite different in the second scenario in which $\gamma < \sqrt{C_M \delta} t^*$, that is, when both transportation and moving costs are low. Given that $C_M t(t^* - t) = 0$ at both $t = 0$ and $t = t^*$, the equation $C_M t(t^* - t) - \gamma^2/4\delta = 0$ has two positive real roots in t, denoted t_1^e and t_2^e, which are both smaller than t^*:

$$t_1^e \equiv \frac{t^* - E}{2} \quad \text{and} \quad t_2^e \equiv \frac{t^* + E}{2},$$

where

$$E \equiv \sqrt{(t^*)^2 - \gamma^2 / C_M \delta}$$

stands for the size of the domain of values of t for which expectations matter; it shrinks as the discount rate γ increases or as the speed of adjustment δ decreases. Indeed, for $t \in (0, t_1^e)$ as well as for $t \in (t_2^e, t^*)$, both eigenvalues are real positive numbers, and the steady state $(1/2, 0)$ is an unstable node as before.

However, for $t \in (t_1^e, t_2^e)$, the eigenvalues become complex numbers with a positive real part so that the steady state is an unstable focus. Consequently, as illustrated in Figure 9.6, for any λ_0 close to but smaller than $1/2$, there is one trajectory going to $\lambda = 1$.[31] It is in such a case that expectations decide along which trajectory the system moves, and thus belief (9.71) is self-fulfilling. As shown in Figure 9.6, a symmetric trajectory going to $\lambda = 0$ exists under the symmetric belief. In other words, expectations matter for λ_0 close enough to $1/2$, whereas history matters otherwise.

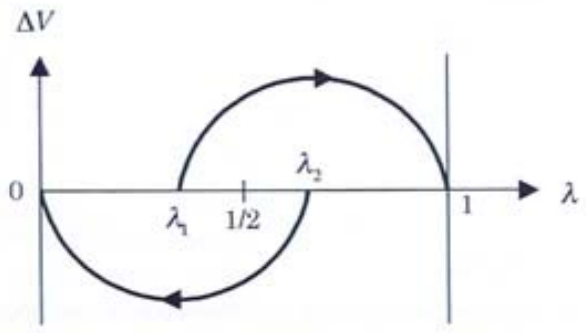


Figure 9.6: The overlap when expectations matter.

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

$$\lambda_1 \equiv \frac{1}{2}(1 - \Lambda) \quad \text{and} \quad \lambda_2 \equiv \frac{1}{2}(1 + \Lambda)$$

where

$$\Lambda \equiv \exp\left(-\frac{\gamma\pi}{\sqrt{4\delta C_M t(t^* - t) - \gamma^2}}\right)$$

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

The overlap is nonempty as long as $t \in (t_1^e, t_2^e)$. Thus, the width of the overlap is increasing in δ, C_M, and t^*, whereas it decreases with γ. Moreover, it is $\cap$-shaped with respect to t, reaching a maximum at $t = t^*/2$. Consequently, we have shown the following result:

Proposition 9.6 *Let λ_0 be the initial spatial distribution of workers. If $t < t^*$ and $\gamma < \sqrt{C_M \delta t^*}$, then $t_1^e \in (0, t^*/2)$, $t_2^e \in (t^*/2, t^*)$, $\lambda_1 \in (0, 1/2)$, and $\lambda_2 \in (1/2, 1)$ exist, such that workers' beliefs about their future utility flows influence the process of agglomeration if and only if $t \in (t_1^e, t_2^e)$ and $\lambda_0 \in [\lambda_1, \lambda_2]$.*

Hence, history alone matters when t is large enough or small enough. In other words, the agglomeration process evolves as if moving costs were zero when obstacles to trade are high or low. By contrast, *when transportation costs take intermediate values and regions are not initially too different, the region that becomes the core is determined by workers' expectations and not by history.*

The range (t_1^e, t_2^e) may be explained as follows. Suppose, indeed, that the economy is such that $\lambda_0 < 1/2$ and ask what is needed to reverse an ongoing agglomeration process leading toward $\lambda = 0$. If the evolution of the economy were to change direction, workers would experience falling, instantaneous, indirect utility flows for some time period as long as $\lambda < 1/2$. The instantaneous, indirect utility flows would start growing only after λ became larger than $1/2$. Put differently, workers would first experience utility losses followed by utility gains. Because the losses would come before the gains, they would be less discounted. This reasoning provides the root for the intuition behind Proposition 9.6. When circular causation leads to substantial wage increases (that is, for intermediate values of t), the benefits of agglomerating at $\lambda = 1$ can compensate workers for the losses they incur during the transition phase, thus making the reversal of migration possible. On the contrary, for low or high values of t, the benefits of agglomerating at $\lambda = 1$ do not compensate workers for the losses.

Finally, when $\lambda(0)$ is closer to one endpoint, the period over which the skilled bear losses is larger, a higher rate of time preference giving more weight to

them. Similarly, large moving costs (δ small) extend the time period over which workers' well-being is reduced, thus strengthening the weight of history in the agglomeration on process.

9.6 CONCLUDING REMARKS

Although the details of the agglomeration process vary with the model and the pricing policy, both the CES utility–iceberg transport cost and the quadratic utility–linear transport cost models of monopolistic competition show that a tendency toward agglomeration exists when transportation costs are sufficiently low. This finding suggests that the secular fall in transportation costs has intensified the geographical concentration of economic activities and confirms earlier analyses by Bairoch (1965) and Kaldor (1970, 340), who has observed the following:

> As communication between different regions becomes more intensified (with improvements in transport and marketing organisation), the region that is initially more developed industrially may gain from the progressive opening of trade at the expense of the less developed region whose development will be inhibited by it.

The decline in transport costs could have suggested that firms would become indifferent about their location. However, footloose activities, which are inherently independent of first nature, rely increasingly on second nature mechanisms. There are at least two reasons behind this phenomenon. First, as transportation costs decrease, firms have an incentive to concentrate their production in a smaller number of sites to better exploit scale economies in production. Second, low transportation costs tend to make price competition fierce, thus inducing firms to differentiate their products to relax price competition. This in turn leads firms to seek locations in which they have the best access to the largest pool of potential customers. In view of the results discussed in Chapters 7 and 8, this tendency suggests that *product differentiation is a powerful force toward agglomeration.*[32]

The process of economic unification within the European Union (EU) makes it ideal for testing the validity of the core–periphery model's predictions. As tariff barriers have been dropped and nontariff barriers reduced by a systematic policy aimed at improving integration of national markets, trade costs have dramatically decreased inside the EU. Recent empirical works seem to confirm the basic result of the core–periphery model. From 1968 to 1990, Amiti (1998) observed both an increase in the geographical concentration of economic activities within most EU state members and, for a vast majority of sectors, a tendency toward more agglomeration within the EU as a whole. Similarly, Brülhart (1998) noted an increase in the geographical concentration for 14 out of 18 industrial sectors between 1980 and 1990 within the EU. The cross-sectional analysis conducted by Haaland et al. (1999) in 13 European countries evidences cumulative causation in the sense that agglomeration of production and agglomeration of expenditure influence each other. Interestingly, these authors have also found

that the most important determinant in the European geography is the location and size of demand, thus confirming a result obtained by Justman (1994) for the United States.

To our knowledge, one of the most elaborated existing studies has been conducted by Combes and Lafourcade (2001) who performed a structural estimation of a multiregional, multisectorial model with vertical linkages. More precisely, they consider 71 industrial sectors and 341 employment areas in France; transport costs are evaluated by means of a distance and time cost function in 1978 and 1993, built from the road network, gas price, highways tolls, and carriers' contracts. Their work shows that a decrease of about 40% in road transport costs was associated with a strengthening in regional specialization and inequalities.

Yet, these tentative conclusions must be qualified. Indeed, the introduction of various types of modifications in the core–periphery model casts doubt on the monotonicity of the agglomeration–transportation costs relationship. First, Venables (1996) noticed that the core–periphery structure rests very much on the assumption of spatial mobility of the skilled workers. This may be a reasonable assumption within some parts of the world (e.g., the United States) but not necessarily in others (e.g., the European Union). It is then more realistic to assume that all workers stay put. In this case, we have seen in Section 9.3 that the existence of an intermediate sector is another reason for a core–periphery structure. However, Krugman and Venables (1995) have invited us to reconsider the conclusions of Krugman (1991a) in that the relationship between the agglomeration rate of the modern sector and the level of transportation costs would be $\cap$-shaped instead of monotone decreasing. Because of the immobility of workers, the agglomeration of firms into a single region intensifies competition on the corresponding regional labor market. The tendency of firms to attract workers from the traditional sector in their region but not from the other region turns out to be a dominant centrifugal force when transportation costs are low enough. Thus, if agglomeration occurred for intermediate values of the transport costs, it would collapse eventually when these costs decreased further. More generally, Puga (1999) has shown that, if workers do not move between regions when their mobility costs between sectors are relatively low, a drastic fall in transportation costs eventually leads to geographic dispersion.

Second, when transport costs for the $\mathbb{M}$-goods (e.g., the industrial goods) are low relative to those for the $\mathbb{T}$-goods (e.g., the agricultural goods), the (relative) price of the $\mathbb{T}$-goods may increase faster than the price of the $\mathbb{M}$-goods, and thus the real income does not necessarily rise in the large region. This makes the concentration of activities into a single region less attractive for both workers and firms, which may lead to the collapse of the core–periphery structure (Calmette and Le Pottier 1995; Fujita et al. 1999, chap. 7). However, although transportation costs of both types of goods have declined since the beginning of the Industrial Revolution, it appears that what matters ultimately

for the regional distribution of economic activities is not only the absolute levels of transport costs but also their relative values (Kilkenny 1998).

And, finally, Helpman (1998), Tabuchi (1998), and Ottaviano et al. (2002) have shown that positive urban costs, which take the form of housing and commuting costs (see Chapter 3), may also generate a ∩-shaped relationship. The reason is that a high concentration of workers into a single urban area gives rise to high land rents, as seen in Chapter 3, thus encouraging the relocation of workers and firms into the peripheral region. In these three cases, what fosters redispersion is the existence of factor price differentials in favor of the periphery that will trigger a new dispersion of activities once transport costs have sufficiently declined.

It is worth noting, however, that such models neglect possible firm relocations to the suburbs of the core metropolitan area rather than the periphery, as discussed in Sections 6.6 and 7.5. Indeed, suburbanization allows firms to enjoy low land rents and wages while preserving a high accessibility to a center supplying strategic inputs (e.g., a CBD). In fact, both urban and suburban locations seem to have been observed for a long time in many parts of the world. For example, in the case of Europe, Hohenberg, and Lees (1985, 129) observed the following:

> Without question, the long-run trend was for manufacturing to shift away from urban to rural locations, or at least to expand more in the latter, although cyclical changes in the economy could affect the trend. But in some cases, an urban location was always desirable, for example, where industries involved relatively heavy use of capital – fixed, circulating, or human – or the need for close entrepreneurial control. Less abstractly, the need to respond quickly to fashion or to other market changes, as well as the use of skilled labor, valuable raw materials, or expensive equipment argued for an urban location.

Moreover, all the models proposed in this chapter assume a given and fixed set of economic activities. However, as mentioned in Chapter 1, the decline of the industrial sector within big cities does not necessarily signify their economic and social decline. The continuous decrease in communications and transportation costs gives rise to new information-oriented economic activities that are typically developed within large metropolises (Feldman and Audretsch 1999; Henderson 1997b; Duranton and Puga 2002; Varga 2000).

In summary, for us, it is far from obvious that large metropolitan areas and regions are going to decline as transportation and communications costs keep decreasing.

APPENDIX

The following proof is due to Frédéric Robert-Nicoud. For notational convenience, set

$$\phi \equiv \Upsilon^{-(\sigma-1)} \tag{A.1}$$

and denote respectively by ϕ_{break} and ϕ_{sustain} the values of ϕ corresponding to

Υ_{break} and $\Upsilon_{sustain}$. Since $\sigma > 1$, ϕ is inversely related to Υ and belongs to the interval (0, 1].

Using (9.29) and (A.1), it is easy to show that $\phi_{sustain}$ is a solution to the equation

$$\phi^{-\mu\rho-1}[(1+\mu)\phi^2 + 1 - \mu] - 2 = 0.$$

Setting

$$f(\phi) \equiv \phi^{-\mu\rho-1}[(1+\mu)\phi^2 + 1 - \mu] - 2$$

it is straightforward that $f(0) \to \infty$ and $f(1) = 0$. Furthermore, it is readily verified that $f(\phi)$ has a single minimizer given by

$$\phi^* = \left[\frac{(1-\mu)(1+\mu\rho)}{(1+\mu)(1-\mu\rho)}\right]^{1/2} < 1.$$

Finally, it can be shown that $f'(0) < 0$ and $f'(1) > 0$. When ϕ varies from 0 to 1, all these facts put together imply that $f(\phi)$ decreases from arbitrarily large values to reach its minimum at $\phi^* > \phi_{sustain}$ since $f(\phi^*) < 0$ and, then, increases but takes negative values to reach the value 0 at $\phi = 1$.

Using (9.30) and (A.1), we obtain

$$\phi_{break} = \frac{(\sigma - 1 + \sigma\mu)(1+\mu)}{(\sigma - 1 - \sigma\mu)(1-\mu)}.$$

Evaluating $f(\phi)$ at ϕ_{break} reveals that $f(\phi_{break}) < 0$. Since $f(\phi_{sustain}) = 0$ and $f(\phi)$ decreases over $(0, \phi^*)$, it must be that

$$\phi_{break} > \phi_{sustain},$$

which yields the desired inequality

$$\Upsilon_{break} < \Upsilon_{sustain}.$$

NOTES

1. An earlier analysis that anticipated several aspects of Krugman's work was developed by Faini (1984). Ideas close to economic geography have already appeared in Krugman (1979) but were not fully worked out. Casetti (1980) has presented a simple dynamic analysis that permits the emergence of various equilibria and catastrophic transitions in a two-region economy. His article was written in 1970 but published in 1980, long before the emergence of new economic geography. Following a different approach, Arthur (1990; 1994, chap. 4) has shown the importance of "positive feedbacks" and history for the formation of regional clusters. Finally, in the presence of technological advantage of one region over the other, Markusen (1983) has shown that trade liberalization may generate more incentives

for labor migration, thus implying a partial agglomeration of workers within one region.

2. Yet, one should not forget that space allows the existence of pure profits even in free-entry equilibrium (Eaton and Lipsey 1978).
3. For more details, see Vives (1990).
4. This observation had already been made by Eaton and Lipsey (1977, 78–9).
5. This had already been noticed by Casetti (1980).
6. There is an older literature in trade theory that aims at modeling the transport sector just as another industry instead of using the iceberg approach. According to Neary (2001), this approach never led to simple results and is largely forgotten. However, this is enough to cast doubt on the generality of the results derived under the iceberg assumption because the two approaches do not seem to yield comparable results.
7. The most that can be obtained within this framework without resorting to numerical solutions probably has been achieved by Puga (1999).
8. When the number of firms is an integer, there are strategic interations (d'Aspremont, Dos Santos Ferreira, and Gérard-Varet 1996).
9. This agrees with the results obtained in industrial organization when firms are free to choose their product (Anderson et al. 1992, chap. 8).
10. It is worth noting that the summation in (9.11) may be interpreted as a market potential of region r. Indeed, each term in (9.11) gives the demand in s for a variety produced in r, which depends positively on the local income, negatively weighted by the accessibility of this region to region s. Finally, the competitiveness of region r is represented by $p_r(i)^{-\sigma}$.
11. We made a similar assumption in Section 8.4.
12. This process can be interpreted as a spatial Marshallian adjustment.
13. See Fujita et al. (1999, chap. 5) for more details.
14. Using the fact that function (9.27) always has a negative slope at $\Upsilon = 1$, we can show the shape of the curve as depicted in Figure 9.2.
15. It is worth noting that this result agrees with what we saw in Chapter 7 in the context of oligopolistic, but partial, equilibrium models.
16. This terminology is dictated by the secular fall in transportation costs.
17. We now assume that the modern sector uses labor through a Cobb–Douglas production function.
18. When $\beta = \delta$, (8.25) degenerates into a standard quadratic utility. See Ludema and Wooton (2000) for a model of economic geography with such a utility and quantity-setting firms.
19. This assumption allows one to focus on interior solutions only. Doing so has some costs in terms of generality but, as will be seen, larger benefits in terms of simpler analysis. The assumption is also consistent with the idea that each individual is interested in consuming both types of goods.
20. In the context of Section 9.2.5, this amounts to rescaling the demand intercept a in (9.52) and entails no loss of generality.
21. This is similar to what we obtained in Chapter 8 in the case of externalities. When the externality parameters are zero (ε in Section 9.2 and θ in Section 9.4), symmetry is the only equilibrium outcome. Sonnenschein (1982) has shown, to the contrary, a related result: perfect competition in each location is unable to explain

agglomerations. More precisely, if the initial distribution of firms is uneven along a given circle, then the spatial adjustment of firms in the direction of higher profit leads the economy toward a uniform long-run equilibrium in which each local economy is perfectly competitive. Mossay (2001) has revisited this problem in the context of local exchange economies and has shown that the spatial adjustment of consumers in the direction of higher utility also leads the economy toward a uniform long-run equilibrium when goods are gross substitutes. By contrast, when local markets are monopolistically competitive, Mossay (2001) has also shown that the spatial adjustment of workers renders the uniform long-run equilibrium unstable and, instead, generates local agglomerations.

22. It may also be shown that $\partial t^*/\partial \delta < 0$, which confirms what we just said.
23. Note that very small d implies that $t_{\text{trade}} < t^*$, and thus agglomeration always arises under trade.
24. Note that this comparison is made for a given level of transport costs. Starting from any τ for which equilibrium involves dispersion and lowering transport costs so that agglomeration arises in A, it can be shown that a sufficiently large decrease in τ also makes the unskilled in B better off.
25. Recall that we assume that transport costs are resource costs and not tariffs.
26. Although the argument above seems to rest very much on the linearity of demands, it is likely to hold true for any functional form such that the elasticity of demand for a variety increases with its price. This happens for any demand function that is concave or convex but less convex than a negative exponential – a condition often met in spatial pricing theory (Greenhut et al. 1987, 25). A notable violation of this property is given by isoelastic demands.
27. Under perfect competition and constant returns, those who remain in the source region are worse off when migration occurs. However, the global welfare always rises, thus allowing those left behind to be compensated by the migrants and the native-born in the destination region (Razin and Sadka 1997). As we have just seen, this is not necessarily true under increasing returns and monopolistic competition.
28. In the Krugman version of the core–periphery model, this is equivalent to assuming that expenditure equals income at each time (Baldwin 2001; Ottaviano 1999).
29. This condition has been introduced by Fukao and Bénabou (1993).
30. In this case, expectations do not play any role, and thus the myopic adjustment process studied in the previous section provides a good approximation of the qualitative evolution of the economy under forward-looking behavior (see also Baldwin 2001).
31. This trajectory is part of a spiral originating at $\lambda = 1/2$.
32. Similar arguments apply to labor: wage competition is a centrifugal force as is price competition, and a better access to a diversified labor pool for both firms and workers is a centripetal force.