# 4 The Dixit-Stiglitz Model of Monopolistic Competition and Its Spatial Implications

In any model in which increasing returns play a crucial role, one must somehow handle the problem of market structure. Traditional urban models deal with the issue by assuming that increasing returns are purely external to firms, allowing the modeler to continue to assume perfect competition. The approach taken in this book, however, avoids any direct assumption of external economies: Externalities emerge as a consequence of market interactions involving economies of scale at the level of the individual firm. Thus we must somehow model an imperfectly competitive market structure. The workhorse model of this kind is, of course, the Dixit-Stiglitz model of monopolistic competition (Dixit and Stiglitz 1977). Dixit-Stiglitz monopolistic competition is grossly unrealistic, but it is tractable and flexible; as we will see, it leads to a very special but very suggestive set of results.

This chapter develops a spatial version of the Dixit-Stiglitz model, that is, one with multiple locations and transport costs between those locations. This spatial Dixit-Stiglitz model is a crucial ingredient in almost everything that follows.

We consider an economy with two sectors, agriculture and manufacturing. The agricultural sector is perfectly competitive and produces a single, homogeneous good, whereas the manufacturing sector provides a large variety of differentiated goods. Of course, the label "agriculture" need not always be interpreted literally; the sector's defining characteristic is that it is the "residual," perfectly competitive sector that is the counterpart to the action taking place in the increasing-returns, imperfectly competitive manufacturing sector.

We imagine that there are a very large number of potential manufactured goods, so many that the product space can be represented as continuous, enabling us to sidestep integer constraints on the number of goods. Although each consumption and production activity takes place at a specific location, first we describe each type of activity without explicitly referring to the location.

## 4.1 Consumer Behavior

Every consumer shares the same Cobb-Douglas tastes for the two types of goods:

$$U = M^{\mu}A^{1-\mu}, \tag{4.1}$$

where $M$ represents a composite index of the consumption of manufactured goods, $A$ is the consumption of the agricultural good, and $\mu$ is a constant representing the expenditure share of manufactured goods. The quantity index, $M$, is a subutility function defined over a continuum of varieties of manufactured goods; $m(i)$ denotes the consumption of each available variety; and $n$ is the range of varieties produced, often called the "number" of available varieties. We assume that $M$ is defined by a constant-elasticity-of-substitution (CES) function:

$$M = \left[\int_0^n m(i)^{\rho}di\right]^{1/\rho}, \quad 0 < \rho < 1. \tag{4.2}$$

In this specification, the parameter $\rho$ represents the intensity of the preference for variety in manufactured goods. When $\rho$ is close to 1, differentiated goods are nearly perfect substitutes for each other; as $\rho$ decreases toward 0, the desire to consume a greater variety of manufactured goods increases. If we set $\sigma \equiv 1/(1-\rho)$, then $\sigma$ represents the elasticity of substitution between any two varieties.

Given income $Y$ and a set of prices, $p^A$ for the agricultural good and $p(i)$ for each manufactured good, the consumer's problem is to maximize utility (4.1) subject to the budget constraint,

$$p^A A + \int_0^n p(i)m(i)di = Y.$$

This problem can be solved in two steps.[1] First, whatever the value of the manufacturing composite, $M$, each $m(i)$ needs to be chosen so as to minimize the cost of attaining $M$. This means solving the following minimization problem:

$$\min \int_0^n p(i)m(i)di \quad \text{s.t.} \quad \left[\int_0^n m(i)^{\rho}di\right]^{1/\rho} = M. \tag{4.3}$$

The first-order condition to this expenditure minimization problem gives equality of marginal rates of substitution to price ratios,

$$\frac{m(i)^{\rho-1}}{m(j)^{\rho-1}} = \frac{p(i)}{p(j)}, \tag{4.4}$$

for any pair $i, j$ that leads to $m(i) = m(j)(p(j)/p(i))^{1/(1-\rho)}$. Substituting this equation into the original constraint,

$$\left[\int_0^n m(i)^\rho di\right]^{1/\rho} = M,$$

and bringing the common term, $m(j)p(j)^{1/(1-\rho)}$, outside the integral, we have that

$$m(j) = \frac{p(j)^{1/(\rho-1)}}{\left[\int_0^n p(i)^{\rho/(\rho-1)} di\right]^{1/\rho}} M. \tag{4.5}$$

This is simply the compensated demand function for the $j$th variety of manufacturing product.

We can also derive an expression for the minimum cost of attaining $M$. Expenditure on the $j$th variety is $p(j)m(j)$, so using (4.5) and integrating over all $j$ gives

$$\int_0^n p(j)m(j)dj = \left[\int_0^n p(i)^{\rho/(\rho-1)} di\right]^{(\rho-1)/\rho} M. \tag{4.6}$$

It is now natural to define the term multiplying $M$ on the right-hand side of this expression as a price index, so that the price index times the quantity composite is equal to expenditure. Denoting this price index for manufactured products by $G$ we have

$$G \equiv \left[\int_0^n p(i)^{\rho/(\rho-1)} di\right]^{(\rho-1)/\rho} = \left[\int_0^n p(i)^{1-\sigma} di\right]^{1/(1-\sigma)} \tag{4.7}$$

where $\rho \equiv (\sigma - 1)/\sigma$ or $\sigma \equiv 1/(1 - \rho)$. The price index, $G$, measures the minimum cost of purchasing a unit of the composite index $M$ of manufacturing goods, so just as $M$ can be thought of as a utility function, $G$ can be thought of as an expenditure function. Demand for $m(i)$ can now be written more compactly (using (4.7) in (4.5)) as

$$m(j) = \left(\frac{p(j)}{G}\right)^{1/(\rho-1)} M = \left(\frac{p(j)}{G}\right)^{-\sigma} M. \tag{4.8}$$

The upper-level step of the consumer's problem is to divide total income between agriculture and manufactures in aggregate, that is, to choose $A$ and $M$ so as to

$$\max U = M^{\mu}A^{1-\mu} \quad \text{s.t.} \quad GM + p^{A}A = Y, \tag{4.9}$$

which yields the familiar results that $M = \mu Y/G$ and $A = (1-\mu)Y/p^{A}$. Pulling the stages together, we obtain the following uncompensated consumer demand functions. For agriculture,

$$A = (1-\mu)Y/p^{A}, \tag{4.10}$$

and for each variety of manufactures

$$m(j) = \mu Y \frac{p(j)^{-\sigma}}{G^{-(\sigma-1)}} \quad \text{for } j \in [0, n]. \tag{4.11}$$

Notice that, holding $G$ constant, the price elasticity of demand for every available variety is constant and equal to $\sigma$.

We can now express maximized utility as a function of income, the price of agricultural output, and the manufactures' price index, giving the indirect utility function

$$U = \mu^{\mu}(1-\mu)^{1-\mu}YG^{-\mu}(p^{A})^{-(1-\mu)}. \tag{4.12}$$

The term $G^{\mu}(p^{A})^{(1-\mu)}$ is the cost-of-living index in the economy.

So far this is a straightforward exercise in demand theory. What is unusual in the Dixit-Stiglitz model—and plays a crucial role in our analysis—is that the range of manufactures on offer becomes an endogenous variable. This means that it is important to understand the effects on the consumer of changes in $n$, the number of varieties.

Increasing the range of varieties on offer reduces the manufactures' price index (because consumers value variety) and hence the cost of attaining a given level of utility. This can be seen most clearly if we assume that all manufactures are available at the same price, $p^{M}$. Then the price index, (4.7), simply becomes

$$G = \left[\int_{0}^{n} p(i)^{1-\sigma}di\right]^{1/(1-\sigma)} = p^{M}n^{1/(1-\sigma)}. \tag{4.13}$$

The price index's responsiveness to the number of varieties depends on the elasticity of substitution between varieties, $\sigma$, and we see that the lower is $\sigma$—the more differentiated are product varieties—the greater is the reduction in the price index caused by an increase in the number of varieties. The effect on welfare is then given by the indirect utility function, (4.12).

Changing the range of products available also shifts demand curves for existing varieties. This can be seen by looking at the demand curve for a single variety, equation (4.11). Because an increase in $n$ reduces $G$, it shifts each demand curve downward. This effect is important as we come to determine the equilibrium number of varieties produced. It says that as we increase the number of varieties, product market competition intensifies, shifting demand curves for existing products downward and reducing the sales of these varieties.

## 4.2 Multiple Locations and Transportation Costs

Depending on what we are trying to model, it is sometimes convenient to think of the economy as consisting of a finite set of locations (regions or countries), sometimes to think of it as spread across a continuous space. For present purposes, however, it is sufficient to think in terms of discrete locations, of which we suppose there are $R$. For the moment, assume that each variety is produced in only one location and that all varieties produced in a particular location are symmetric, having the same technology and price. We denote the number of varieties produced in location $r$ by $n_r$, and the mill or f.o.b. price of one of these varieties by $p_r^M$.

Agricultural and manufactured goods can be shipped between locations and may incur transport costs in shipment. To avoid modeling a separate transportation industry, we assume the "iceberg" form of transport costs introduced by von Thünen and Paul Samuelson.[2] Specifically, if a unit of the agricultural good [any variety of manufactured goods] is shipped from a location $r$ to another location $s$, only a fraction, $1/T_{rs}^A$ [$1/T_{rs}^M$], of the original unit actually arrives; the rest melts away en route. The constant $T_{rs}^A$ [$T_{rs}^M$] represents the amount of the agricultural [manufactured] good dispatched per unit received.

The iceberg transport technology implies that if a manufacturing variety produced at location $r$ is sold at price $p_r^M$, then the delivered (c.i.f.) price, $p_{rs}^M$, of that variety at each consumption location $s$ is given by

$$p_{rs}^M = p_r^M T_{rs}^M. \tag{4.14}$$

The manufacturing price index may take a different value in each location; we denote this by writing the price index for location $s$ as $G_s$.

Iceberg transport costs together with the assumption that all varieties produced in a particular location have the same price mean that, using equation (4.7), this price index can be written as

$$G_s = \left[ \sum_{r=1}^{R} n_r (p_r^M T_{rs}^M)^{1-\sigma} \right]^{1/(1-\sigma)}, \quad s = 1, \ldots, R. \tag{4.15}$$

Consumption demand in location $s$ for a product produced in $r$ now follows (from 4.11) as

$$\mu Y_s (p_r^M T_{rs}^M)^{-\sigma} G_s^{(\sigma-1)}, \tag{4.16}$$

where $Y_s$ is income for location $s$. This gives consumption, but to supply this level of consumption, $T_{rs}^M$ times this amount has to be shipped. Summing across locations in which the product is sold, the total sales of a single location $r$ variety, denoted, $q_r^M$, therefore amount to:

$$q_r^M = \mu \sum_{s=1}^{R} Y_s (p_r^M T_{rs}^M)^{-\sigma} G_s^{\sigma-1} T_{rs}^M. \tag{4.17}$$

This simply says that sales depend on income in each location, the price index in each location, transport costs, and the mill price. Notice that because the delivered prices of the same variety at all consumption locations change proportionally to the mill price, and because each consumer's demand for a variety has a constant price elasticity $\sigma$, the elasticity of the aggregate demand for each variety with respect to its mill price is also $\sigma$, regardless of the spatial distribution of consumers.

## 4.3 Producer Behavior

Next we turn to the production side of the economy. The agricultural good, we assume, is produced using a constant-returns technology under conditions of perfect competition. Manufacturing, however, we assume to involve economies of scale. These economies of scale arise at the level of the variety; there are no economies of scope or of multiplant operation. Technology is the same for all varieties and in all locations and involves a fixed input of $F$ and marginal input requirement $c^M$. Thus, assuming for the moment that the only input is labor, the production of a quantity $q^M$ of any variety at any given location requires labor input $l^M$, given by

$$l^M = F + c^M q^M. \tag{4.18}$$

Because of increasing returns to scale, consumers' preference for variety, and the unlimited number of potential varieties of manufactured goods, no firm will choose to produce the same variety supplied by another firm. This means that each variety is produced in only one location, by a single, specialized firm, so that the number of manufacturing firms in operation is the same as the number of available varieties.

### 4.3.1 *Profit Maximization*

Next, consider a particular firm producing a specific variety at location $r$ and facing a given wage rate, $w_r^M$, for manufacturing workers there. Then, with a mill price $p_r^M$, its profit is given by

$$\pi_r = p_r^M q_r^M - w_r^M (F + c^M q_r^M), \tag{4.19}$$

where $q_r^M$ is given by the demand function, (4.17). Each firm is assumed to choose its price taking the price indices, $G_s$, as given. The perceived elasticity of demand is therefore $\sigma$, so profit maximization implies that

$$p_r^M (1 - 1/\sigma) = c^M w_r^M, \qquad \text{or} \quad p_r^M = c^M w_r^M / \rho, \tag{4.20}$$

for all varieties produced at $r$.

We suppose that there is free entry and exit in response to profits or losses. Given the pricing rule, the profits of a firm at location $r$ are

$$\pi_r = w_r^M \left[ \frac{q_r^M c^M}{\sigma - 1} - F \right]. \tag{4.21}$$

Therefore, the zero-profit condition implies that the equilibrium output of any active firm is

$$q^* \equiv F(\sigma - 1)/c^M, \tag{4.22}$$

and the associated equilibrium labor input is

$$l^* \equiv F + c^M q^* = F\sigma. \tag{4.23}$$

Both $q^*$ and $l^*$ are constants common to every active firm in the economy. Therefore, if $L_r^M$ is the number of manufacturing workers at location $r$, and $n_r$ is the number of manufacturing firms ($\equiv$ the number of the varieties produced) at $r$, then

$$n_r = L_r^M / l^* = L_r^M / F\sigma. \tag{4.24}$$

The results (4.20) and (4.22) are somewhat odd but play a crucial role throughout our analysis. They say that *the size of the market affects neither the markup of price over marginal cost nor the scale at which individual goods are produced.* As a result, *all scale effects work through changes in the variety of goods available.* Obviously this is a rather strange result: Normally we think both that larger markets mean more intensive competition, and that one of the ways the economy takes advantage of the extent of the market is by producing at larger scale. The Dixit-Stiglitz model says, however, that all market-size effects work through changes in variety.

This result is an artifact of the constant-elasticity demand functions, together with the nonstrategic behavior implied by our assumption that firms take the price indices, $G_s$, to be constant as they solve their profit maximization problem. If we were to relax the assumption of nonstrategic behavior, each firm would then recognize that its choice changes the price index, and this recognition of market power would tend to reduce the firm's output and increase its price-cost margin. If we adopt a specific form of oligopolistic interaction, such as Cournot or Bertrand competition, then we can derive explicit expressions for the pricing rule, and in both these cases the price-cost margin is a decreasing function of each firm's market share.[3] Under these assumptions an increase in market size has a procompetitive effect. It causes entry of firms, which reduces price-cost margins and means that firms must operate at larger scale (and lower average cost) to break even. We have already seen (section 4.1) how variety effects create a negative relationship between market size and the price index; the procompetitive effect is a second force operating in the same direction.

Throughout our analysis, however, we choose to ignore this second effect. Having constant price-cost markups and firm scale is a dramatic simplification, allowing us to model cleanly issues that might otherwise seem quite intractable.

### *4.3.2 The Manufacturing Wage Equation*

We have seen that the condition that firms make no profits is equivalent to the condition that they produce $q^*$. Using the demand functions, (4.17), firms at location $r$ attain this level of output if the following equation is satisfied:

$$q^* = \mu \sum_{s=1}^{R} Y_s (p_r^M)^{-\sigma} (T_{rs}^M)^{1-\sigma} G_s^{\sigma-1}. \tag{4.25}$$

We can turn this equation around and say that active firms break even if and only if the price they charge satisfies

$$(p_r^M)^\sigma = \frac{\mu}{q^*} \sum_{s=1}^{R} Y_s (T_{rs}^M)^{1-\sigma} G_s^{\sigma-1}. \tag{4.26}$$

Using the pricing rule (4.20) this can be expressed as

$$w_r^M = \left(\frac{\sigma - 1}{\sigma c^M}\right) \left[\frac{\mu}{q^*} \sum_{s=1}^{R} Y_s (T_{rs}^M)^{1-\sigma} G_s^{\sigma-1}\right]^{1/\sigma}. \tag{4.27}$$

We refer to this as the *wage equation* and use it often. It gives the manufacturing wage at which firms in each location break even, given the income levels and price indices in all locations and the costs of shipping into these locations. As can be seen, this wage is higher the higher are incomes in the firms' markets, $Y_s$, the better is the firm's access to these markets (lower $T_{rs}^M$), and the less competition the firm faces in these markets. (Recall that the price index is decreasing in the number of varieties sold.)

Two important observations need to be made about the wage equation. First, we assume that active firms *always* make no profits, so that this equation gives the actual manufacturing wage in any location that has a nonzero number of firms. In the long run, this wage equals the supply price of labor to manufacturing but in the short run may differ from it. Any such difference gives rise to adjustment dynamics, which are spelled out in later chapters. Essentially then, we are assuming that the entry and exit of firms occurs very fast—so profits are always 0—but relocation of workers among sectors or locations occurs more slowly, with a dynamic that we will model explicitly.

Second, the manufacturing wage as given by (4.27) is defined even in locations that have no manufacturing. It then measures the maximum wage that could be paid by a firm considering production in the location.

### 4.3.3 *Real Wages*

Real income at each location is proportional to nominal income deflated by the cost-of-living index, $G_r^{\mu}(p_r^A)^{1-\mu}$. This means that the real wage of manufacturing workers in location, $r$ denoted $\omega_r^M$, is

$$\omega_r^M = w_r^M G_r^{-\mu}(p_r^A)^{-(1-\mu)}. \tag{4.28}$$

## 4.4 Some Normalizations

The manufacturing price index and the wage equation pop up frequently in this book. Happily, we can simplify them if we choose units of measurement appropriately. First, notice that we are free to choose units of measurement for output—be it units, tens of units, kilos, or tons. We choose units such that the marginal labor requirement satisfies the following equation:

$$c^M = \frac{\sigma - 1}{\sigma} \ (= \rho). \tag{4.29}$$

This normalization means that the pricing equation, (4.20), becomes

$$p_r^M = w_r^M \tag{4.30}$$

and also that $q^* = l^*$.

Second, as we have seen, the number of firms is simply an interval of the real line, $[0, n]$, and without loss of generality, we can choose units of measurement for this range. For sections II and III of this book we choose convenient units by setting the fixed input requirement $F$ to satisfy the following equation:

$$F = \mu/\sigma. \tag{4.31}$$

The number of firms in each location is related to the size of the manufacturing labor force in the location according to equation (4.24), which becomes

$$n_r = L_r^M/\mu. \tag{4.32}$$

These choices of units also set firm scale. The output level at which firms make no profit (equation (4.22)) becomes

$$q^* = l^* = \mu. \tag{4.33}$$

Using these normalizations the price index and wage equation can now be written in a more convenient form. The price index becomes

$$
\begin{aligned}
G_r &= \left[\sum_{s=1}^{R} n_s (p_s^M T_{sr}^M)^{(1-\sigma)}\right]^{1/(1-\sigma)} \\
&= \left[\frac{1}{\mu}\sum_{s=1}^{R} L_s^M (w_s^M T_{sr}^M)^{(1-\sigma)}\right]^{1/(1-\sigma)}.
\end{aligned} \tag{4.34}
$$

The wage equation becomes

$$
\begin{aligned}
w_r^M &= \left(\frac{\sigma - 1}{\sigma c^M}\right)\left[\frac{\mu}{q^*}\sum_{s=1}^{R} Y_s (T_{rs}^M)^{1-\sigma} G_s^{\sigma-1}\right]^{1/\sigma} \\
&= \left[\sum_{s=1}^{R} Y_s (T_{rs}^M)^{1-\sigma} G_s^{\sigma-1}\right]^{1/\sigma}.
\end{aligned} \tag{4.35}
$$

We use these two equations repeatedly, both to characterize equilibrium and to investigate its stability. Essentially we have chosen units in a way that shifts attention from the number of manufacturing firms and product prices to the number of manufacturing workers and their wage rates.

## 4.5 The Price Index Effect and the Home Market Effect

The price indices and wage equations (4.34) and (4.35) do not define a full economic model, but they nevertheless imply some of the most important relationships that drive the results that follow, and it is worth examining them in some detail to draw out these relationships.

Consider a two-location version of these equations. Writing the equations out in full, we have the price indices, (4.34),

$$
\begin{aligned}
G_1^{1-\sigma} &= \frac{1}{\mu}[L_1 w_1^{1-\sigma} + L_2 (w_2 T)^{1-\sigma}], \\
G_2^{1-\sigma} &= \frac{1}{\mu}[L_1 (w_1 T)^{1-\sigma} + L_2 w_2^{1-\sigma}],
\end{aligned} \tag{4.36}
$$

and the wage equations, (4.35),

$$
\begin{aligned}
w_1^{\sigma} &= Y_1 G_1^{\sigma-1} + Y_2 G_2^{\sigma-1} T^{1-\sigma}, \\
w_2^{\sigma} &= Y_1 G_1^{\sigma-1} T^{1-\sigma} + Y_2 G_2^{\sigma-1},
\end{aligned} \tag{4.37}
$$

where we have dropped the superscripts $M$, because we are looking only at manufacturing, and have denoted transport costs between locations by the single number $T$, and assumed, as we do throughout, that no transport costs are incurred within each location. These pairs of equations are symmetric, and so have a symmetric solution: That is, if $L_1 = L_2$ and $Y_1 = Y_2$, then there is a solution with $G_1 = G_2$ and $w_1 = w_2$. By inspection, it is easy to see that these symmetric equilibrium values satisfy the following relationships,

$$1 + T^{1-\sigma} = \frac{\mu}{L}\left(\frac{G}{w}\right)^{1-\sigma} = \frac{w}{Y}\left(\frac{G}{w}\right)^{1-\sigma} \tag{4.38}$$

where absence of subscripts denotes that these are symmetric equilibrium values.

We can explore the relationships contained in the price indices and wage equations by linearizing them around the symmetric equilibrium. Around this point an increase in a variable in one location is always associated with a change, of opposite sign but of equal absolute magnitude, in the corresponding variable in the other country. So letting $dG = dG_1 = -dG_2$, and so on, we derive, by differentiating the price indices and wage equations respectively,

$$(1-\sigma)\frac{dG}{G} = \frac{L}{\mu}\left(\frac{G}{w}\right)^{\sigma-1}(1-T^{1-\sigma})\left[\frac{dL}{L} + (1-\sigma)\frac{dw}{w}\right], \tag{4.39}$$

$$\sigma\frac{dw}{w} = \frac{Y}{w}\left(\frac{G}{w}\right)^{\sigma-1}(1-T^{1-\sigma})\left[\frac{dY}{Y} + (\sigma-1)\frac{dG}{G}\right]. \tag{4.40}$$

From the first equation, we can see the direct effect of a change in the location of manufacturing on the price index of manufactured goods. Suppose that the supply of labor to manufacturing is perfectly elastic, so that $dw = 0$. Bearing in mind that $1 - \sigma < 0$ and $T > 1$, equation (4.39) implies that a change $dL/L$ in manufacturing employment has a negative effect on the price index, $dG/G$. We call this the *price index effect*. It means that the location with a larger manufacturing sector also has a lower price index for manufactured goods, simply because a smaller proportion of this region's manufacturing consumption bears transport costs.

Next, let us consider how relative demand affects the location of manufacturing. It is convenient to define a new variable, $Z$,

$$Z \equiv \frac{1 - T^{1-\sigma}}{1 + T^{1-\sigma}}, \tag{4.41}$$

which is a sort of index of trade cost, with values between 0 and 1. If trade is perfectly costless, $T = 1$, then $Z$ takes the value 0; if trade is impossible, it takes value 1. Using the definition of $Z$ and eliminating $dG/G$ from equations (4.39) and (4.40) gives

$$\left[\frac{\sigma}{Z} + Z(1 - \sigma)\right]\frac{dw}{w} + Z\frac{dL}{L} = \frac{dY}{Y}. \tag{4.42}$$

We learn a number of things from this equation.

First, suppose that our wider economic model gives us a perfectly elastic supply of labor to manufacturing, so $dw = 0$. We then have a relationship known as the *home market effect*. A 1 percent change in demand for manufactures ($dY/Y$) causes a $1/Z$ ($> 1$) percent change in the employment in, and hence production of, manufactures, $dL/L$. That is, other things being equal, the location with the larger home market has a more than proportionately larger manufacturing sector, and therefore also exports manufactured goods.[4]

Second, although we have just derived the home market effect for the case when labor supply is perfectly elastic, this need not be the case; if the labor supply curve slopes upward, some of the home market advantage is taken out in higher wages rather than exports. Thus, *locations with a higher demand for manufactures may pay a higher nominal wage.*[5]

But notice that we have already seen that an increase in $L$ is, other things being the same, associated with a decrease in $G$. So if $Y$ is high in some region, we may expect the real wage to be high both because the nominal wage is high and because the price index is low. Hence *locations with a higher demand for manufactures tend, other things being equal, to offer a higher real wage to manufacturing workers.*

Of course other things need not be equal, but we have just sketched out several of the key elements of the cumulative causation that, in our models, tends to lead to agglomeration. Areas with large manufacturing sectors tend to have low price indexes for manufactures, because of the price index effect; areas with large demand for manufactures tend to have disproportionately large manufacturing sectors, because of the home market effect. If we fill in just one more relationship—that manufacturing workers themselves demand manufactures, so that locations with large concentrations of manufacturing also tend to have large demand for manufactured goods—we are almost there. But let us spell out the details in chapter 5.

## 4.6 The "No-Black-Hole" Condition

We have seen that increasing the size of the manufacturing sector tends to raise real income. However, we often want to put an upper bound on the strength of this effect. The condition we use is best explained by looking at a closed economy: a situation where $Z = 1$.

Consider the real income of a manufacturing worker, (4.28). Suppose that the price of agricultural output is constant, and totally differentiate this to give

$$\frac{d\omega}{\omega} = \frac{dw}{w} - \mu\frac{dG}{G}, \tag{4.43}$$

where we have once again dropped the superscripts $M$ and, because we are looking at a single economy, also the location subscripts. Now using (4.39) and (4.40) with $Z = 1$, we obtain

$$\begin{aligned} \frac{d\omega}{\omega} &= (1 - \mu)\frac{dY}{Y} + \left[\frac{\mu\sigma}{\sigma - 1} - 1\right]\frac{dL}{L} \\ &= (1 - \mu)\frac{dY}{Y} + \left[\frac{\mu - \rho}{\rho}\right]\frac{dL}{L}, \end{aligned} \tag{4.44}$$

which says the following: Suppose that we add more workers to the manufacturing sector of a closed economy, holding expenditure on the industry constant ($dY = 0$) and hence holding constant the nominal income generated. What effect does this have on real wages of workers in the sector? Clearly, because expenditure on manufactures is held constant, so is the wage bill, implying that an increase in $L$ reduces the wage $w$ equiproportionately. However, the increase in manufacturing employment increases the number of varieties of manufacturing products, thus reducing $G$ and tending to raise real income. This latter effect can conceivably outweigh the former, so that an increase in the number of workers would actually raise their real wage.

We in general are not interested in economies in which increasing returns are that strong, if only because, as we will see, in such economies the forces working toward agglomeration always prevail, and the economy tends to collapse into a point. To avoid such "black-hole location" theory, we usually impose what we call the *assumption of no black holes*:

$$\frac{\sigma - 1}{\sigma} = \rho > \mu. \tag{4.45}$$

We now have the building blocks of our approach and are ready to start examining some geography.

## Notes

1. A two-stage budgeting proceduring is applicable because preferences are separable between agriculture and manufactures and $M$, the subutility function for manufactures, is homothetic in the quantities $m(i)$. See Deaton and Muellbauer (1980) for discussion of conditions under which two-stage budgeting is appropriate.

2. The "iceberg" transport technology was formally introduced by Samuelson (1952). Von Thünen, however, supposed the cost of grain transportation to consist largely of the grain consumed on the way by the horses pulling the wagon (von Thünen, 1826, chap. 4). Hence, the von Thünen model may be considered as the predecessor of the "iceberg" transport technology.

3. See Smith and Venables 1988 for derivation of these expressions.

4. The home market effect should apply whether or not a cumulative process of agglomeration is at work. Indeed, Krugman 1980, which originally introduced the effect, did so in the context of a model in which relative market sizes were purely exogenous. Recent work by Davis and Weinstein (1997) has attempted to measure the empirical importance of the home market effect in patterns of international trade and has found surprisingly strong impacts.

5. Because $0 \leq Z \leq 1$, the coefficient on $dw/w$ is positive.

*This page intentionally left blank*

# 5 Core and Periphery

In the last chapter we laid out some basic machinery for modeling a monopolistically competitive economy—in essence, a set of technical tricks that allow us to handle the problems of market structure posed by the assumption that there are increasing returns at the level of the individual firm. We are now in a position to use that machinery to develop our first model of economic geography.

The analysis we introduce here is not intended to be realistic. Aside from the basic artificiality of the Dixit-Stiglitz model of monopolistic competition (an artificiality that is, alas, a necessary part of nearly all the models in this book), in this chapter we make a number of additional unrealistic assumptions that we drop or modify in later chapters. Our purpose here is to show, as clearly and simply as possible, how the interactions among increasing returns at the level of the firm, transport costs, and factor mobility can cause spatial economic structure to emerge and change. Some of the conclusions from this first pass turn out to be sensitive to those assumptions, but let us postpone that discussion until later. For now, let us simply get into the model.

## 5.1 Assumptions

We consider an economy of the type set out in chapter 4. It has two sectors, monopolistically competitive manufacturing $M$ and perfectly competitive agriculture $A$. Each of these sectors employs a single resource, workers and farmers respectively, and we assume that each of these sector-specific factors is in fixed supply.

The geographical distribution of resources is partly exogenous, partly endogenous. Let there be $R$ regions. The world has $L^A$ farmers, and each region is endowed with an exogenous share of this world agricultural labor force denoted $\phi_r$. The manufacturing labor force, by contrast, is mobile over time; at any point in time we denote the share of region $r$ in the world worker supply $L^M$ by $\lambda_r$. It is convenient to choose units[1] so that $L^M = \mu$, $L^A = 1 - \mu$.

Transport costs among regions take a very special form. Manufactured goods are subject to iceberg transport costs of the form introduced in chapter 4; if one unit of a good is shipped from $r$ to $s$, only $1/T_{rs}$ units arrive. Shipment of agricultural goods, by contrast, is assumed costless. This is a very unrealistic assumption: In the real world, the cost of transporting one dollar's worth of raw materials is normally higher than that of transporting a dollar's worth of manufactured goods! However, assuming costless transport of food makes our life much simpler for the moment; we turn to the consequences of dropping that assumption in chapter 7.

Because agricultural goods can be freely transported, and because these goods are produced with constant returns, agricultural workers have the same wage rate in all regions. We use this wage rate as the numeraire, so $w_r^A = 1$. Wages of manufacturing workers, however, may differ both in nominal and in real terms. Let us define $w_r$ and $\omega_r$ to be the nominal and real wage rate, respectively, of manufacturing workers in region $r$.

What determines how workers move between regions? Rather than try to produce a sophisticated theory of dynamics, we simply assume that they move toward regions that offer high real wages and away from regions that offer below-average real wages. Specifically, we define the average real wage as

$$\bar{\omega} = \sum_r \lambda_r \omega_r \tag{5.1}$$

and assume the ad hoc dynamics[2]

$$\dot{\lambda}_r = \gamma(\omega_r - \bar{\omega})\lambda_r. \tag{5.2}$$

(Notice that the extra $\lambda_r$ is necessary to ensure that the changes in all region's shares sum to 0.)

In our model, then, the distribution of manufacturing across regions is given at any point in time but evolves over time to the extent that real wages differ across regions. Regional real wages, however, themselves depend on the distribution of manufacturing, so we turn next to that dependence.

## 5.2 Instantaneous Equilibrium

There are a number of different ways to describe the determination of equilibrium at a point in time. We find it most useful to think of that equilibrium as the simultaneous solution of 4R equations, which determine the income of each region, the price index of manufactures consumed in that region, the wage rate of workers in that region, and the real wage rate in that region.

### 5.2.1 *Income*

The income equation is simple. Because transportation of agricultural goods is costless, agricultural workers earn the same wage everywhere, equal to 1 because it is the numeraire. Recalling that we have chosen units so that there are $\mu$ manufacturing workers and $1 - \mu$ agricultural workers in total, the income of region $r$ is

$$Y_r = \mu\lambda_r w_r + (1 - \mu)\phi_r. \tag{5.3}$$

### 5.2.2 *Price Index*

The second ingredient is the price index of manufactures in each location, which is as constructed in chapter 4 and given in equation (4.34). Because the number of manufacturing workers in location $s$ is $L_s^M = \mu\lambda_s$, the price index becomes

$$G_r = \left[\sum_s \lambda_s (w_s T_{sr})^{1-\sigma}\right]^{1/1-\sigma}. \tag{5.4}$$

Equation (5.4) exhibits the price index effect that we saw in chapter 4. Suppose that wages in different regions were the same. Then it is apparent from looking at the equation that the price index in $r$ would tend to be lower, the higher the share of manufacturing that is in regions with low transport costs to $r$. In particular, were there only two regions, a shift of manufacturing into one of the regions would tend, other things equal, to lower the price index in that region—and thus make the region a more attractive place for manufacturing workers to be. This is a version of the forward linkages that we discussed briefly in chapter 3, and turns out to be one of the forces that may lead to emergence of geographical structure in the economy.

### 5.2.3 *Nominal Wages*

As we saw in chapter 4, it is possible to derive the level of wages at which manufacturing in region $r$ breaks even. This wage equation is given by equation (4.35), which we restate as,

$$w_r = \left[ \sum_s Y_s T_{rs}^{1-\sigma} G_s^{\sigma-1} \right]^{1/\sigma}. \tag{5.5}$$

Like the equation for the price index, this equation is worth looking at for a moment. Suppose that the price indexes in all regions were similar. Then (5.5) would say that the nominal wage rate in region $r$ tends to be higher if incomes in other regions with low transport costs from $r$ are high. The reason, of course, is that firms can afford to pay higher wages if they have good access to a larger market. Thus our model exhibits a form of the backward linkages that drove the base-multiplier model sketched out in chapter 3; these reinforce the forward linkages described above.

### 5.2.4 *Real Wages*

Finally, it is straightforward to define the real wages of workers: Because manufactured goods receive a share $\mu$ of their expenditure, we have

$$\omega_r = w_r G_r^{-\mu}. \tag{5.6}$$

The nominal wage is deflated by the cost-of-living index, as in (4.28), but with the price of agriculture equal to unity everywhere.

### 5.2.5 *Determination of Equilibrium*

This model's instantaneous equilibrium can be thought of as determined by the simultaneous solution of the equations for income (5.3), the equations for price indices (5.4), the wage equations (5.5), and the real-wage equations (5.6): 4R equations in all.[3] Obviously we cannot say much about the solution of these equations in the general case. We can, however, get considerable insight by examining an obvious special case: that of a two-region economy in which agriculture is evenly divided between regions. In that special case, the obvious question is whether manufacturing is equally divided between the two regions or concentrated in one region: that is, whether the economy becomes divided between a manufacturing "core" and an agricultural "periphery." This special case has therefore come to be known as the *core-periphery model*; let us see how it works.

## 5.3 The Core-Periphery Model: Statement and Numerical Examples

The core-periphery model is the special case of the model described above when there are only two regions and agriculture is evenly divided between those two regions. This means that we need not explicitly write out shares of agriculture, because they are both $\frac{1}{2}$; and we can also simplify notation slightly by letting $T$ be the transport cost between the two regions and letting an unsubscripted $\lambda$ represent region 1's share of manufacturing (with $1 - \lambda$ representing region 2's share). Thus there are eight equations for instantaneous equilibrium:

$$Y_1 = \mu\lambda w_1 + \frac{1-\mu}{2}, \tag{5.7}$$

$$Y_2 = \mu(1-\lambda)w_2 + \frac{1-\mu}{2}, \tag{5.8}$$

$$G_1 = [\lambda w_1^{1-\sigma} + (1-\lambda)(w_2 T)^{1-\sigma}]^{1/1-\sigma}, \tag{5.9}$$

$$G_2 = [\lambda(w_1 T)^{1-\sigma} + (1-\lambda)w_2^{1-\sigma}]^{1/1-\sigma}, \tag{5.10}$$

$$w_1 = [Y_1 G_1^{\sigma-1} + Y_2 G_2^{\sigma-1} T^{1-\sigma}]^{1/\sigma}, \tag{5.11}$$

$$w_2 = [Y_1 G_1^{\sigma-1} T^{1-\sigma} + Y_2 G_2^{\sigma-1}]^{1/\sigma}, \tag{5.12}$$

$$\omega_1 = w_1 G_1^{-\mu}, \tag{5.13}$$

and

$$\omega_2 = w_2 G_2^{-\mu}. \tag{5.14}$$

This model still does not look particularly tractable: eight simultaneous nonlinear equations! We will see shortly, however, that the core-periphery model does indeed yield clear analytical results to the determined economist. However, to know what kind of results to look for, it is very helpful to look first at some numerical examples.

Figures 5.1, 5.2, and 5.3 plot $\omega_1 - \omega_2$, the difference between the two regions' real wage rates in manufacturing, against $\lambda$, the region 1 share

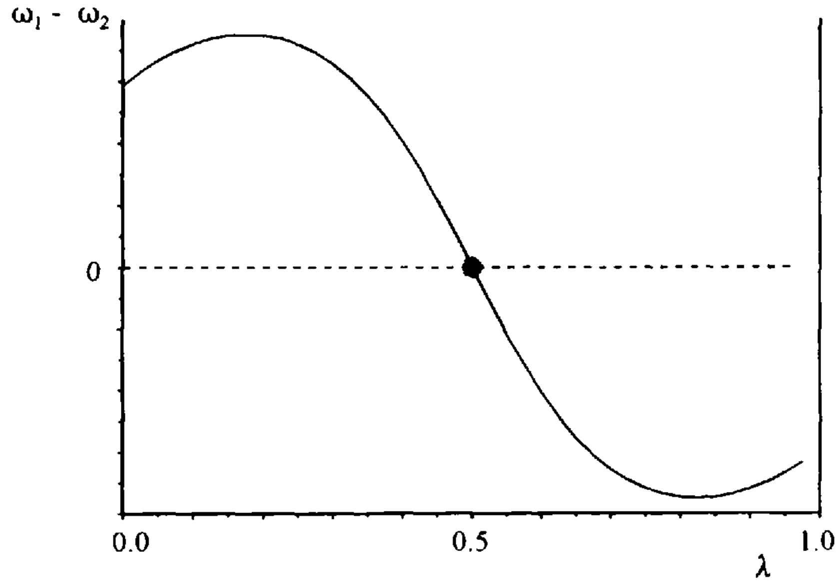


**Figure 5.1**
Real wage differentials, $T = 2.1$

of manufacturing. All three figures are calculated for $\sigma = 5$, $\mu = 0.4$. However, the transport cost $T$ is different in each: Figure 5.1 shows a high transport cost case, $T = 2.1$, figure 5.2 a low case, $T = 1.5$, and figure 5.3 an intermediate case, $T = 1.7$.

In figure 5.1, the wage differential is positive if $\lambda$ is less than $1/2$, negative if $\lambda$ is greater than $1/2$. This means that if a region has more than

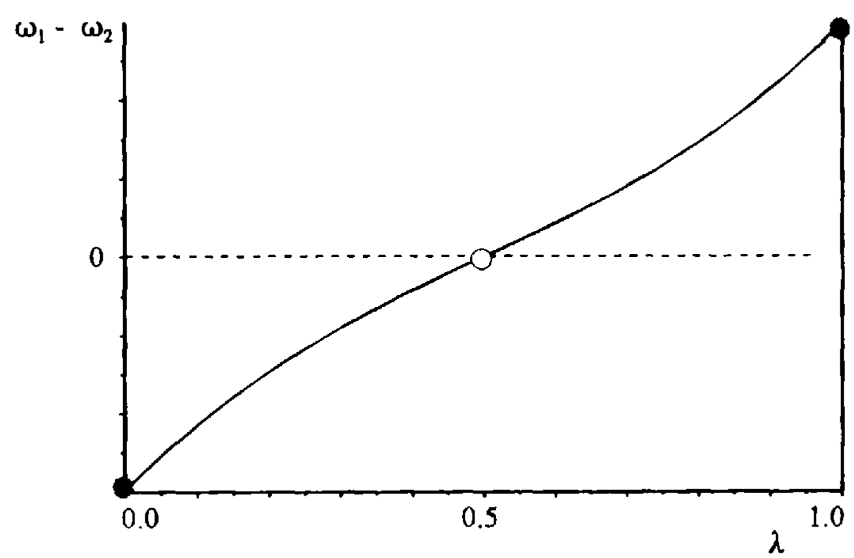


**Figure 5.2**
Real wage differentials, $T = 1.5$

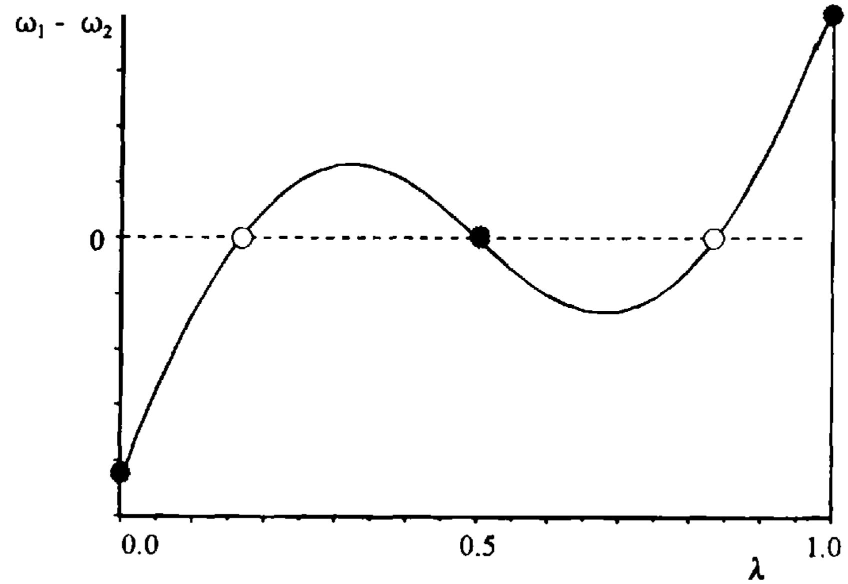


**Figure 5.3**
Real wage differentials, $T = 1.7$

half the manufacturing labor force, it is less attractive to workers than the other region. Clearly, in this case the economy converges to a long-run symmetric equilibrium in which manufacturing is equally divided between the two regions.

In figure 5.2, by contrast, the wage differential slopes strictly upward in $\lambda$: The higher the share of manufacturing in either region, the more attractive the region becomes. This upward slope results, of course, from the two linkage effects discussed in section 5.2: Other things equal, a larger manufacturing labor force makes a region more attractive both because the larger local market leads to higher nominal wages (backward linkage) and because the larger variety of locally produced goods lowers the price index (forward linkage). The important point here is that although an equal division of manufacturing between the two regions is still an equilibrium, it is now unstable: If one region should have even a slightly larger manufacturing sector, that sector would tend to grow over time while the other region's manufacturing shrank, leading eventually to a core-periphery pattern with all manufacturing concentrated in one region.

Finally, figure 5.3, for an intermediate level of transport costs, shows a more complicated picture. The symmetric equilibrium is now locally stable, as in figure 5.1. However, two unstable equilibria now flank it: If $\lambda$ starts from either a sufficiently high or a sufficiently low initial

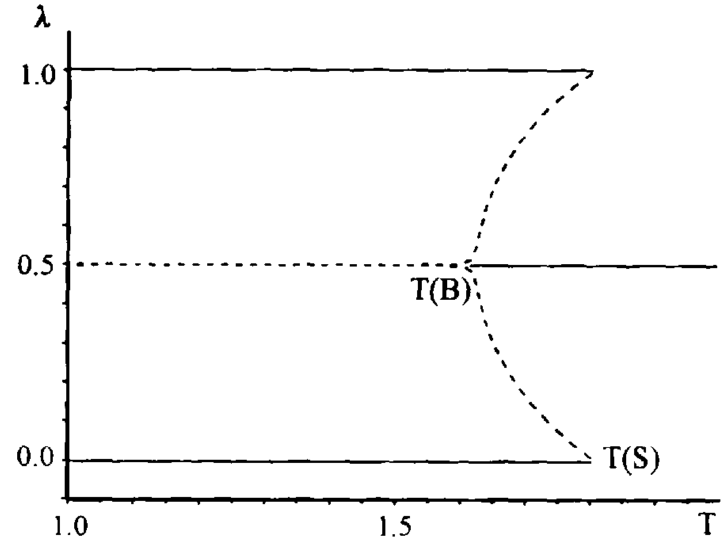


**Figure 5.4**
Core-periphery bifurcation

value, the economy converges not to the symmetric equilibrium but to a core-periphery pattern with all manufacturing in only one region. This picture then has five equilibria: three stable (the symmetric equilibrium and manufacturing concentration in either region) and two unstable.

From these three cases it is straightforward to understand figure 5.4, which shows how the types of equilibria vary with transport costs. As in figure 3.1, solid lines indicate stable equilibria, broken lines unstable. At sufficiently high transport costs, there is a unique stable equilibrium in which manufacturing is evenly divided between the regions. When transport costs fall below some critical level, new stable equilibria emerge in which all manufacturing is concentrated in one region. When they fall below a second critical level, the symmetric equilibrium becomes unstable.

The similarities to the base-multiplier model are clear. In particular, there are two critical points. The sustain point (labeled as point $T(S)$ in figure 5.4) is the point at which a core-periphery pattern, once established, can be sustained. And the break point $T(B)$ is the point at which symmetry between the regions must be broken because the symmetric equilibrium is unstable.

We can also now see how to approach the model analytically. We want to determine the conditions under which a core-periphery pattern is possible—the sustain point—and the conditions under which it is necessary—the break point.

## 5.4 When Is a Core-Periphery Pattern Sustainable?

Suppose we start with all manufacturing concentrated in one region, say region 1. To determine whether this is an equilibrium, we ask whether a small group of workers moving from region 1 to region 2 would receive a higher real wage than that received by the workers remaining behind. If so, a core-periphery geography is not an equilibrium: Manufacturing will shift over time to the peripheral region. If not, a core-periphery pattern *is* an equilibrium: The concentration of manufacturing will be self-sustaining.

In short, to assess whether a core-periphery pattern is sustainable, we need to posit a situation in which $\lambda = 1$ and ask whether in that case $\omega_1$ is greater or less than $\omega_2$. If $\omega_1 \geq \omega_2$, then the core-periphery pattern is sustainable, because manufacturing workers will not move out of region 1.[4]

Suppose we set $\lambda = 1$. Simply guess that $w_1 = 1$; in that case

$$Y_1 = (1 + \mu)/2, \qquad Y_2 = (1 - \mu)/2,$$
$$G_1 = 1, \qquad G_2 = T, \tag{5.15}$$

and we can then confirm from (5.11) that $w_1 = 1$ is indeed an equilibrium value. Notice that income is higher in location 1 than in location 2: It has all the income generated by manufacturing employment. And notice also that the price index is higher in 2 than in 1, because location 2 has to import all its manufactures. These two facts are the basis of the backward and forward linkages that support the core-periphery pattern.

Because $w_1 = 1$ and $G_1 = 1$, it then follows that $\omega_1 = 1$ as well. So all we need to do is determine $\omega_2$ and see whether it is more or less than 1. Substituting into the nominal and real wage equations, (5.12) and (5.14), we have

$$\omega_2 = T^{-\mu}\left[\frac{1+\mu}{2}T^{1-\sigma} + \frac{1-\mu}{2}T^{\sigma-1}\right]^{1/\sigma}. \tag{5.16}$$

Equation (5.16) looks complex but lends itself immediately to interpretation in terms of forward and backward linkages. The first term in the equation, $T^{-\mu}$, represents the forward linkage: It comes from the fact that the price index in region 2 is $T$ times as high as that in region 1 because manufactured goods must be imported. The term is less than unity: Having to import manufactures makes location 2 relatively expensive, and therefore unattractive, as a place for manufacturing workers to locate.

The second term represents the nominal wage at which a firm locating in 2 would break even. The income level in location 1 is weighted by $T^{1-\sigma}$, which is less than unity; this weighting is a result of the transport cost disadvantage that a firm in 2 would face in supplying location 1. The income level in 2 is symmetrically weighted by $T^{\sigma-1}$, greater than unity, because of the transport cost disadvantage that firms in 1 bear in supplying location 2. Although these effects are symmetric, they mean that a firm considering locating in 2 does well in the smaller market but badly in the larger; hence there is a backward linkage via demand from the concentration of production to the nominal wage rate firms can afford to pay.

What does equation (5.16) tell us about the sustainability of the core-periphery structure? First, consider the role of transportation costs. To do this it is helpful to rewrite (5.16), in the form

$$\omega_2^{\sigma} = \frac{1+\mu}{2} T^{1-\sigma-\mu\sigma} + \frac{1-\mu}{2} T^{\sigma-1-\mu\sigma}. \tag{5.17}$$

Clearly, when $T = 1$ (no transport costs), $\omega_2 = 1$: Location is irrelevant. If we consider a small transport cost increase from that point, we find (by totally differentiating (5.17) and evaluating the derivative at $T = 1$, $\omega_2 = 1$)

$$\frac{d\omega_2}{dT} = \frac{\mu(1-2\sigma)}{\sigma} < 0. \tag{5.18}$$

At small levels of transport costs agglomeration must therefore be sustainable, because $\omega_2 < 1 = \omega_1$.

Suppose, on the other hand, that we consider very large $T$. The first term in (5.17) clearly becomes arbitrarily small. There are two possibilities for the second term. If $(\sigma - 1) - \mu\sigma < 0$, then this term also becomes arbitrarily small, so $\omega_2$ tends to 0. But recall that $(\sigma - 1)/\sigma > \mu$ is the no-black-hole condition we discussed in chapter 4. It is now clear how to interpret this alternative case: If $(\sigma - 1)/\sigma < \mu$, the agglomeration forces are so strong that a core-periphery pattern is always an equilibrium.

If the no-black-hole condition holds, so $(\sigma - 1) - \mu\sigma > 0$, then the second term in (5.17) becomes arbitrary large. Figure 5.5 shows what happens in this case. The curve defining $\omega_2$ as a function of $T$ slopes downward in the vicinity of $T = 1$, but then turns upward. The point where it crosses 1 defines the sustain value of $T$: Below this value the core-periphery pattern is an equilibrium, and above it, it is not.

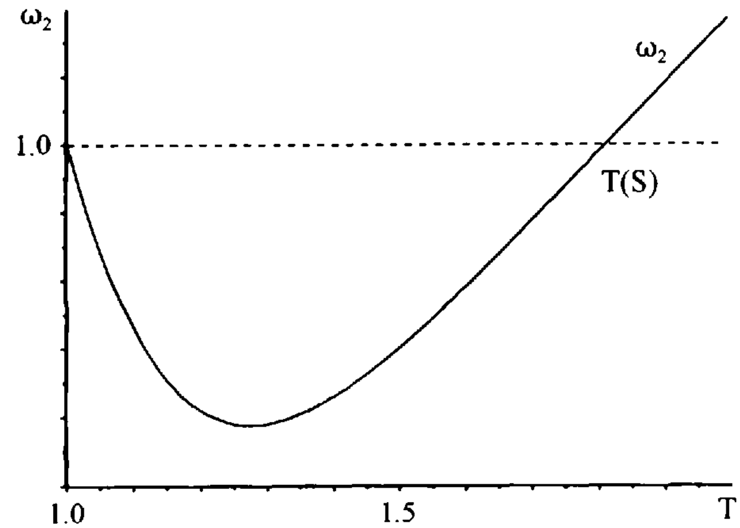


**Figure 5.5**
Sustain point

How does this sustain value of $T$ depend on parameters? A lower value of $\sigma$ (and $\rho$) stretches the curve in figure 5.5 to the right, raising the range of values of $T$ at which the core-periphery structure is sustainable. Conversely, as $\sigma$ (and $\rho$) get very large, the sustain value of $T$ becomes close to unity, because very small transport costs then choke off trade, so that manufacturing must operate in both locations to supply local demand.

It is also easy to see that the likelihood that $\omega_2 < 1$, so that a core-periphery equilibrium exists, depends on a sufficiently large role of manufacturing in the economy. Suppose that $\mu = 0$, so (5.16) reduces to $[(T^{1-\sigma} + T^{\sigma-1})/2]^{1/\sigma}$. Providing $T > 1$, this is always greater than 1, so there cannot be a core-periphery pattern. More generally, at lower values of $\mu$ the curve in figure 5.5 is rotated upward, reducing the range of values of $T$ that sustain the core-periphery geography. When the manufacturing sector is large, and so can generate significant forward linkages via supply and backward linkages via demand, it generates sufficient centripetal forces to sustain a concentrated equilibrium over a wide range of transportation costs.

## 5.5 When Is the Symmetric Equilibrium Broken?

We saw from figures 5.1–5.4 that the symmetric equilibrium is stable for high enough values of transport costs but becomes unstable at low values. The figures also illustrate how we can go about finding this break point: It occurs when the model's parameters are such that the $\omega_1 - \omega_2$ curve is horizontal at the symmetric equilibrium. To find it, we have to differentiate totally the equilibrium—characterized by equations (5.7)–(5.14)—with respect to $\lambda$, and hence find the equilibrium response $d(\omega_1 - \omega_2)/d\lambda$.

This task is not as formidable as it sounds, because we differentiate around the symmetric equilibrium (as we did in section 4.5). At this equilibrium we know the values of all the endogenous variables of the model. They are

$$\lambda = 1/2, \quad Y_1 = Y_2 = 1/2, \quad w_1 = w_2 = 1,$$
$$G_1^{1-\sigma} = G_2^{1-\sigma} = \left[\frac{1 + T^{1-\sigma}}{2}\right]. \tag{5.19}$$

These can be checked by recalling that $\lambda = 1/2$ is the definition of the symmetric equilibrium, and then seeing that equations (5.7)–(5.12) are satisfied at these values. In the subsequent discussion, we write the values of variables at the symmetric equilibrium as $G$, $Y$, etc., dropping the location subscript.

The fact that we are differentiating around the symmetric equilibrium brings another simplification. Any change in an endogenous variable in region 1 is matched by an equal but opposite sign change in the corresponding variables in region 2. This means that we do not have to keep track of region 1 and region 2 variables separately. Instead we write $dY \equiv dY_1 = -dY_2$, and similarly for changes in other variables.

To see how this works, consider the total derivative of the income equations, (5.7) and (5.8). These are

$$dY_1 = \mu w_1 d\lambda + \mu\lambda dw_1, \quad dY_2 = -\mu w_2 d\lambda + \mu(1 - \lambda)dw_2, \tag{5.20}$$

but around the symmetric equilibrium, these can be described by the single equation,

$$dY = \mu d\lambda + \frac{\mu}{2} dw. \tag{5.21}$$

Proceeding analogously, the total differential of the price indices (5.9) and (5.10) is

$$(1 - \sigma)\frac{dG}{G} = G^{\sigma-1}(1 - T^{1-\sigma})\left[d\lambda + \frac{(1-\sigma)dw}{2}\right]. \tag{5.22}$$

The term $1 - T^{1-\sigma}$ appears repeatedly because it captures the effects of an increase in a variable in one region and corresponding decrease in the other. We can simplify by defining a variable $Z$,

$$Z \equiv \frac{[1 - T^{1-\sigma}]}{[1 + T^{1-\sigma}]} = \frac{[1 - T^{1-\sigma}]}{2G^{1-\sigma}}, \tag{5.23}$$

where the second equation comes from the value of G at the symmetric equilibrium, (5.19). We have already seen $Z$ in chapter 4. It is an index of trade barriers, taking value 0 when there are no transport costs ($T = 1$) and 1 when transport costs are prohibitive ($T \to \infty$). Using this expression, (5.22) becomes

$$\frac{dG}{G} = \frac{2Z}{1 - \sigma} d\lambda + Zdw. \tag{5.24}$$

Applying the same techniques, the total differential of the wage equations (5.11) and (5.12) and real wage equations (5.13) and (5.14) are

$$\sigma dw = 2ZdY + (\sigma - 1)Z\frac{dG}{G}, \tag{5.25}$$

$$G^{\mu}d\omega = dw - \mu\frac{dG}{G}. \tag{5.26}$$

We want to find $d\omega/d\lambda$, and we can do so by eliminating $dG/G$, $dw$, and $dY$ from equations (5.21), (5.24), (5.25) and (5.26). This is a long but straightforward set of substitutions, details of which are given in the appendix. The required expression giving the change in real wages caused by a movement of workers is

$$\frac{d\omega}{d\lambda} = 2ZG^{-\mu}\left(\frac{1 - \rho}{\rho}\right)\left[\frac{\mu(1 + \rho) - Z(\mu^2 + \rho)}{1 - \mu Z(1 - \rho) - \rho Z^2}\right], \tag{5.27}$$

where we have replaced $\sigma$ by $\rho$, $\rho = (\sigma - 1)/\sigma$. (The expression is slightly more compact this way.) The symmetric equilibrium is stable if $d\omega/d\lambda$ is negative and unstable if it is positive. We can see easily that the denominator is positive, because $Z$ lies in the interval 0 (free trade) to 1 (autarky) and both $\mu$ and $\rho$ are less than unity. The sign of the expression therefore depends on the numerator of the term in square brackets. When $Z$ is close to 0 (transport costs are low), this is certainly positive, so the symmetric equilibrium is unstable. Increasing $Z$ reduces the size of the numerator, and when $Z = 1$ (so that transport

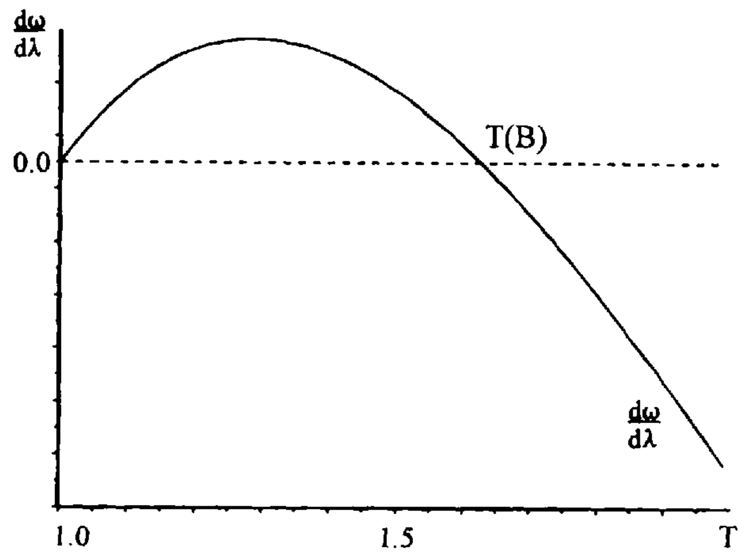


**Figure 5.6**
Break point

costs are infinite), the numerator is positive if $\rho < \mu$, or changes sign and becomes negative if $\rho > \mu$. This gives exactly the same two cases as we saw in our discussion of the sustain point. The symmetric equilibrium is always unstable if the no-black-hole condition fails, $\rho < \mu$. Otherwise, the symmetric equilibrium is stable at sufficiently high levels of transport costs.

Figure 5.6 gives $d\omega/d\lambda$ as a function of $T$, for the case in which the no-black-hole condition holds, $\rho > \mu$. At free trade ($T = 1, Z = 0$) relocation of labor ($d\lambda$) has no effect on the regional real wage differentials ($d\omega$), essentially because with no transport costs the regions are not economically distinct. At intermediate levels of $T$, the forward and backward linkages associated with the relocation of workers raise the real wage in the location to which workers are moving, so $d\omega/d\lambda > 0$ and the symmetric equilibrium is unstable. As $T \rightarrow \infty$ (autarky), an increase in one region's industrial labor force reduces the real wage there, because it increases the supply of manufactures that cannot now be exported. The break point is at point $T(B)$, where $d\omega/d\lambda$ changes sign.

We can use equation (5.27) to derive an explicit expression for the break point value of $T$. Setting the numerator of the term in square brackets equal to 0 and using the definition of $Z$, this expression is

$$\frac{d\omega}{d\lambda} = 0 \quad \text{if} \quad T^{\rho/(1-\rho)} = \frac{(\rho + \mu)(1 + \mu)}{(\rho - \mu)(1 - \mu)}. \tag{5.28}$$

**Table 5.1**
Critical values of $T$: Break points $T(B)$ and sustain points $T(S)$

| | $\mu = 0.2$ | | $\mu = 0.4$ | | $\mu = 0.6$ | |
|---|---|---|---|---|---|---|
| | $T(B)$ | $T(S)$ | $T(B)$ | $T(S)$ | $T(B)$ | $T(S)$ |
| $\sigma = 3$ ($\rho = 0.67$) | 1.67 | 1.72 | 3.05 | 4.47 | 8.72 | 3124.7 |
| $\sigma = 5$ ($\rho = 0.8$) | 1.26 | 1.27 | 1.63 | 1.81 | 2.30 | 5.00 |
| $\sigma = 7$ ($\rho = 0.86$) | 1.158 | 1.164 | 1.36 | 1.44 | 1.68 | 2.44 |

The parameter values satisfying this equation define the break values at which the symmetric equilibrium becomes unstable. What do we know about these values? First, the break value of $T$ is unique and, if we maintain the the no-black-hole condition, it occurs at a positive level of trade costs, $T > 1$. Second, the break value is increasing in $\mu$: The larger the share of manufacturing workers in the economy, the greater the range of $T$ in which the symmetric equilibrium is unstable. It is also decreasing in $\rho$ (and therefore also $\sigma$), a low $\rho$ corresponding to a high degree of product differentiation and large price cost markups, and hence strong forward and backward linkages.

The dependence of break and sustain points on parameters is most easily summarized using some numerical examples. Each cell in table 5.1 reports first the break point and then the sustain point at different values of $\mu$ and $\sigma$. Because both critical values are increasing in $\mu$ and decreasing in $\sigma$, the range of transport costs in which the core-periphery geography occurs is greater the larger is the share of manufactures in the economy, and the larger are firms' price cost markups. Notice that the sustain point always occurs at a higher value of $T$ than does the break point, because the bifurcation is a tomahawk, as illustrated in figure 5.4.

## 5.6 Implications and Conclusions

One could say that the dynamic spatial model laid out in this chapter—and its two-region core-periphery version in particular—plays much the same role in our approach to economic geography that the $2 \times 2 \times 2$ model plays in constant-returns trade theory. That is, it is a model simple enough to yield readily to analysis, yet enough is going on in the model that it yields a number of suggestive and interesting conclusions. From it we learn how economies of agglomeration can emerge from the interactions among economies of scale at the level of the individual producer, transport costs, and factor mobility. We also get a clear illustration both of the tension between centripetal and centrifugal forces and of the potential for discontinuous change that tension creates. Finally, we get a first view of the way that dynamic analysis can serve as a powerful tool of simplification, allowing us to sort through and in the end limit the possibilities static analysis suggests.

For all its virtues, however, the core-periphery model—like the $2 \times 2 \times 2$ model in trade!—can be a bit too seductive: Some of its implications turn out to be sensitive to assumptions one would not want to defend. In the next two chapters we therefore push out the model's boundaries, first by considering the implications of multiple regions, then by turning to a more realistic structure of transport costs.

## Appendix: Symmetry Breaking

We want to find the effect of a change $d\lambda$ on the symmetric equilibrium. In the text we totally differentiated the equilibrium around the symmetric point and derived:

$$dY = \mu d\lambda + \frac{\mu}{2} dw, \tag{5.21}$$

$$\frac{dG}{G} = \frac{2Z}{1 - \sigma} d\lambda + Z dw, \tag{5.24}$$

$$\sigma dw = 2Z dY + (\sigma - 1) Z \frac{dG}{G}, \tag{5.25}$$

and

$$G^{\mu} d\omega = dw - \mu \frac{dG}{G}. \tag{5.26}$$

We use (5.21) to eliminate $dY$ from (5.25), and then write (5.24) and (5.25) as the system

$$\begin{bmatrix} 1 & -Z \\ Z & \dfrac{\sigma - \mu Z}{1 - \sigma} \end{bmatrix} \begin{bmatrix} \dfrac{dG}{G} \\ dw \end{bmatrix} = \begin{bmatrix} \dfrac{2Z}{1 - \sigma} d\lambda \\ \dfrac{2Z\mu}{1 - \sigma} d\lambda \end{bmatrix}, \tag{5A.1}$$

from which

$$\frac{dG}{G} = \frac{d\lambda}{\Delta} \frac{2\sigma Z}{(1 - \sigma)^2} [1 - \mu Z], \tag{5A.2}$$

and

$$dw = \frac{d\lambda}{\Delta} \frac{2Z}{(1 - \sigma)} [\mu - Z], \tag{5A.3}$$

with the determinant, $\Delta$, taking the form

$$\Delta = \frac{Z^2(1 - \sigma) - Z\mu + \sigma}{1 - \sigma}. \tag{5A.4}$$

Using these in the equation for $d\omega$, (5.26), gives

$$\frac{d\omega}{d\lambda} = \frac{2ZG^{\mu}}{\sigma - 1} \left[ \frac{\mu(2\sigma - 1) - Z(\sigma(1 + \mu^2) - 1)}{\sigma - \mu Z - Z^2(\sigma - 1)} \right]. \tag{5A.5}$$

Equation (5.27) of the text is derived by replacing $\sigma$ with $1/(1 - \rho)$ and using the definition of $Z$, (5.23).

## Notes

1. We can choose units in which to measure each of the two different types of labor, in addition to the choice of units for output and for firms that we made in chapter 4.

2. Although we have no deep justification for this particular formulation, these dynamics are equivalent to the "replicator dynamics" routinely used in evolutionary game theory. Indeed, our model can, if one likes, be regarded as an evolutionary game. See Weibull 1995.

3. When we work with numerical examples, it is necessary to solve this system. Although it is not necessarily the most efficient procedure, simple iteration generally works: That is, start with guesses at $w_r$, $r = 1, \ldots, R$; calculate in sequence the implied $Y$ and $G$ vectors; calculate new values of $w$; and repeat until convergence.

4. We could of course pursue the symmetric case: set $\lambda = 0$ and see if $\omega_1 \leq \omega_2$.