Institut for Nationaløkonomi

Handelshøjskolen i København

Working paper 17-98

UNDERSTANDING ´INCREASING RETURNS TO SCALE AND ECONOMIC GEOGRAPHY´

Dieter M. Urban

Department of Economics - Copenhagen Business School

Nansensgade 19, 5. DK - 1366 København K.

Understanding

'Increasing Returns to Scale and Economic Geography'

Part I: A graphical exposition

Part II: An analytical solution

by

Dieter M. Urban
Copenhagen Business School

Abstract part I: This paper provides a simple graphical exposition and a rigorous analytical method for monopolistic competition, increasing returns to scale, geography and trade models with transport costs which explain agglomeration or convergence of industries. In the main text, the agglomeration and convergence forces are graphically exposed, whereas the appendix provides the analytical treatment of the model. New light is shed on the Dixit-Stiglitz-Krugman model by an analogy to a heterogeneous agent pure exchange economy.

Abstract part II: This paper provides an analytical solution to the Krugman (1991a) model explaining industry agglomeration. It is shown there exists a unique short-run equilibrium and multiple long-run equilibria. The latter proves the existence of a "poverty trap" in this model: depending on the initial level of industries we will either see industries spreading evenly in the plane, or moving away from one of the regions. However, it is also shown that this "poverty trap" will not appear if the economy starts developing from an equal distribution of industries.

Mailing Address: Dieter M. Urban, Department of Economics, Copenhagen Business School, Nansensgade 19, 5th floor, DK-1366 Copenhagen K, Denmark.

Understanding

'Increasing Returns to Scale and Economic Geography'

Part I: A graphical exposition

by

Dieter M. Urban
Copenhagen Business School

Abstract part I: This paper provides a simple graphical exposition and a rigorous analytical method for monopolistic competition, increasing returns to scale, geography and trade models with transport costs which explain agglomeration or convergence of industries. In the main text, the agglomeration and convergence forces are graphically exposed, whereas the appendix provides the analytical treatment of the model. New light is shed on the Dixit-Stiglitz-Krugman model by an analogy to a heterogeneous agent pure exchange economy.

September 25, 1998

Understanding Geography and Trade*

By

Dieter M. URBAN
Department of Economics
Copenhagen Business School, Denmark.

Abstract

This paper provides a simple graphical exposition and a rigorous analytical method for monopolistic competition, increasing returns to scale, geography and trade models with transport costs which explain agglomeration or convergence of industries. In the main text, the agglomeration and convergence forces are graphically exposed, whereas the appendix provides the analytical treatment of the model. New light is shed on the Dixit-Stiglitz-Krugman model by an analogy to a heterogeneous agent pure exchange economy.

JEL Classification: F12, R60.
Keywords: convergence, agglomeration, poverty trap.

Mailing Address: Dieter M. Urban, Department of Economics, Copenhagen Business School, Nansensgade 19, 5th floor, DK-1366 Copenhagen K, Denmark.

—————————————————

*The author thanks Kjetil Bjorvatn, Niels Blomgren-Hansen, Pascalis Raimondos-Moller, and participants of the EEA conference in Istanbul 1996, and of seminars at Copenhagen Business School, Copenhagen University, PennState University, and Uppsala University for their comments. Usual caveats apply.

1 Introduction

The economic geography literature works with several \workhorses". The most prominent one is perhaps the Dixit-Stiglitz-Krugman monopolistic competition increasing returns to scale model with transport costs explaining industry agglomeration in space. There are two general versions of it: the regional economic model by Krugman (1980,1991) and the international trade model by Krugman and Venables (1995). The regional economic model assumes migration of workers between regions dragging industries with them. Across nations, however, labour is less mobile. Hence, the international trade model assumes instead that labor is intersectorally mobile and agglomeration of industries occurs, if labor moves from a constant returns to scale agricultural sector to an increasing returns to scale manufacturing sector. However, the models are otherwise identical in structure and the driving forces for the di®erent kind of labor mobilities are very similar, too.

These models have two disadvantages. They are non-linear in structure and may entail multiple equilibria. Hence, most of the analysis is numerical with some suggestive analytical treatments. Rigorous algebra is in particular important for de¯ning conditions for which an even distribution of workers and industries turns into an uneven one (agglomeration condition). Krugman (1991) gives such a condition for total agglomeration, Venables (1995) and Krugman and Venables (1995) give a condition, for which the equal distribution is not a stable equilibrium (\algebra of symmetry-breaking").

Still, the analysis is far from both the rigorosity and the economic intuition of the driving e®ects which is for example provided in the Heckscher and Ohlin type of economic geography models.[1] Urban (1996) provides a complete analytical solution to the Krugman (1991) model and detects a \poverty trap" that was overlooked by the seminal article. Because Urban (1996) does not provide any economic intuition for his results, his paper is supplemented by this companion paper giving a simple graphical

[1] See, for example, Norman and Venables (1995).

exposition of the model.

This graphical apparatus gives a clear-cut intuition for the agglomeration and convergence forces embedded in the model, whereas the analytics of the seminal paper do not reveal this intuition. In particular, the seminal paper argues on the basis of home market and extent of competition e®ect. We con¯rm the mechanics of the home market e®ect, but show that the extent of competition e®ect is better described as a regional composition e®ect of goods (number of goods e®ect). Indeed, the model deviates from a standard neoclassical model only because of the endogeneity of product space. The graphical apparatus provides an exposition of the seminal model suitable for undergraduate classes.

A need for a clari¯cation of the mechanics of the Krugman (1991) model may also be derived from Davis (1998). He shows that a minor change in the model set-up - transport cost for agricultural goods - undermines the home market e®ect unexpectedly. Unfortunately, his formal proof - though elegant - adds little to the understanding of the mechanics of the model. Instead of his proof by contradiction, we follow the standard approach in gerneral equilibrium theory focusing on the excess demand system and the comparative static e®ects of worker migration on relative prices and wages. This approach allows us to compare the Krugman increasing returns to scale model directly with a standard neoclassical pure exchange economy.

The analogy to a pure exchange economy sheds new light on the interpretation of trade costs and region size in the model. The analogy may suggest to rethink the role of increasing returns in the Dixit-Stiglitz-Krugman model. This is somewhat in contrast to Fujita and Thisse (1996) who attribute a major role in the explanation of agglomeration economics to increasing returns. Whereas we will show that the mechanism of the Dixit-Stiglitz-Krugman model is analogue to a pure exchange economy with two types of heterogeneous agents and three goods with the following properties: 1) each type of consumer is equally endowed with one good, owes the world endowment of the second and nothing of the third. 2) The consumer strictly prefers the good he is well endowed with to the good he is not endowed with.

The main text gives an almost entirely graphical treatment filled with intuition, whereas the appendix contains the analytical treatment of the model. Section 2 repeats the familiar model set-up for convenience; section 3 gives the equilibrium conditions; section 4 draws an insightful analogy to a pure exchange economy; section 5 explains why there is agglomeration or convergence; and section 6 concludes.

2 The Model Set-Up

In this section, the basic structure of a typical geography and trade model as developed by Krugman (1991) is presented. The model has two regions with one increasing returns to scale sector (industry) and one constant returns to scale sector (agriculture) in each region. The increasing returns to scale sector is monopolistically competitive. Furthermore, transport costs for industrial goods introduce a geographical dimension into the model.

There are two types of consumers $j = 1, 2$, which are only different by there place of residence. Home region's consumers are indexed by 1, foreign region's consumers by 2. Regions are defined as areas for which it is costless to trade industrial goods within them, but costly to trade industrial goods across the border. Furthermore, there is no short run mobility of production factors across borders. However, a long run mobility of labor across borders is considered.

The two types of consumers j have identical Cobb-Douglas utility functions of the form

$$U_j = C_{Mj}^{\mu} C_A^{1-\mu}; \quad 0 < \mu \leq 1; \tag{1}$$

where C_A is consumption of the agricultural good produced with constant returns to scale and C_{Mj} is an aggregate basket of industrial goods produced in both regions under monopolistic competition and increasing returns to scale. The industrial goods basket C_{Mj} is further specified by a Dixit-Stiglitz (1977) subutility function:

$$C_{Mj} = \left[\sum_{i=1}^{n_1 + n_2} c_{ij}^{(\sigma - 1 = \sigma)} \right]^{\sigma = (\sigma - 1)} ; \ \sigma > 1: \tag{2}$$

The demand of consumer j for a single industrial firm i's product is denoted c_{ij}. There are n_1 firms in the home region and n_2 firms in the foreign. The number of firms is assumed to be sufficiently large. The elasticity of substitution between the industrial goods is denoted σ.

There is factor specificity for industrial production by workers and agricultural production by peasants. Peasants work according to a constant returns to scale technology in perfectly competitive markets. The price for agricultural products serves as numeraire and price equals wage.

Industrial workers have an increasing returns to scale technology of a simple structure: there is a fixed cost α and a constant marginal cost β for each firm i. The firm i uses L_{Mi} units of labour for producing x_i goods:

$$L_{Mi} = \alpha + \beta x_i: \tag{3}$$

Every firm produces a different variety in order to exploit potential monopoly profits. In equilibrium, firms will not succeed, however, because free costs of firm entry and exit will assure zero profits. Because firms are assumed to be symmetric, we drop the firm index i for convenience. However, we will use the index $i = 1; 2$ in order to distinguish the home and the foreign firms, respectively.

The number of industrial workers in region 1, L_1, and in region 2, L_2, are for simplicity assumed to add up to μ:[2]

$$L_1 + L_2 = \mu: \tag{4}$$

Without loss of generality, I define the domestic region (region 1) to be the smaller one, i.e. there are less industrial workers than in the foreign region (region 2). The

[2] See Krugman (1991), footnote 1, for a justification of this assumption.

total amount of peasants is $1 - \mu$, they are assumed to be equally distributed in both regions, and they are not mobile. Every worker and peasant supplies one unit of work and earns a salary 1, if peasant, and w_i, if worker in region i.

Finally, there are trade costs of the Samuelson iceberg-type τ, such that only a fraction τ of one produced unit of an industrial good arrives at its foreign destination ($0 < \tau \leq 1$). There are neither trade costs for goods delivered to domestic customers, nor trade costs for agricultural goods.

3 Equilibrium Conditions

In this section, we state some well-known economic relationships that stem from firm optimization, consumer optimization, the zero-profit condition, the labour market equilibrium condition and the goods market equilibrium conditions.

Given the usual assumption that the firm takes into account the impact of its pricing decision on its own demand, but not on other firms pricing decisions, then the well-known pricing rule for the firm holds:[3]

$$p_i = \gamma \cdot \beta \cdot w_i; \qquad (5)$$

where $\gamma \equiv \sigma/(\sigma - 1)$, and p_i denotes the mill price of region i firms' goods[4]. Prices are constant mark-ups over wages due to the assumptions of constant elasticities of substitution and constant marginal cost. Prices and wages are proportional. An increase in wages drives up prices and vice versa. Therefore, prices and wages can be used interchangeably, henceforth.

The optimal output of the firm is known to be determined by the zero profit condition:

[3] This equation is only an approximation which is fairly good for a large number of firms. (It does not imply that n_i has to be large, because n_i is normalized to number of firms per country population.) See the discussion in Yang and Heijdra (1993), Dixit and Stiglitz (1993), and d'Aspremont et al. (1996).

[4] The pricing decision for export goods p_i^{Export} requires the firm to demand the domestic price plus the additional transport cost: $p_i^{Export} = \gamma \cdot \beta \cdot w_i/\tau; \quad i = 1; 2$

$$x_i = \frac{\alpha \cdot (\sigma - 1)}{\beta} \equiv \bar{x}; \qquad (6)$$

The equilibrium output of the firm is independent of the number of workers or the number of firms in a region. In fact, it is an exogenously given constant $\bar{x}$. All the interesting effects, which drive the agglomeration or convergence process stem from the demand side. This is again a result of the simplifying assumption of constant marginal cost.[5]

The equilibrium number of firms per region follows from the labour market clearing conditions and the output decision:

$$n_i = \frac{L_i}{\alpha\sigma}. \qquad (7)$$

This is the third important economic relationship to be kept in mind. If the number of workers increases in a region, workers drag industries with them, and the number of goods increases proportionally.

Since profits are zero, aggregate income in a region is the sum of the income of all workers and peasants in that region:

$$y_i(w_i; L_i) = w_i \cdot L_i + \frac{1 - \mu}{2}. \qquad (8)$$

This implies also that the wage bill of all firms in a region equals the sales of all firms in that region.

$$y_i(p_i; n_i) = p_i \cdot n_i \cdot \bar{x} + \frac{1 - \mu}{2} \qquad (9)$$

These economic relationships are useful, because every behavioural equation of the model can be interchangeably expressed in terms of the two prices p_i and the number of goods per region n_i or equivalently in terms of the two wages w_i and the labour distribution L_i[6]. Hence, it will be completely sufficent to describe the model in terms of

[5] Hence, one effect, which one might think of, is missing in the model: the larger region does not have larger firm sizes and hence lower production cost under increasing returns to scale. See Krugman (1980), footnote 3, on this issue.

[6] From now on L denotes the amount of domestic workers and $\mu - L$ the amount of foreign workers. Respectively, n denotes the number of domestic goods, whereas the number of foreign goods is given by: $\frac{\mu}{\alpha\sigma} - n$.

prices and number of goods which will be determined by the goods market equilibrium conditions. Another example are the CES-price indices P_j that can be written in both ways as function of prices and number of goods or wages and the worker distribution:[7]

$$P_j = P_j(p_1, p_2, n_j) = P_j(w_1, w_2, L_j). \tag{10}$$

An increase in the number of domestic goods will lower the price index even at given prices, because there will be less goods to be paid transport cost for than before ($\partial P_j / \partial n_j < 0$ and $\partial P_j / \partial L_j < 0$). With these ingredients, a short run equilibrium can be defined as an equilibrium of the goods market, the labour market and zero firm profits at a given distribution of labour. Such an equilibrium can be found as the solution of the excess demand curves of the domestic firms $f(p_1, p_2, n)$, the foreign firms $f^*(p_1, p_2, n)$, and the agricultural sector $g(p_1, p_2, n)$[8,9]:

$$\begin{aligned} f(p_1, p_2, n) &: \quad D(p_1, P_1) \cdot y_1 + \frac{E(p_1, P_2) \cdot y_2}{\tau} = \bar{x} \qquad (11)\\ f^*(p_1, p_2, n) &: \quad D(p_2, P_2) \cdot y_2 + \frac{E(p_2, P_1) \cdot y_1}{\tau} = \bar{x} \\ g(p_1, p_2, n) &: \quad (1-\mu)(y_1 + y_2) = (1-\mu); \end{aligned}$$

where

$$D(p_i, P_i)$$

describes the fraction of region i's income spent on region i's firms and

$$E(p_i, P_j);\ i = j$$

describes the fraction of region j's income spent on imports from region i's firms.[10] The home regions' income y_1 and the foreign regions' income y_2 are defined according

[7] The explicit functional forms are given in appendix A.

[8] The agricultural sectors can be merged into one equilibrium condition, because there are no transport cost for agricultural goods.

[9] The explicit functional forms of the equation system (11) are given in appendix B.

[10] The explicit definition for $D(p_i, P_j)$ is:
$D(p_i, P_i) = \frac{\alpha\sigma p_i^{-\sigma}}{P_i^{1-\sigma}}$
and the explicit definition for $E(p_i, P_j)$ is:
$E(p_i, P_j) = \frac{\alpha\sigma (p_i/\tau)^{-\sigma}}{P_j^{1-\sigma}}$

to equation (9), and the price indices are defined according to equation (10). These functions describe the excess demand in the three goods markets. The two industrial goods excess demand functions add up domestic demand and exports of a firm and subtract its output. The agricultural goods demand is a constant fraction of world income, whereas supply equals the number of peasants.

Urban (1996) proves the uniqueness of the equilibrium of the system (11). Therefore, we can depict the equilibrium of the goods markets in figure 1.

Figure 1 about here

Figure 1 shows the three implicit functions $f(p_1; p_2; n)$; $f^*(p_1; p_2; n)$ and $g(p_1; p_2; n)$ in the p_2-p_1-space for a given n. The three schedules show the equilibrium in the domestic industrial goods market, the foreign industrial goods market and the market for agricultural products, respectively. The intersection of the three curves is the short run equilibrium. One of the equations is redundant due to Walras law. I drop the excess demand function for the foreign industrial goods market $f^*(p_1; p_2; n)$.

The equilibrium condition for domestic firms (f-schedule) is upward sloping, because any increase in domestic prices for given foreign prices reduces demand for domestic industrial products ($\partial f/\partial p_1 < 0$)[11]. In order to restore equilibrium at a given constant supply, foreign prices must also rise (because $\partial f/\partial p_2 > 0$).

The equilibrium condition for agricultural products (g-schedule) is downward sloping. A constant fraction of world income is spent on agricultural products. Supply is proportional to the number of peasants in the world and thus a constant in this model. If industrial wages in one region rise, world income is rising thus rising demand for agricultural products at constant supply. In order to restore equilibrium, the industrial wages in the other region must fall, until world income is back at the original level. As mentioned above, prices follow wages.

[11] The rigorous mathematical derivation of the derivatives in this paragraph is part of appendix C.

The way economic geography papers are written, leaves in our opinion a rest of a mystique as to what the model really is about. Before we analyze this system, we will depart from the previous model by setting up a completely different well-known microeconomic three goods and two heterogeneous agents pure exchange economy. This new set-up will lead to the same demand system as (11). It will be this analogy which sheds new light on the geography and trade models. In particular, it will contribute to the question, how to interprete trade costs and region size in this model.

4 What are Trade Costs?

As motivation for this section, we pose the question, whether a small open economy like Denmark in the vicinity of both Scandinavia and Germany is a big or a small country and whether such a country has a relatively high or a low parameter τ.

We set up a well-known pure exchange economy and show that this problem yields the goods market equilibrium conditions (11) under the following two conditions: 1) the consumer j is equally endowed with one good, owes the world endowment of the second and nothing of the third. 2) The consumer strictly prefers the good he is well endowed with to the good he is not endowed with.

There are two heterogeneous types of agents $j = 1; 2$ in a pure exchange economy with three goods $x_1; x_2; X_A$, where the total endowments of the economy are normalized to $\bar{x}_1 \equiv \frac{\alpha(\sigma-1)n_1}{\sigma}$, $\bar{x}_2 \equiv \frac{\alpha(\sigma-1)n_2}{\sigma}$, and $\bar{X}_A \equiv 1-\mu$. The endowments are distributed in the following way to the types of consumers: Consumer j is only endowed with good x_j, but not with good x_i, whereas X_A, the numeraire good, is equally distributed among the two types of consumers. This gives rise to the wealth constraint

$$y_j = p_j \bar{x}_j + \frac{1-\mu}{2}; \qquad (12)$$

where y_j is consumer wealth of all consumers of type j. Finally, the utility function for consumer j is given by the following expression:

$$U_j(c_{jj}; c_{ij}; C_{Aj}) = \left(\alpha_j c_{jj}^{\frac{\sigma-1}{\sigma}} + \alpha_i c_{ij}^{\frac{\sigma-1}{\sigma}}\right)^{\frac{\sigma}{\sigma-1}\mu} C_{Aj}^{1-\mu}; \; j \neq i; \tag{13}$$

where c_{ij} denotes consumption of consumer j for good x_i, C_{Aj} is consumption of consumer j for good X_A and

$$\alpha_j \equiv \left(\frac{n_j}{n_1 + n_2}\right)^{\frac{1}{\sigma}}$$
$$\alpha_i \equiv \left(\frac{n_i \tau^{\sigma-1}}{n_1 + n_2}\right)^{\frac{1}{\sigma}}$$

are some weighting factors in the utility function such that a consumer j prefers the good x_j to the good x_i. In other words, the consumer prefers the good which she ownes to the good that has to be bought from the other type of consumer. We can think of the following interpretation: the two consumers live in different regions and the preferences are biased towards the domestically available goods. The parameter τ is then a proxy for the degree to which preferences are biased towards domestic goods. The lower the τ; the stronger are domestic goods preferred. Additionally, the weighting factor includes also the size of the country as proxied by the number of goods in the Krugman (1991) model. To yield the same set of equilibrium prices (and wages) country size can be traded off with the degree of preference bias. For instance, China can afford to have more specific tastes than Denmark, and can still achieve higher relative prices (and wages).

The utility function (13) is maximized according to the budget constraint:

$$c_{1j} p_1 + c_{2j} p_2 + C_{Aj} = y_j \tag{14}$$

It is straight forward to show that exactly the demand system (11) emerges.[12] It is this analogy that sheds new light on the Dixit-Stiglitz-Krugman model. In particular,

[12] The analogy holds only for the short run equilibrium of the model in Krugman (1991). The long run equilibrium could be replicated by our pure-exchange economy, if preferences are endogenous. For example, some type 1 consumers turn into type 2 consumers in the next generation (next time period), if type 2 experiences higher utility in the short run equilibrium today. These consumers will not only change their preferences, but also the endowment. This simple example shows that the application of the Krugman (1991) model is not restricted to economic geography, but also applies to heterogeneous social groups within one economy.

this is somewhat in contrast to Fujita and Thisse (1996) who attribute a major role in the explanation of agglomeration economics to increasing returns, whereas we achieve similar effects in a neoclassical pure exchange economy. The production side of the Dixit-Stiglitz-Krugman model seems not so important for its effect on agglomeration economics than the demand side.[13]

Now, we return to the question, whether a small open economy like Denmark is a big or a small region. On the one hand, if trade costs are taken literally to be transport costs as Krugman (1991) seems to suggest, then Denmark is a big region. This is so, because Denmark has a relatively high population density. I.e. a firm finds a relatively large market in its surroundings.[14] Furthermore, Denmark should not be considered a homogeneous region independent of South Sweden and North Germany, because transport costs e.g. from Kolding to Flensburg are not substantially different from transport costs from Kolding to Skagen.

On the other hand, if trade costs are interpreted as a parameter of preference bias towards domestic goods, then the question arises, how different are Denish tastes, relative to European tastes. Clearly, Denmark should then be regarded as a homogeneous region with a unique cultural background and relatively homogeneous tastes inside, but (maybe) somewhat different tastes relative to people in other countries. For example, Denish books are strictly preferred by Denish people relative to, say, Swedish people (who can read Denish books with some inconvenience). Furthermore, Denmark would be a small region, because absolute number of inhabitants is the relevant meassure of region-size in this case. After having clarified some interpretations of the model, we will return to the original set-up of sections (2.1) and (2.2), analyze the system (11), and clarify the agglomeration and convergence forces in the model.

[13] What matters in the Krugman-Dixit-Stiglitz model is the endogeneity of product space (see Matsuyama, 1995).

[14] This statement requires just one qualification. Transport costs are higher for the islands. Still, it is not obvious that products are more expensive, for example, in Copenhagen than in Kolding.

5 Why is there Agglomeration or Convergence?

Suppose for the moment that a worker changes for some arbitrary reason her residence. What is going to happen to the wages and prices? If wages rise in the immigration region and fall in the emmigration region, there will be an incentive for more workers to follow (agglomeration). If wages fall in the immigration region and rise in the emmigration region, there will be an incentive for this worker to go back (convergence)[15]. Because of equation (5), wages are proportional to prices. Hence, we have to examine, how prices change, if a worker moves thereby changing the regional distribution of firms and goods. We will begin with the impact of a movement of a worker on the equilibrium condition for domestic industrial goods.

If an industrial worker moves from the foreign to the domestic region at a given level of prices and wages, then the demand-change for domestically produced industrial goods will be described by the following expression:

$$\begin{aligned} \frac{\partial f}{\partial L} &= w_1 D(p_1; P_1) - w_2 \cdot \frac{E(p_1; P_2)}{\tau} \\ &\quad + y_1 \frac{\partial D(p_1; P_1)}{\partial P_1} \cdot \frac{\partial P_1}{\partial L} + \frac{y_2}{\tau} \cdot \frac{\partial E(p_1; P_2)}{\partial P_2} \cdot \frac{\partial P_2}{\partial L}; \end{aligned} \tag{15}$$

where we made use of (9), (10), and (11). A movement of labour from the bigger to the smaller region has two effects on domestic demand for industrial goods.

The income effect: There is one worker more in the domestic region, who spends her income on domestic goods and hence is there one person less in the foreign region, who could spend some income on exports. Note that the price for the domestic good is reduced for this person due to the absence of trade costs in intra-regional trade.

[15] In the context of economic geography, convergence means the tendency of increasing returns to scale industries to allocate equally in plane, whereas convergence in the growth literature means the tendency of growth rates of GDP of poorer countries to be bigger than the one in richer countries (This is absolute convergence as defined by Barro and Sala-i-Martin, 1995) The two notions are interrelated, if increasing returns industries have higher technological progress than agriculture. An agglomeration process of industries would then also imply divergence of growth rates and a convergence process of industries would also mean convergence of growth rates.

This means that this person now spends relatively more on the domestic good than before. This e®ect is captured in the ¯rst and second term of equation (15).

The number of goods e®ect: If a worker moves from the foreign to the domestic region, ¯rms will relocate, too. Because there are now more ¯rms in the domestic region and less in the foreign, the composition of the price index also changes. At given prices, domestic consumers have fewer import goods to pay transport cost for (and vice versa for the foreign consumer). This lowers the domestic price index for industrial goods and rises the foreign one. If the domestic price index is lowered, the relative price of domestic goods to the price index is increased. This decreases ceteris paribus domestic demand for products of domestic ¯rms. The reverse holds for exports. This e®ect is captured in the third and fourth term of equation (15). Because these two e®ects are the key to the understanding of agglomeration and convergence in this model, we will repeat them looking from a di®erent ankle that coincides with the graphical exposition that follows in this section.

Suppose that a worker moves from the smaller domestic to the bigger foreign region (rather than vice versa as in the explanation above). On the one hand, the migrated worker increases total income in the foreign region and reduces total income in the domestic region. This typically reduces demand for the domestic ¯rm, because the migrated worker buys less goods from the former home region. This is an agglomeration force, because it rises domestic prices and wages at given foreign prices and wages.

On the other hand, the decrease in the number of workers in the domestic region reduces the number of ¯rms and goods produced. The domestic CES price index accounts for this e®ect by an increase, because the composition of the index changes towards foreign goods, which are more expensive because of the transport cost and because of the higher labor cost in the bigger foreign region. The increase in the domestic price index reduces the relative price for domestic goods. This increases domestic demand for industrial goods and is thus a convergence force in the model, because a decrease in domestic demand lowers domestic prices and wages at given foreign prices and wages. It is ambiguous, i.e. dependent on the parameter of the model, which force dominates ($\partial f / \partial L \gtrless 0$). This can be demonstrated in figure 2.

Figure 2 about here

Figure 2 shows in panel (a) the agglomeration case. Suppose that starting from the equal distribution equilibrium, workers move out continuously from the domestic region making it the smaller region. If the emmigration causes a reduction in demand for domestic goods, this causes prices and wages in the domestic region to fall at given prices and wages in the foreign region. Hence, the f-schedule shifts leftward.

Additionally, the g-schedule twists anti-clockwise around the ($w_1 = 1$, $w_2 = 1$)-point. If wages in the domestic region are smaller, than the emmigration will cause a rise in world income which leads to excess demand in the market for agricultural goods. In order to restore equilibrium, domestic wages have to fall at given foreign wages. Both movements of the f- and the g-schedule lead to a fall of domestic wages relative to foreign wages. Hence, the emmigration is self-enforcing. The economy ends up at complete agglomeration ($L = 0$).

Figure 2 shows in panel (b) the convergence case. Now, emmigration causes a rise in domestic industrial goods demand which increases domestic prices and wages at given foreign prices. Hence, the f-schedule shifts rightward. If wages in the domestic region are higher than in the foreign region, then the emmigration causes a fall in world income which leads to excess supply in the market for agricultural products. In order to restore equilibrium, domestic wages have to rise at given foreign wages. The movements of both schedules together induce a rise in domestic wages relative to foreign wages. Hence, the incentive to move out is reversed and the equal distribution equilibrium is stable in the long run.

The last conclusion can graphically be demonstrated more clearly in figure 3.

Figure 3 about here

Figure 3, panel (a) repeats the arrowed line from figure 2 (b). Panel 3 (b) is constructed from panel 3 (a) by drawing a ray through the origin to the equilibrium wage combination for a specific worker distribution. Taking the tangens of the angle of this array gives the relative nominal wage $\left(\frac{w_2(L)}{w_1(L)}\right)$ at this labour distribution. Repeating this procedure for every possible labour distribution gives a schedule depicting the relative nominal wage as a function of the labour distribution. This curve is drawn in panel (b) and describes the short run equilibrium condition.[16] The points A,B,C and D correspond in the two panels of figure 3. The schedule is either upward or downward sloping.

Still, this is not the end of the story, because so far only relative nominal wages are considered. A migration decision is rather based on relative real wages. Let's look at the long run steady state condition: equal real wages.

$$\omega_1 \equiv \frac{w_1}{P_1(p_1; p_2; n)^{\mu}} = \frac{w_2}{P_2(p_1; p_2; n)^{\mu}} \equiv \omega_2 \qquad (16)$$

Let's suppose that all nominal wages and prices are equal in both regions. Does this guarantee equal real wages, too? The answer is no. If the domestic region is smaller, i.e. has less workers and less industrial products, more industrial products have to be imported. This rises the domestic CES price index above the foreign one, because the transport cost mark-up has to be paid for more products.

Figure 4 about here

But then the domestic real wage is smaller than the foreign. Hence, the domestic nominal wage needs to be bigger than the foreign nominal wage for the real wages to be equal, if the domestic region has less workers. The equal real wage condition is downward sloping in the $\left(\frac{w_2}{w_1}; L\right)$-space and is depicted in figure 4 as the solid line.[17]

[16] The mathematical derivation of this line is found in appendix D and denoted $h(W; L)$.

[17] The mathematical form of the equal real wage condition is derived in appendix E and is denoted $k(W; L)$.

Real wages are bigger in the foreign region above this line (agglomeration) and bigger in the domestic region below this line (convergence).

Urban (1996) proves that three cases are possible. The short run equilibrium condition is always above the equal real wage condition; then real wages are bigger in the bigger region and workers of the smaller region have an incentive to move to the bigger region thus self-enforcing the agglomeration process (see ¯gure 4, panel a). The short run equilibrium condition is always below the equal real wage condition; then real wages are bigger in the smaller region and the equal distribution equilibrium is dynamically stable (see ¯gure 4, panel c).[18] The equal real wage condition cuts the short run equilibrium condition from above (see ¯gure 4, panel b). There will be an unstable intermediate steady state equilibrium S_1 next to the symmetric steady state equilibrium S_2. For any labour distribution smaller than the one corresponding to S_1, there is an agglomeration process going on. For any labour distribution bigger than the one corresponding to S_1, the system converges to the equal distribution equilibrium.[19] This case can be considered a \poverty trap", because it depends on the initial distribution of industries, whether a region becomes industrialized or not.

Urban (1996) gives the precise conditions for each of these cases in his proposition 2. If an economy starts out with low scale economies (low ¾), a big agricultural sector (high ¹), and high transport costs (low ¿), then the economy is likely to be described by the convergence scenario (¯gure 4 (a)). As transport costs are falling, industries are developing, and economies of scale are rising, the economy will most likely end up in the agglomeration scenario (¯gure 4 (c)). Whether the \poverty trap" scenario is passed on the way of development (¯gure 4 (b)), depends on whether the ¯rst industries started to be spread even in plane or were already clustered in a few places. Then, the tendency of clustering might have appeared in some regions, whereas it might not have appeared in others. This intermediate stage (poverty trap) might explain di®erent stages in the degree of agglomeration.

[18] These two cases are reported in Krugman (1991).

[19] This case is found in Urban (1996).

6 Summary

This paper analyses the reasons for agglomeration or convergence in a typical geography and trade model such as Krugman (1991). Agglomeration happens, if the movement of one worker from the domestic to the foreign region increases aggregate demand for industrial products in this region at given prices and constant output. This happens, if an income e®ect overcompensates a number of goods e®ect. The income e®ect arises, because this worker spends after his relocation more on goods of foreign ¯rms and less on domestic ¯rms in order to save trade costs.

The number of goods e®ect arises, because an increase in demand at given prices rises pro¯ts. This induces new ¯rms to enter the market and thus to increase the number of goods in the foreign region, whereas the number of goods is falling in the domestic region. This rises ceteris paribus the price index in the smaller region, because trade costs have to be paid for more goods than before, and lowers the price index in the bigger region. If the price index is risen in the domestic region, this lowers the relative price of domestic goods which increases domestic demand.

For reassuring equilibrium, the (producer) price has to rise in the foreign region and to fall in the domestic region. Because prices are set as constant mark-ups over marginal cost, wages follow the price movement. Hence, more workers have an incentive to follow the ¯rst one thereby even enforcing the divergence of wages.

Convergence takes place, if the movement of one worker increases demand for products in the smaller region. The movement of one worker reduces the pro¯ts of ¯rms in the smaller region. This drives some ¯rms out of the market and reduces the number of goods. The reduction of the number of goods increases the CES price index in the smaller region, because more foreign goods have to be bought, which bear transport cost. This decreases the relative price of an industrial product in the smaller region and increases demand. This e®ect needs to overcompensate the income e®ect, which decreases the total income of a region, if a worker moves away.

There exists an interior equilibrium for some parameter range. If a region has a certain critical mass of industries, industries tend to spread equally in space. If a region has not this critical mass of industries (\poverty trap"), this region is going to dry out of industries completely.

Finally, it is shown that the Dixit-Stiglitz-Krugman model behaves like a pure exchange economy with two heterogeneous types of consumers, three goods and the following properties: 1) Each type of consumer is equally endowed with one good, owes the world endowment of the second, and nothing of the third. 2) The consumer strictly prefers the good he is well endowed with to the good he is not endowed with.

References

d'Aspremont, Claude, Rudolphe Dos Santos Ferreira, and Louis-André Gérard-Varet, (1996), On the Dixit-Stiglitz Model of Monopolistic Competition, *American Economic Review*, Vol. 86, p. 623-629.

Barro, Robert J., and Xavier Sala-i-Martin, (1995), Economic Growth, McGraw Hill, New York.

Davis, Donald R., (1998), The home market, trade, and industrial structure, *American Economic Review*, forthcoming.

Dixit, Avinash K., and Joseph E. Stiglitz, (1977), Monopolisitic Competition and Optimum Product Diversity, *American Economic Review*, Vol. 67, 297-308.

Dixit, Avinash K., and Joseph E. Stiglitz, (1993), Monopolistic Competition and Optimal Product Diversity: Reply, *American Economic Review*, Vol. 83, No. 1, p. 302-304.

Fujita, Masahisa, and Jacque-Francois Thisse, (1996), Economics of Agglomeration, CEPR-Discussion Paper No. 1344.

Krugman, Paul R., (1980), Scale Economies, Product Differentiation, and the Pattern of Trade, *American Economic Review*, Vol. 70, p. 950-959.

Krugman, Paul R., (1991), Increasing Returns and Economic Geography, *Journal of Political Economy*, Vol. 99, No. 3, p. 483-499.

Krugman, Paul R., and Anthony J. Venables, (1995), Globalization and the Inequality of Nations, *Quarterly Journal of Economics*, Vol. 110, No. 4, p. 857-880.

Matsuyama, Kiminori, (1995), Complementarities and Cumulative Processes in Models of Monopolistic Competition, *Journal of Economic Literature*, Vol. 33, p. 317-334.

Norman, Victor D., and Anthony J. Venables, (1995), International Trade, Factor Mobility, and Trade Costs, *The Economic Journal*, Vol. 105, p. 1488-1504.

Urban, Dieter M., (1996), Increasing Returns and Economic Geography: An Analytical Note, mimeo.

Venables, Anthony J., (1995), Economic Integration and the Allocation of Firms, *American Economic Review*, Vol. 85, No. 2, p. 296-300.

Yang, Xiaokai, and Ben J. Heijdra, (1993), Monopolistic Competition and Optimal Product Diversity: Comment, *American Economic Review*, Vol. 83, No. 1, p. 295-301.

Appendix

Appendix 1: Definition of Price Indices

The explicit functional forms for the price indices are:

$$
\begin{aligned}
P_1(w_1; w_2; L) &= \gamma\beta\left(\frac{L}{\mu} w_1^{1-\sigma} + \tau^{\sigma-1}\left(\frac{1-L}{\mu}\right) w_2^{1-\sigma}\right)^{\frac{1}{1-\sigma}} \\
P_1(p_1; p_2; n) &= \left(n p_1^{1-\sigma} + \left(\frac{1}{\alpha\sigma} - n\right)(p_2/\tau)^{1-\sigma}\right)^{\frac{1}{1-\sigma}} \\
P_2(w_1; w_2; L) &= \gamma\beta\left(\left(\frac{1-L}{\mu}\right) w_2^{1-\sigma} + \tau^{\sigma-1}\frac{L}{\mu} w_1^{1-\sigma}\right)^{\frac{1}{1-\sigma}} \\
P_2(p_1; p_2; n) &= \left(\left(\frac{1}{\alpha\sigma} - n\right) p_2^{1-\sigma} + n(p_1/\tau)^{1-\sigma}\right)^{\frac{1}{1-\sigma}}:
\end{aligned}
\tag{17}
$$

Appendix 2: The Excess Demand Functions

The equation system (11) is written in explicit functional form:

$$
\begin{aligned}
\frac{p_1^{-\sigma}\left(p_1 n\left(\bar{x} + \frac{1-\mu}{2}\right)\right)}{n p_1^{1-\sigma} + \tau^{\sigma-1}\left(\frac{1}{\alpha\sigma} - n\right) p_2^{1-\sigma}} + \frac{p_1^{-\sigma}\left(p_2\left(\frac{1}{\alpha\sigma} - n\right)\left(\bar{x} + \frac{1-\mu}{2}\right)\right)}{n p_1^{1-\sigma} + \tau^{1-\sigma}\left(\frac{1}{\alpha\sigma} - n\right) p_2^{1-\sigma}} &= \frac{\bar{x}}{\mu} \\
\frac{\tau^{\sigma-1} p_2^{-\sigma}\left(p_1 n\left(\bar{x} + \frac{1-\mu}{2}\right)\right)}{\tau^{\sigma-1} n p_1^{1-\sigma} + \left(\frac{1}{\alpha\sigma} - n\right) p_2^{1-\sigma}} + \frac{p_2^{-\sigma}\left(p_2\left(\frac{1}{\alpha\sigma} - n\right)\left(\bar{x} + \frac{1-\mu}{2}\right)\right)}{n p_1^{1-\sigma} + \tau^{\sigma-1}\left(\frac{1}{\alpha\sigma} - n\right) p_2^{1-\sigma}} &= \frac{\bar{x}}{\mu} \\
p_1 \cdot n + p_2 \cdot \left(\frac{1}{\alpha\sigma} - n\right) &= \frac{1}{\bar{x}}:
\end{aligned}
\tag{18}
$$

Appendix 3: Partial Derivatives of Excess Demand Functions

This part of the appendix calculates the signs of the partial derivatives of the excess demand functions of equation system (11) which is written in explicit functional form in appendix B.

1) The determination of the sign of $\frac{\partial f}{\partial p_1}$:

We will first rewrite the first equation in (18) in the following way:

$$
f = \frac{f_1}{f_2} + \frac{f_3}{f_4} - \bar{x}; \tag{19}
$$

where we used the notation n_1 and n_2 and defined

$$
\begin{aligned}
f_1 &\equiv \mu\left(\bar{x} + \frac{1-\mu}{2 p_1 n_1}\right) > 0; \\
f_2 &\equiv 1 + \tau^{\sigma-1}\left(\frac{n_2}{n_1}\right)\left(\frac{p_2}{p_1}\right)^{1-\sigma} > 0; \\
f_3 &\equiv \tau^{\sigma-1}\mu y_2 > 0; \\
f_4 &\equiv \tau^{\sigma-1} n_1 p_1 + n_2 p_2^{1-\sigma} p_1^{\sigma} > 0:
\end{aligned}
$$

The partial derivatives are thus:

$$
\begin{aligned}
\frac{\partial f_1}{\partial p_1} &= -\frac{1-\mu}{2p_1^2 n_1 \bar{x}} < 0; \\
\frac{\partial f_2}{\partial p_1} &= \tau^{\sigma-1}(\sigma-1)\left(\frac{n_2}{n_1}\right) p_2^{1-\sigma} p_1^{\sigma-2} > 0; \\
\frac{\partial f_3}{\partial p_1} &= 0; \\
\frac{\partial f_4}{\partial p_1} &= \tau^{\sigma-1} n_1 + \sigma n_2 p_2^{1-\sigma} p_1^{\sigma-1} > 0.
\end{aligned}
\tag{20}
$$

The partial derivative of f is then given by the following expression:

$$
\frac{\partial f}{\partial p_1} = \frac{\frac{\partial f_1}{\partial p_1} f_2 - f_1 \frac{\partial f_2}{\partial p_1}}{f_2^2} - \frac{f_3}{f_4^2 \frac{\partial f_4}{\partial p_1}} < 0;
\tag{21}
$$

which is readily checked to be negative. Q.E.D.
2) The determination of the sign of $\frac{\partial f}{\partial p_2}$:

First, we slightly rewrite (19) by defining $\tilde{f}_3$ and $\tilde{f}_4$ to replace f_3 and f_4, respectively:

$$
\begin{aligned}
\tilde{f}_3 &\equiv \mu \tau^{\sigma-1} p_1^{-\sigma} \left(p_2 \cdot n_2 \cdot \bar{x} + \frac{1-\mu}{2} \right) > 0; \\
\tilde{f}_4 &\equiv \tau^{\sigma-1} n_1 p_1^{1-\sigma} + n_2 p_2^{1-\sigma} > 0.
\end{aligned}
\tag{22}
$$

Then, we can formulate the following partial derivatives:

$$
\begin{aligned}
\frac{\partial f_1}{\partial p_2} &= 0; \\
\frac{\partial f_2}{\partial p_2} &= \tau^{\sigma-1}(1-\sigma)\left(\frac{n_2}{n_1}\right) p_1^{1-\sigma} p_2^{-\sigma} < 0; \\
\frac{\partial \tilde{f}_3}{\partial p_2} &= \mu \tau^{\sigma-1} p_1^{-\sigma} n_2 \bar{x} > 0; \\
\frac{\partial \tilde{f}_4}{\partial p_2} &= (1-\sigma) n_2 p_2^{-\sigma} < 0.
\end{aligned}
\tag{23}
$$

The partial derivative of f is then given by the following expression:

$$
\frac{\partial f}{\partial p_2} = -\frac{f_1}{f_2^2 \frac{\partial f_2}{\partial p_2}} + \frac{\frac{\partial \tilde{f}_3}{\partial p_2} \tilde{f}_4 - \tilde{f}_3 \frac{\partial \tilde{f}_4}{\partial p_2}}{\tilde{f}_4^2} > 0;
\tag{24}
$$

which is readily checked to be positive. Q.E.D.

The sign of the partial derivative $\partial f / \partial L$ is ambiguous. The signs of the partial derivatives of $g(p_1; p_2; L)$ can be readily seen from the 3rd equation of (18).

Appendix 4: Goods Market Equilibrium Condition

This part of the appendix derives an implicit functional form of the goods market equilibrium condition in W-L-space which is repeatedly shown in figures 3 and 4.

Reformulating the first two equations from appendix B by using (5) and (6) and (7) yields the following equation system which guarantees goods market equilibrium for domestic industrial products and agricultural products:

$$\mu = w_1 L + (\mu - L) w_2; \tag{25}$$

$$1 = \frac{w_1^{-\sigma}\left(w_1 L + \frac{1-\mu}{2}\right)}{\frac{L}{\mu} w_1^{1-\sigma} + \tau^{\sigma-1}\left(\frac{\mu - L}{\mu}\right) w_2^{1-\sigma}} + \frac{\tau^{\sigma-1} w_1^{-\sigma}\left(w_2 (\mu - L) + \frac{1-\mu}{2}\right)}{\tau^{\sigma-1}\frac{L}{\mu} w_1^{1-\sigma} + \left(\frac{\mu - L}{\mu}\right) w_2^{1-\sigma}}.$$

Expressing this system of equations in relative wages $W = \frac{w_2}{w_1}$ and rearranging yields:

$$\frac{\left(L + \frac{1-\mu}{2w_1}\right)}{\left(\frac{L}{\mu} + \tau^{\sigma-1}\left(\frac{\mu - L}{\mu}\right) W^{1-\sigma}\right)} + \frac{\tau^{\sigma-1}\left(W \cdot (\mu - L) + \frac{1-\mu}{2w_1}\right)}{\left(\tau^{\sigma-1}\frac{L}{\mu} + \left(\frac{\mu - L}{\mu}\right) W^{1-\sigma}\right)} = 1; \tag{26}$$

$$\frac{L + W(\mu - L)}{\mu} = \frac{1}{w_1}.$$

Plugging the second into the first equation of (26) allows to define an implicit function $h(W;L)$ in the relative nominal wage W and the labor distribution L, which fully characterises goods market equilibrium in both sectors:

$$h(W;L) \equiv \frac{2\mu L + (1-\mu)(L + W(\mu - L))}{(L + \tau^{\sigma-1}(\mu - L) W^{1-\sigma})} \tag{27}$$

$$+ \frac{\tau^{\sigma-1}(2\mu W(\mu - L) + (1-\mu)(L + W(\mu - L)))}{(\tau^{\sigma-1} L + (\mu - L) W^{1-\sigma})} - 2 = 0.$$

The implicit functional form $h(W;L)$ can be solved for L giving two solutions. Only one of them can be in positive prices because of the uniqueness of the short run equilibrium (Proof see Urban (1996), proposition 1). Hence, the short run equilibrium condition must be either monotonically increasing or monotonically decreasing in $W-L$-space. By inspection of the algebraic form of the solution to L that is a rational function, the schedule of $h(W;L)$ must also be continuous.

Appendix 5: Equal Real Wage Condition

This part of the appendix gives a functional form for the equal real wage condition in figures 3, 4 and 5. The definition of the relative real wage can be rewritten in terms of the nominal real wage and the labor distribution by using (17) and (16):

$$\omega(L;W) = \frac{(\tau^{\sigma-1} L + (\mu - L) W^{1-\sigma})^{\frac{1}{1-\sigma}}}{W(L + \tau^{\sigma-1}(\mu - L) W^{1-\sigma})^{\frac{1}{1-\sigma}}}. \tag{28}$$

In the steady state, the relative real wage needs to be equal to one ($\omega(L;W) = 1$). This equation is solved for L and an implicit function $k(W;L)$ in W and L is defined for which the real wage is one:

$$k(W;L) \equiv -\frac{\mu\left(W^{1-\sigma} - \tau^{\sigma-1} W^{(1-\sigma)\left(1+\frac{1}{\mu}\right)}\right)}{\tau^{\sigma-1} - W^{1-\sigma} - W^{\frac{1-\sigma}{\mu}} + \tau^{\sigma-1} W^{(1-\sigma)\left(1+\frac{1}{\mu}\right)}} - L = 0. \tag{29}$$

This is the implicit functional form for the equal real wage condition used in figure 4.

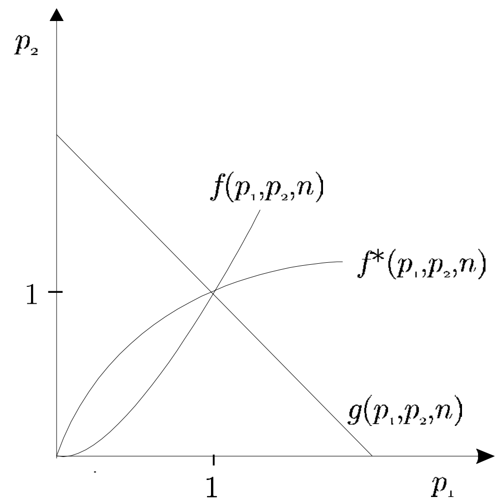

Figure 1: Short Run Equilibrium

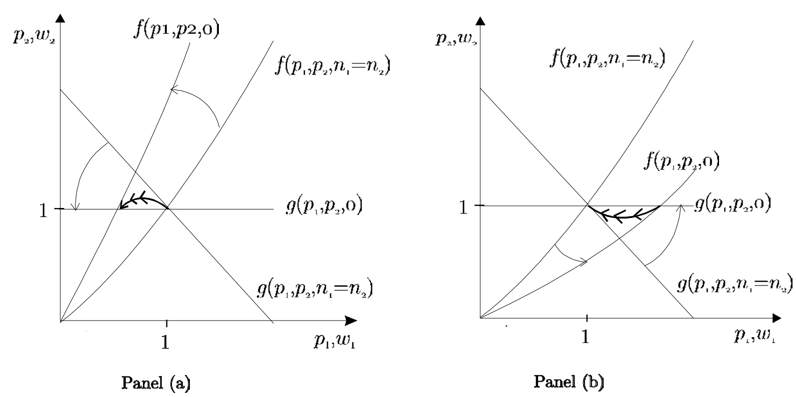

Figure 2: Comparative Statics of Worker Migration

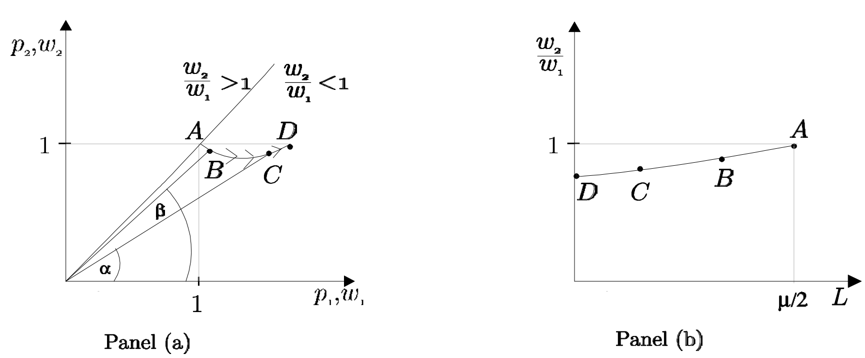

Figure 3: Short Run Equilibrium Condition

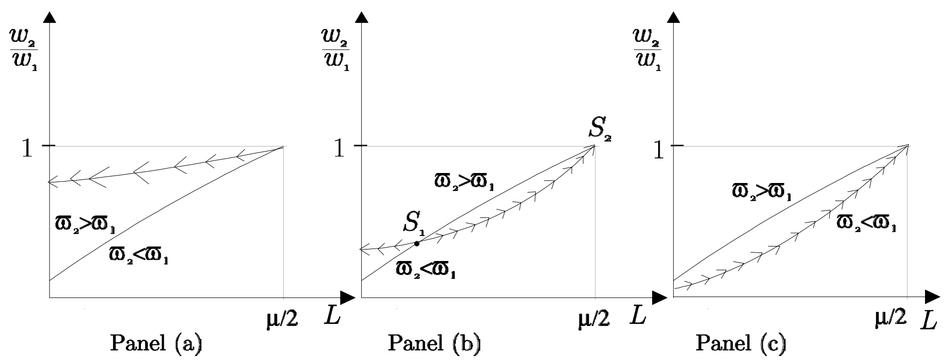

Figure 4: Agglomeration or Convergence

Understanding

'Increasing Returns to Scale and Economic Geography'

Part II: An analytical solution

by

Dieter M. Urban
Copenhagen Business School

Abstract part II: This paper provides an analytical solution to the Krugman (1991a) model explaining industry agglomeration. It is shown there exists a unique short-run equilibrium and multiple long-run equilibria. The latter proves the existence of a "poverty trap" in this model: depending on the initial level of industries we will either see industries spreading evenly in the plane, or moving away from one of the regions. However, it is also shown that this "poverty trap" will not appear if the economy starts developing from an equal distribution of industries.

November 19, 1998

Increasing Returns and Economic Geography: An Analytical Note¤

By

Dieter M. URBAN
Department of Economics
Copenhagen Business School, Denmark.

Abstract

This paper provides an analytical solution to the Krugman (1991a) model explaining industry agglomeration. It is shown there exists a unique short-run equilibrium and multiple long-run equilibria. The latter proves the existence of a \poverty trap" in this model: depending on the initial level of industries we will either see industries spreading evenly in the plane,or moving away from one of the regions. However, it is also shown that this \poverty trap" will not appear if the economy starts developing from an equal distribution of industries.

JEL Classi¯cation: F12, R60.
Keywords: convergence, agglomeration, poverty trap.

Mailing Address: Dieter M. Urban, Department of Economics, Copenhagen Business School, Nansensgade 19, 5th °oor, DK-1366 Copenhagen K, Denmark.

¤The author thanks Niels Blomgren-Hansen, Karsten Junius, Je®rey Nilsen and Pascalis Raimondos-M¿ller for their comments. Usual caveats apply.

1 Introduction

The economics of agglomeration have become a growing ¯eld of theoretical research.[1] Problems of industry agglomeration are also attracting more and more attention by applied economists. An explanation of agglomeration processes is provided by geography and trade models of the Krugman (1991a) type. These models are appealing, because (i) they generate endogenously a cumulative process, similar to the one informally described in Myrdal (1957), (ii) they do not rely on (unobservable) exogenous externalities, and (iii) they are based on internal increasing returns to scale that is thought to be an important source for aggloemration among regional economists (see Fujita and Thisse, 1996).

Yet, these models lack the analytical rigor that can be found in corresponding models with constant returns to scale production.[2] The reason is that these models are based on a non-linear equation system which does not necessarily obey the standard convexity assumptions and may therefore yield multiple equilibria. A consequence is that the literature is often either based on numerical simulations (e.g. Venables, 1996), or on the analysis of corner solutions (e.g. Krugman, 1991a). If geography and trade models are analytically rigorous, then they usually rely on factor price equalization ruling out some agglomeration and convergence forces (e.g. Helpman and Krugman, 1985, and Martin and Rogers, 1995).

The purpose of this paper is to give a complete analytical treatment of the seminal geography and trade model by Krugman (1991a). In doing so, we uncover some interesting properties of the model that explain the existence of a \poverty trap". In Krugman's model, the relocation of industries between two regions is driven by worker migration due to real wage di®erences. Krugman (1991a) concentrates on the case, in which it pays for a ¯rm to attract workers to a region that had no industries before (total agglomeration). Krugman and Venables (1995) and Venables (1995) analyze the case, in which it pays for a ¯rm to defect from an equal distribution of industries

[1] A recent survey on the economics of agglomeration is Fujita and Thisse (1996).
[2] See, for example, Norman and Venables (1995).

(symmetry-breaking). The contribution of this paper is to ¯nd analytically all steady state industry distributions in between these two extreme cases. Furthermore, we are able to describe the industry reallocation dynamics at any initial distribution of industries. In doing so, we identify a particular parameter range for which a \poverty trap" arises: if a region has a certain initial threshold level of industries, they tend to spread evenly in the plane (convergence); On the contrary, if a region does not have this threshold level, all its industries move away (total agglomeration). In this sense, the model shows that initial conditions matter in Krugman's (1991a) model not only for which region is drying out of industries (determination of the \winner" region), but also whether industries agglomerate or converge (threshold property of the agglomeration process). The integration of \similar" regions (e.g. European Community) causes convergence of industry distribution, the integration of a region with a lot of industries and one with just a few (e.g. German uni¯cation) causes agglomeration.

Finally, we show that the \poverty trap" case will not appear if the economy starts developing from an equal distribution of industries. In this particular case, the analysis of symmetry-breaking, as described in Venables (1995) and Krugman and Venables (1995), is su±cient to fully characterize the model.

The mathematical problem solved in this paper is the determination of the exact number of equilibria in a simple ¯xed point problem with multiple solutions. The well-known ¯xed point theorems provide little help for this problem, because they prove the uniqueness (non-uniqueness) of an equilibrium. We ¯nd a speci¯c solution to this problem for models based on polynomials (i.e. models with CES or Cobb-Douglas functional forms). This solution applies not only to economic geography models, but also to models that exhibit poverty traps and growth (Urban, 1998).

The rest of this paper is organized as follows. Section 2 describes brie°y the Krugman (1991a) model. Subsection 2.1 proves the uniqueness of the short-run equilibrium and subsection 2.2 analyzes the long-run equilibrium and proves the existence of a \poverty trap". Some conclusions can be found in section 3.

2 The Krugman (1991a) Model

The model has two sectors, two regions and two consumers. The two sectors – agriculture and industry – differ by their market form: the market for agricultural goods is perfectly competitive; the market for industrial goods is monopolistically competitive. Regions are defined as areas for which it is costless to trade industrial goods within them, but costly to trade industrial goods across them. Consumers $j = 1; 2$ differ only by their place of residence. It is assumed that there is no short-run mobility of production factors from one region to the other. However, mobility of production factors is allowed in the long-run.

Consumers have identical Cobb-Douglas utility functions containing the agricultural good and the aggregate basket of the industrial goods. The income share attributed to the industrial goods basket is denoted by μ. The industrial goods basket is further specified by a Dixit-Stiglitz (1977) subutility function with σ denoting the elasticity of substitution between varieties ($\sigma > 1$). The agricultural good is taken as the numeraire. Based on these standard utility functions, the price index is well known and it is given by the following expression:

$$P_j = \left(\frac{n_i}{n_1 + n_2} (p_i^{ex})^{1-\sigma} + \frac{n_j}{n_1 + n_2} p_j^{1-\sigma} \right)^{\frac{1}{1-\sigma}}; \quad i = j; \tag{1}$$

where the number of goods in region i is denoted by n_i; the domestic prices of firms in region j are denoted by p_j; and the fob-prices for export goods are denoted by p_i^{ex}.

Both goods are produced by using only labour. Furthermore, it is assumed that there is factor specificity for industrial production by workers, and for agricultural production by peasants. Peasants work according to a constant returns to scale technology. Workers work according to an increasing returns to scale technology, where marginal cost is constant and where there exists some fixed cost. Firms are assumed to have zero profits.

The sum of workers in both regions ($L_1 + L_2$) is normalised to μ. Thus, the total amount of peasants is $1 - \mu$ and they are assumed to be equally distributed in both regions and immobile. Every peasant and worker supplies one unit of work and earns a salary 1, if peasant, and w_i, if worker in region i.

Finally, there are transport cost of the Samuelson iceberg-type, such that only a fraction τ of one produced unit of an industrial good arrives at its foreign destination ($0 < \tau < 1$). There are no transport cost for goods delivered to domestic customers or for the agricultural goods.

2.1 The Short-Run Equilibrium

Having described briefly the model, we now present the economic relationships that result from consumer optimization, firm optimization, labour market clearing, product market clearing, and the zero profit condition.

Under the standard assumption that a firm does not take into account the impact of its pricing decision on other firms' pricing decisions and on regional income, the pricing rule for the firm will be: [3]

$$p_i = \frac{\sigma}{\sigma - 1} \cdot \beta \cdot w_i \quad \text{and} \quad p_i^{ex} = \frac{\sigma}{\sigma - 1} \cdot \beta \cdot w_i / \tau; \tag{2}$$

where $\sigma/(\sigma - 1)$ is the mark-up of prices over marginal cost, and β is the reciprocal of the marginal product of labour. Note, that the pricing decision for export goods p_i^{ex} requires the firm to take into account the additional transport cost.

The optimal output of the firm is determined by the zero profit condition:

$$x_i = \frac{\alpha \cdot (\sigma - 1)}{\beta} \equiv \bar{x}; \tag{3}$$

where α is the fixed cost parameter. It is seen that the equilibrium output of the firm is independent of the number of workers or the number of firms in a region. The equilibrium number of firms per region follows from the labour market clearing

[3] This equation is an approximation which is only fairly good for a large number of firms (see the discussion in Yang and Heijdra (1993), Dixit and Stiglitz (1993), and d'Aspremont et. al. (1996)). The latter authors show that the mark-up of prices over marginal cost is underestimated by the above approximation. Any correction for this would not change the principle story as long as there is a positive relationship between prices and wages and a positive relationship between number of firms and number of workers.

conditions and the output decision:

$$n_i = \frac{L_i}{\alpha\sigma}. \tag{4}$$

If the number of workers increases in a region, workers drag industries with them and the number of goods increases proportionally. Since pro¯ts are zero, aggregate income y_i in a region i will be the sum of income of all workers and peasants in that region: [4]

$$y_i = w_i \cdot L_i + \frac{1-\mu}{2} = n_i p_i \bar{x} + \frac{1-\mu}{2}. \tag{5}$$

Finally, the conditions for equilibrium in the goods market are as follows: [5]

$$\frac{\mu p_1^{1-\sigma} y_1}{n_1 p_1^{1-\sigma} + t n_2 p_2^{1-\sigma}} + \frac{t \mu p_1^{1-\sigma} y_2}{t n_1 p_1^{1-\sigma} + n_2 p_2^{1-\sigma}} = \bar{x}; \tag{6}$$

$$(y_1 + y_2)(1-\mu) = 1-\mu; \tag{7}$$

where, for notational simplicity, we write $t = \tau^{1-\sigma}$. Equation (6) gives the market clearing condition for the domestic industrial goods, adding up domestic demand and exports of a ¯rm and setting it equal to output. Equation (7) gives the market clearing condition for the agricultural good: the demand of the agricultural good (which is a constant fraction of the world income) equals the supply of the argicultural good (which equals the total number of peasants).

The short-run equilibrium is described by (2), (4), (5), (6) and (7) that determine w_i, n_i, y_i, and p_i for a given labour distribution L_i and the parameters of the model $(\tau; \mu; \sigma)$.[6] However, the above system might yield several solutions at a given distribution of labour.[7] In what follows, it is shown that only one of them is in positive prices and quantities.

[4] The second equality sign holds, because workers income | or, in other words, labour cost | equals ¯rms sales in a region.

[5] The equilibrium condition for the foreign ¯rm is omitted due to Walras' law. This is the ¯rst step, where we di®er from Krugman (1991a). He drops the equilibrium condition for the agricultural sector instead. Our proceeding will allow us to summarize the whole model in a single equation.

[6] The parameters α and β will drop out on the way.

[7] To ¯nd an indication for multiple equilibria, insert ¯rst equation (5) into equations (6) and (7). Second, suppose $\sigma = 2$: One of the two emerging equations entails a polynomial of degree 3 in prices. It is known that an equation system with such a polynomial does not necessarily entail a unique solution. However, some solutions can be complex and thus economically irrelevant.

Proposition 1: The equation system (2), (4), (5), (6) and (7) has a unique solution for w_i, n_i, y_i, and p_i at a given labour distribution L_i.

Proof: First, we note that (6) and (7) can be rewritten as a system of excess demand functions g and f; where $g \equiv g_1 + g_2 - \bar{x}$ (with g_1 and g_2 being respectively the first and second term at the left hand side of (6)) and $f \equiv x_1p_1n_1 + x_2p_2n_2 - \mu$. It is easily seen that the excess demand functions g and f fulfill the gross substitute property. This means that g depends positively on the industrial price p_2; and f depends positively on both industrial prices p_1 and p_2 (at a given number of goods). However, for any excess demand system that fulfills the gross substitute property there exists a unique equilibrium price vector for a given number of goods in both regions (see proposition 17.F.3 in Mas-Colell, et.al. (1995), p. 613). Having shown that, all other endogenous variables are linear transformations of prices and the number of goods and must thus be unique, too. Q.E.D.

To facilitate the analysis, we define the relative nominal prices as $v \equiv \frac{p_2}{p_1}$ and the relative distribution of industries as $z \equiv \frac{n_2}{n_1}$. Because the relative distribution of industries equals the relative labour distribution according to equation (4) and relative prices of industrial goods equal relative industrial wages according to equation (2), v describes also relative wages ($v = \frac{w_2}{w_1}$) and z describes the relative distribution of labour ($z = \frac{L_2}{L_1}$). The system (6) and (7) can then be rewritten by using (5) as follows:

$$\frac{\mu\left(1 + \frac{1-\mu}{2} \cdot \frac{1}{n_1p_1\bar{x}}\right)}{(1 + tzv^{1-\sigma})} + \frac{t\mu\left(zv + \frac{1-\mu}{2} \cdot \frac{1}{n_1p_1\bar{x}}\right)}{(t + zv^{1-\sigma})} = 1 \tag{8}$$

$$\frac{1}{n_1p_1\bar{x}} = \frac{1 + vz}{\mu}$$

Substituting the second equation into the first, and solving for z, we obtain a condition that describes the goods market equilibrium as an implicit relationship of the relative labour distribution z and the relative nominal wages v. We call this condition, the short-run equilibrium condition $h(v;z)$:

$$h(v;z): \qquad z = \frac{1-\mu+(1+\mu)t^2-2tv^{\sigma}}{v\left[1-\mu+(1+\mu)t^2-2tv^{-\sigma}\right]} \tag{9}$$

The $h(v;z)$ function can be drawn in the v-z space.[8] The symmetry point $(v = 1; z = 1)$ will be a point of this function. The function can be either upward or downward sloping in the v-z space.

2.2 The Long-Run Equilibrium

In the long-run workers are allowed to be mobile, moving to the region which pays the highest real wage.[9] To characterize the migration process at any labour distribution, and thus the firm reallocation incentives, we write the arbitrage condition for the steady state equilibrium:

$$\omega_1 \equiv \frac{w_1}{P_1} = \frac{w_2}{P_2} \equiv \omega_2 \tag{10}$$

Using (1), (2), (4) and the definitions for v and z, we can rewrite this condition as an implicit function $k(v;z)$:

$$k(v;z): \qquad z = \frac{t - v^{\frac{1-\sigma}{\epsilon}}}{v^{1-\sigma}\left[t v^{\frac{(1-\sigma)}{\epsilon}} - 1\right]} \tag{11}$$

We call the above the equal real wage condition. The symmetry point $(v = 1; z = 1)$ will be a point of this function. Furthermore, the $k(v;z)$ function is always downward sloping in the positive orthant of the v-z space.

Whether a worker will migrate or not depends on whether the short-run equilibrium condition lies above or below the equal real wage condition. All cutting points of the two curves are interior steady state equilibria. It is obvious that the symmetry point is always a steady state. In what follows we derive the exact conditions under which workers will migrate towards the one region or the other. A diagrammatic illustration is provided later on.

The full characterization of the long-run equilibria is approached in several steps. First, a condition is derived that determines the parameter values for which the

[8] Without loss of generality, we define region 2 to be the region with the fewest workers (or an equal number of workers). Then, we only need to consider the range $0 \le z \le 1$.

[9] Real wage rates are identical to the value of the indirect utility function of a worker given a distribution of workers. This follows immediately from the definitons of the CES-price index and the indirect utility function, respectively.

symmetry point is a stable steady state.

Lemma 1: The symmetry point $v = 1; z = 1$ is a stable steady state equilibrium if and only if $c^{sb}(\mu;\tau;\sigma) < 0$; where

$$c^{sb}(\mu;\tau;\sigma) = \tau^{\sigma-1} - \frac{1-\mu-\sigma(1-\mu)^2}{1+\mu-\sigma(1+\mu)^2}$$

Proof: See appendix 1.

If the symmetry point is a stable steady state, then there will occur convergence of industry location in the neighbourhood of this point. If the symmetry point is unstable, i.e. $c^{sb}(\mu;\tau;\sigma) \geq 0$; there will occur "symmetry-breaking"[10] and the firms and workers start relocating unevenly in the plane (agglomeration).

Second, we examine the condition for the existence of a corner solution, i.e. $L_1 = 0$ or $L_2 = 0$. A corner solution is a stable equilibrium if and only if real wages are lowest in the region where all industries have disappeared.

Lemma 2: Total agglomeration (either $L_1 = 0$ or $L_2 = 0$) is a stable steady state equilibrium if and only if $c^{ta}(\mu;\tau;\sigma) > 0$, where

$$c^{ta}(\mu;\tau;\sigma) = 2 - \tau^{\mu\sigma}\left[(1+\mu)\tau^{\sigma-1} + (1-\mu)\tau^{-(\sigma-1)}\right]. \qquad (12)$$

Proof: See appendix 2.

The total agglomeration condition describes the set of parameters for which a region dries out of industries completely.[11] If the opposite inequality sign holds, i.e. $c^{ta}(\mu;\tau;\sigma) \leq 0$, then there will be convergence at $z = 0$, i.e. some firms will start relocating from the region with the industries to the region without the industries.

[10] Conditions for symmetry breaking are given in other models of the same type (see Venables, 1995, and Krugman and Venables, 1995).

[11] This condition, although differently derived, is identical to equation (26) in Krugman (1991a). The economic meaning of the parameters $\mu; \tau;$ and σ in this condition is the same as in Krugman (1991a) and therefore we do not discuss it further.

Third, a relationship between the agglomeration/convergence conditions from lemmas 1 and 2 is found.

Lemma 3: The condition for symmetry-breaking is a subset of the condition for total agglomeration. Thus, for any $(\mu;\tau;\sigma)$ with $0 < \mu \leq 1$, $\sigma > 1$, and $0 < \tau < 1$; the following is true:

$$c^{sb}(\mu;\tau;\sigma) \geq 0 \Rightarrow c^{ta}(\mu;\tau;\sigma) > 0 \text{ and}$$
$$c^{ta}(\mu;\tau;\sigma) \leq 0 \Rightarrow c^{sb}(\mu;\tau;\sigma) < 0$$

Proof: See appendix 3.

The importance of this lemma is that there is agglomeration at $z = 0$ and convergence at $z = 1$ for some parameters, but there are no allowed parameter values for which there is convergence at $z = 0$ and agglomeration at $z = 1$.

Finally, the maximum number of interior steady states is determined.

Lemma 4: The system (9) and (11) has at most one interior steady state z^* with $0 < z^* < 1$.

Proof: See appendix 4.

Using these four lemmas, the long-run equilibria can be fully characterized. Proposition 2 does exactly that.

Proposition 2: (i) Workers and firms tend to agglomerate completely in one of the two regions independently of the initial labour distribution, if

$$c^{sb}(\mu;\tau;\sigma) \geq 0:$$

(ii) Workers and firms tend to spread even in both regions (convergence) independently of the initial labour distribution, if

$$c^{ta}(\mu;\tau;\sigma) \leq 0.$$

(iii) There is an unstable intermediate steady state equilibrium at a firm distribution z^* with $0 < z^* < 1$ and two stable steady state equilibria at $z = 0$ and $z = 1$, if

$$c^{sb}(\mu;\tau;\sigma) < 0 < c^{ta}(\mu;\tau;\sigma).$$

Proof: (i) If $c^{sb}(\mu;\tau;\sigma) \geq 0$; then there is agglomeration at $z = 1$ (lemma 1). However, if that happens, then $c^{ta}(\mu;\tau;\sigma) > 0$ (lemma 3) and thus there must also be agglomeration at $z = 1$ (lemma 2). But then there cannot exist any convergence in between $0 < z < 1$; because this would require at least two interior steady states. This, however, contradicts lemma 4. Hence, if $c^{sb}(\mu;\tau;\sigma) \geq 0$, there must be agglomeration for any labour distribution z.

(ii) If $c^{ta}(\mu;\tau;\sigma) \leq 0$, then there is convergence at $z = 0$ (lemma 2). However, if that happens, then $c^{sb}(\mu;\tau;\sigma) < 0$ (lemma 3) and thus there must also be convergence at $z = 1$ (lemma 1). But then there cannot exist any agglomeration in between $0 < z < 1$; because this would require at least two interior steady states. This, however, contradicts lemma 4. Hence, if $c^{ta}(\mu;\tau;\sigma) \leq 0$, there must be convergence for any labour distribution z.

(iii) If $c^{sb}(\mu;\tau;\sigma) < 0$, there is convergence at $z = 1$ (lemma 1); then the schedule of $h(v;z)$ lies below the schedule of $k(v;z)$ in v-z space for z slightly below 1. If $c^{ta}(\mu;\tau;\sigma) > 0$, there is agglomeration at $z = 0$ (lemma 2). Then the schedule of $h(v;z)$ lies above the schedule of $k(v;z)$ in v-z space for $z = 0$. Hence, there must be at least one cutting point z^* of the two schedules in between 0 and 1 (intermediate value theorem). Because of lemma 4, there is exactly one. z^* is unstable, because $z = 0$ and $z = 1$ are stable at this parameter constellation. Q.E.D.

Proposition 2 can be most easily demonstrated in figure 1. The three cases in proposition 2 correspond to the three panels of figure 1.

Figure 1: (about here)

Panel (a) shows that the short-run equilibrium condition $h(v; z)$ is below the equal real wage condition $k(v; z)$ for all labour distributions. Real wages are higher in the larger region and workers and ¯rms of the smaller region have an incentive to move, reenforcing thus the agglomeration process. This corresponds to case (i) in proposition 2.

Panel (b) shows that the short-run equilibrium condition $h(v; z)$ is above the equal real wage condition $k(v; z)$ for all labour distributions. Real wages are higher in the smaller region and the workers and ¯rms tend to spread evenly in the plane (convergence). This corresponds to case (ii) in proposition 2.[12]

Panel (c) shows that the short-run equilibrium condition $h(v; z)$ is below the equal real wage condition $k(v; z)$, if the initial labour distribution z is below a critical mass z^*. Then real wages are higher in the bigger region and the few industries in the smaller region relocate to the bigger region (total agglomeration). On the contrary, the short-run equilibrium condition $h(v; z)$ is above the equal real wage condition $k(v; z)$, if the initial labour distribution z is above the critical mass z^*. Then, real wages are higher in the smaller region and industries start to relocate evenly in the plane (convergence). This corresponds to case (iii) in proposition 2.

The third case will be called a \poverty trap". In the \poverty trap", regions that lack the critical mass of industries are stuck in this state because there is no market force that could foster industrialization. If the region had more industries to begin with, market forces would enforce a convergence process.[13]

[12] The cases (i) and (ii) are reported in form of simulations in ¯gure 1 of Krugman (1991a). Note that we derive the schedules of the functions $h(v; z)$ and $k(v; z)$ analytically. The precise curvature of the two schedules in ¯gure 1 is suggestive.

[13] The \poverty trap" is usually discussed in the context of growth models (see Barro and Sala-i-Martin, 1995, p. 49®, for an overview.) There, a poverty trap means that some countries are stuck with a low level of the capital stock. If they had enough capital to begin with, they would converge to the advanced nations. The reason is the switch from a diminishing returns to scale to an increasing returns to scale technology. This is somewhat related to the result in this model. Here the poverty trap arises, if a certain share of the increasing returns to scale and the constant returns to scale sector is prevailed in the economy. A \poverty trap" arises also in Matsuyama (1991). His economic

The “poverty trap” shows that the location of industries might depend on initial conditions. Even if all parameters – like preferences and technology – are the same, convergence or agglomeration may take place depending on the initial distribution of industries. History matters, not only in the determination of the region to be the “winner”, but also in the determination of whether industries tend to agglomerate or spread evenly in the plane.[14] For example the rather underdeveloped Eastern Germany lost a large proportion of its industries and its workers after unification with West Germany.[15] However, the integration of the rather homogeneous European countries seems to enforce a convergence process.[16] This observation may justify that the EU requires new members to have a sufficient stage of development.[17]

Finally, the “poverty trap” case will not appear, if the economy starts developing with a symmetric distribution of industries. It follows from proposition 2 that the condition of “symmetry-breaking” is then sufficient to fully characterize the model.

3 Conclusion

This paper gives a complete analytical solution to a monopolistic competition, increasing returns to scale model with transport cost amd mobile labour (Krugman, 1991a). It is shown that this model has a unique short-run equilibrium. Furthermore, it is shown that the long run equilibrium is characterized by three groups of parameter ranges: (i) industries agglomerate completely in one of the two regions; (ii) industries tend to spread evenly in the plane; (iii) it depends on the initial distribution of industries whether agglomeration or convergence occurs (“poverty trap”).

---

interpretation of the “poverty trap” is similar to the one in this paper. However, the underlying story is very different.

[14] Krugman (1991b) shows that history or expectations might matter for the determination of the “winner” region, if a forward looking migration process is assumed.

[15] See Lange and Pugh (1998).

[16] See Ben-David (1993).

[17] The model predicts also that an integration process which lowers trade cost may cause the economy to move from the convergence case to the “poverty trap” case. This finding follows from proposition 2. To see this, start from the convergence case (ii). Note that a sufficiently small increase in τ (decrease of trade cost) induces a rise in $c^{ta}(\mu;\tau;\sigma)$ and a violation of the condition $c^{ta}(\mu;\tau;\sigma) < 0$ (which is most easily seen by simulating numerically equation (12)), but not necessarily a violation of $c^{sb}(\mu;\tau;\sigma) < 0$, because we proved that the parameter set of case (iii) is non-empty. Therefore, a decrease in trade cost τ may lead to a move from case (ii) (convergence case) to case (iii) (poverty trap case) for some parameters μ and σ and some initial conditions z_0 with $0 \leq z_0 \leq 1$.

The later ¯nding shows that initial conditions matter in Krugman's (1991a) model not only for the determination of the \winner" region, but also for the determination of agglomeration or convergence. However, the \poverty trap" case will not appear, if the economy starts developing with a symmetric distribution of industries. The condition of \symmetry-breaking" is then su±cient to fully characterize the model.

References

d'Aspremont, Claude, Rudolphe Dos Santos Ferreira, and Louis-André Gérard-Varet, (1996), On the Dixit-Stiglitz Model of Monopolistic Competition, American Economic Review, Vol. 86, p. 623-629.

Barro, Robert J., and Xavier X. Sala-i-Martin (1995), Economic Growth, Mc Graw Hill, New York.

Ben-David, Danny, (1993), Equalizing exchange: Trade liberalization and income convergence, Quarterly Journal of Economics, Vol. 108, p. 653-79.

Dixit, Avinash K. and Joseph E. Stiglitz (1977), Monopolisitic Competition and Optimum Product Diversity, American Economic Review, Vol. 67, 297-308.

Dixit, Avinash K. and Joseph E. Stiglitz (1993), Monopolistic Competition and Optimal Product Diversity: Reply, American Economic Review, Vol. 83, No. 1, p. 302-304.

Fujita, Masahisa and Jacques-Francois Thisse (1996): Economics of Agglomeration, CEPR Discussion Paper, No. 1344.

Helpman, Elhanan and Paul R. Krugman (1985), Market Structure and Foreign Trade, MIT Press, Cambridge.

Itô, Kiyosi (1993), Encyclopedic Dictionary of Mathematics, Vol. 1, 2nd ed., MIT Press, New York.

Krugman, Paul R. (1991a), Increasing Returns and Economic Geography, Journal of Political Economy, Vol. 99, No. 3, p. 483-499.

Krugman, Paul R. (1991b), History versus Expectations, Quarterly Journal of Economics, Vol. 106, p. 651-667.

Krugman, Paul R. and Anthony J. Venables (1995), Globalization and the Inequality of Nations, Quarterly Journal of Economics, Vol. 110, No. 4, p. 857-880.

Lange, Thomas, and Geoffrey Pugh, (1998), The economics of German unification, Edward Elgar, Cheltenham, UK, and Northhampton, USA.

Mas-Colell, Andreu, and Michael D. Whinston and Jerry R. Green (1995), Microeconomic Theory, Oxford University Press, New York and Oxford.

Martin, Philippe and Carol Ann Rogers (1995), Industrial Location and Public Infrastructure, Journal of International Economics, Vol. 39, No. 3/4, p. 335-351.

Matsuyama, Kiminori (1991), Increasing Returns, Industrialization, and Indeterminacy of Equilibria, Quarterly Journal of Economics, Vol. 106, 617-650.

Myrdal, Gunnar, (1957), Economic theory and under-developed regions, Duckworth, London.

Norman, Victor D. and Anthony J. Venables (1995), International Trade, Factor Mobility, and Trade Costs, The Economic Journal, Vol. 105, p. 1488-1504.

Urban, Dieter M., (1998), Neoclassical growth, manufacturing agglomeration, and terms of trade, mimeo.

Venables, Anthony J. (1995), Economic Integration and the Allocation of Firms, American Economic Review, Vol. 85, No. 2, p. 296-300.

Venables, Anthony J. (1996), The Location of Vertically Linked Industries, International Economic Review, Vol. 37, No. 2, p. 341-359.

Yang, Xiaokai and Heijdra, Ben J. (1993), Monopolistic Competition and Optimal Product Diversity: Comment, American Economic Review, Vol. 83, No. 1, p. 295-301.

Appendix

Appendix 1:

Taking the derivative of (9) and evaluating at the symmetry point, yields the following expression:

$$\left.\frac{dz}{dv}\right|_{h(1;1)} = \frac{1 - \mu + (1+\mu)\,t^2 - 2t\,(1 - 2\sigma)}{(1 - t^2)\,\mu - (1 - t)^2}$$

Taking the derivative of (11) and evaluating at the symmetry point, yields:

$$\left.\frac{dz}{dv}\right|_{k(1;1)} = \frac{(1 - \sigma)\,[1 - \mu + t\,(1+\mu)]}{(1 - t)\,\mu}$$

The equal distribution is stable (convergence), if

$$\left.\frac{dz}{dv}\right|_{h(1;1)} < \left.\frac{dz}{dv}\right|_{k(1;1)}:$$

Equalizing the two derivatives $\left.\frac{dz}{dv}\right|_{h(1;1)} = \left.\frac{dz(v)}{dv}\right|_{k(1;1)}$ and solving for t yields three solutions. One is always negative ($t = -1$) and therefore economically irrelevant. The other is $t = 1$; i.e. factor prize equalization holds in the absence of transport cost. And the third is:

$$t = \frac{1 - \mu - \sigma\,(1 - \mu)^2}{1 + \mu - \sigma\,(1 + \mu)^2}$$

This expression used in the conditions above gives the lemma 1. Q.E.D.

Appendix 2:

The equal real wage condition for $L_1 = 0; L_2 = \mu$ is given in equation (11):

$$\frac{1}{z} = \frac{v^{1-\sigma}\left[t v^{\frac{1-\sigma}{\mu}} - 1\right]}{t - v^{\frac{1-\sigma}{\mu}}}$$

or equivalently by using $(1/z) = L_1/L_2 = 0$ and the definition of $v = w_2/w_1$ and $t = \tau^{\sigma-1}$:

$$\frac{w_2}{w_1} = \tau^{\mu}$$

Using the definition for real wages of equation (10) and noting that the relative real wage is one gives the condition for the relative price indices:

$$\frac{P_2}{P_1} = \tau^{\mu}$$

From the short-run equilibrium condition (9), it follows that

$$\frac{1}{z} = \frac{v\,[1 - \mu + (1+\mu)\,t^2 - 2tv^{-\sigma}]}{1 - \mu + (1+\mu)\,t^2 - 2tv^{\sigma}};$$

which yields for $(1/z) = 0$:

$$\frac{w_2}{w_1} = v = \left(0.5\,(1-\mu)\,t^{-1} + 0.5\,(1+\mu)\,t\right)^{-(1/\sigma)}.$$

The relative real wage at the total agglomeration point can thus be written as:

$$\frac{w_1 P_2}{w_2 P_1} = \tau^{\mu}\left(0.5\,(1-\mu)\,\tau^{-(\sigma-1)} + 0.5\,(1+\mu)\,\tau^{\sigma-1}\right)^{1/\sigma}$$

By definition of total agglomeration the real wage in region 1 needs to be smaller than the real wage in region 2, i.e.:

$$\frac{w_1 P_2}{w_2 P_1} < 1$$

Consequently the condition can be stated as in lemma 2. Q.E.D.

Appendix 3:

First, it is shown that the boundary planes of the two conditions from lemmas 1 and 2 do not intersect for allowed parameter constellations of τ, σ, μ:

$$c^{ta}(\mu, \tau, \sigma) = 0$$

$$c^{sb}(\mu, \tau, \sigma) = 0$$

The second equation is solved for τ and then plugged into the first expression to give:

$$\left(\frac{(1-\mu)[1-\sigma(1-\mu)]}{(1+\mu)[1-\sigma(1+\mu)]}\right)^{\frac{\mu\sigma}{\sigma-1}}\left(\frac{(1-\mu)[1-\sigma(1-\mu)]}{1-\sigma(1+\mu)} + \frac{(1+\mu)[1-\sigma(1+\mu)]}{1-\sigma(1-\mu)}\right) = 2$$

This condition is dealt numerically. A grid search procedure in Mathematica shows that $\mu = 0$ or $\sigma = 0$ are the only two solutions to this equation. This implies that there is no cutting point of the two equations for $0 < \mu \leq 1$, and $\sigma > 1$.

Second, it is readily checked that there exist parameter constellations of τ, σ, and μ, such that the three sets defined by

$$\begin{aligned} c^{ta}(\mu, \tau, \sigma) &\leq 0 \text{ and } c^{sb}(\mu, \tau, \sigma) < 0 \\ c^{ta}(\mu, \tau, \sigma) &> 0 \text{ and } c^{sb}(\mu, \tau, \sigma) \geq 0 \\ c^{ta}(\mu, \tau, \sigma) &> 0 \text{ and } c^{sb}(\mu, \tau, \sigma) < 0 \end{aligned}$$

are non-empty. If the boundary planes do not intersect, there can only be three non-empty sets. See the illustration in figure 2.

Figure 2: (about here)

Figure 2 depicts the boundary planes $c^{ta}(\mu, \tau, \sigma) = 0$ and $c^{sb}(\mu, \tau, \sigma) = 0$ for any arbitrary value of τ in the μ-σ space such that they do not intersect. Then there can only be defined three distinct subspaces by these two boundary planes. This implies that the joint conditions $c^{ta}(\mu, \tau, \sigma) \leq 0$ and $c^{sb}(\mu, \tau, \sigma) \geq 0$ describe an empty set. Hence, for any (μ, τ, σ) in the parameter range holds:

$$\begin{aligned} c^{sb}(\mu, \tau, \sigma) \geq 0 &\Rightarrow c^{ta}(\mu, \tau, \sigma) > 0 \text{ and} \\ c^{ta}(\mu, \tau, \sigma) \leq 0 &\Rightarrow c^{sb}(\mu, \tau, \sigma) < 0 \end{aligned}$$

Or in words: the condition for symmetry-breaking is a subset of the condition for total agglomeration. (If symmetry-breaking occurs, then the system always ends up with total agglomeration.) Q.E.D.

Appendix 4:

If the short-run equilibrium condition (9) and the equal real wage condition (11) are set equal to each other and some terms are rearranged, a function in the variable v emerges which needs only to be defined for positive wages ($v > 0$).

$$\sum_{i=1}^{3} a_i v^{b_i} + a_4 = 0; \tag{13}$$

where

$$\begin{aligned} a_1 &= -a_2 = (1-\mu)\left(1 - t^2\right) > 0 \\ a_3 &= -a_4 = t(1+\mu)\left(1 - t^2\right) > 0 \\ b_1 &= -\sigma < 0 \\ b_2 &= \frac{1-\sigma}{\mu} < 0 \\ b_3 &= \frac{1-\sigma(1+\mu)}{\mu} < 0 \end{aligned}$$

If this equation was a polynomial, Descartes' rule of sign[18] would imply that this gives at most 3 solutions for $v > 0$. However, Descartes' rule of sign can still be applied in the following way: Suppose the $b_i; i = 1; 2; 3$ are rational numbers and $N \in \mathbb{N}$ is the common denominator of them. Then set $\hat{b}_i = b_i N$ and define a ξ such that $v = \xi^N$. The equation (13) can thus be rewritten as

$$\sum_{i=1}^{3} a_i \xi^{\hat{b}_i} + a_4 = 0 \tag{14}$$

which is a polynomial and Descartes' rule of sign applies. If the polynomial (14) has at most three solutions for ξ; then it must also have at most three solutions for v (because there is a one to one mapping between v and ξ). One of them is $v = 1$. Suppose that the other two were both interior solutions, i.e. $v \in (0; 1)$: If that was true, then there should also exist two solutions for $v > 1$, as any interior solution for $v \in (0; 1)$ must have a corresponding steady state for $v > 1$ (the result must be independent of the label of the region). However, then there would exist more than 3 solutions. Thus, at most one solution is interior. Q.E.D.

[18] See Itô (1993), p. 36 for a statement of Descartes' rule of sign.

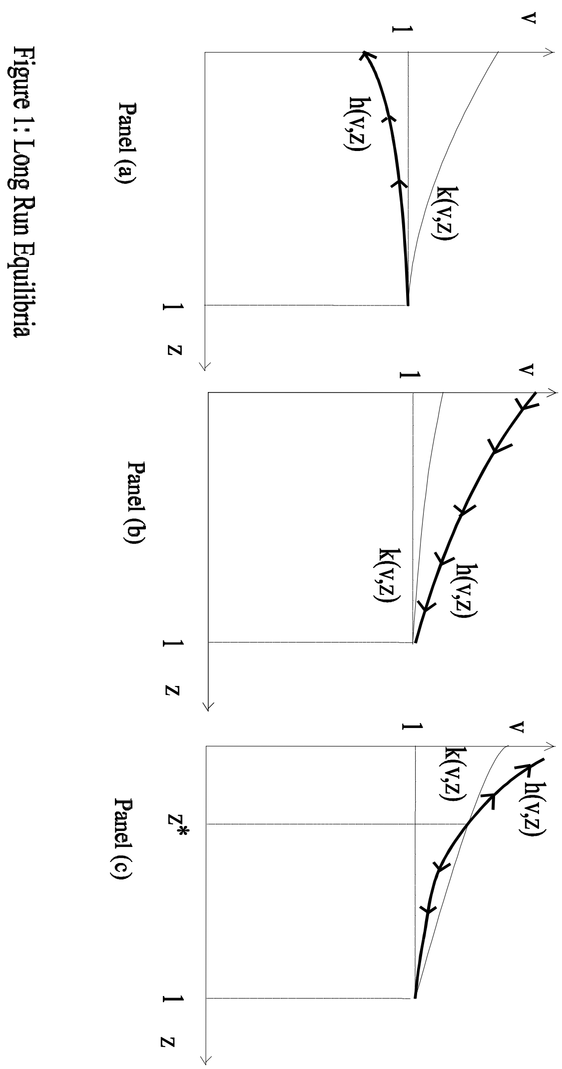
Figure 1: Long Run Equilibria
Panel (a)
Panel (b)
Panel (c)
h(v,z)
k(v,z)
v
z
1
z*

Figure 1:

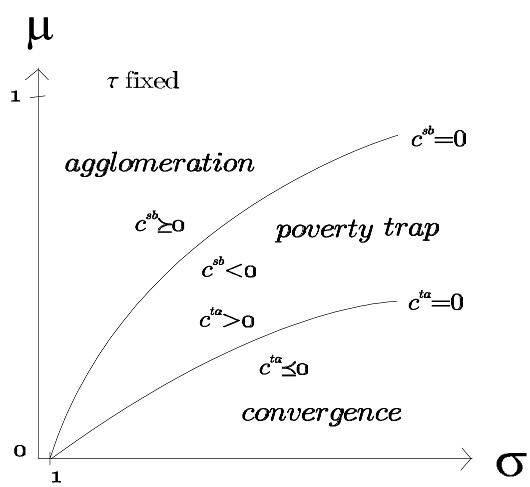

Figure 2: Parameter Range

Figure 2: