The Size of Regions

By

Elhanan Helpman

Working Paper #14-95

The Foerder Institute for Economic Research

THE SIZE OF REGIONS

by

Elhanan Helpman[1]

Working Paper No.14-95

May, 1995

[1] The Eitan Berglas School of Economics, Tel-Aviv University and the Canadian Institute for Advanced Research.

1 Introduction

Regional scientists and urban economists have studied many forces that shape urban and regional structures. They have emphasized the availability of usable land, of local amenities such as beaches or climate, the provision of local public goods such as roads and schools, commuting costs, congestion, pollution, localized economies of scale in production, and more. These features help to understand differences in layouts of cities and differences in city size. They also help to understand regional differences, such as population density or the concentration of manufacturing facilities. And they can be used to explain differences in the economic geography of countries.

Some of these elements, such as the availability of land or congestion, push for the dispersion of economic activity. Others, such as the provision of local public goods or localized economies of scale in production, push for agglomeration. The observed outcomes reflect the balance of these forces. Tolley and Crihfield (1987) suggest that market forces tend to generate too much agglomeration, and their view is quite common among urban economists.

My purpose in this chapter is to have a fresh look at the problem of agglomeration by examining a model of two regions in which the availability of housing (land) is the main driving force for dispersion of economic activity, as is common in regional and urban economics. In addition I will introduce an industrial sector that supplies differentiated products that are traded between the regions at a cost. These goods are produced with brand specific economies of scale. As a result, the industrial sector provides the main driving force for agglomeration. An industrial sector of this sort has been extensively utilized by Paul Krugman in his work on economic geography (e.g., Krugman (1991)).

Although each one of these centrifugal and centripetal forces has been studied before, their combination provides some new insights. In particular, as simple as this model may be, it produces interesting equilibrium outcomes: each region may be occupied in proportion to (or as a function of) its relative amount of housing; or

1

regions may be unequally populated even when they have equal amounts of housing. Importantly, in the last type of equilibrium there is too little agglomeration while the degree of agglomeration is efficient in the former case. The simplicity of the model enables us to explain these results by means of the intensity of preferences for differentiated products, the degree of substitution across brands, and the level of transport costs.

The formal model is developed in section 2. Section 3 is then devoted to an analysis of equilibrium outcomes. It is shown that whenever transport costs are low there exists a unique stable equilibrium in which both regions are occupied. Population density is determined by the relative availability of housing. If the amount of housing is the same in both regions half of the population resides in each one of them. When the demand for housing is high or the elasticity of substitution across brands is high, the same type of unique, stable equilibrium prevails for all levels of transport costs. When the demand for housing is low, however, or the elasticity of substitution across brands is low, there exist two asymmetric equilibria with regions of unequal size in each one of them, even when the supply of housing is the same in both locations. The higher the transport costs the more unequal the regions. And the inequality in the size of regions rises very rapidly with transport costs. These results, about the link between agglomeration and transport costs, are different from Krugman's (1991). Explanations for the differences are provided in section 6, in which the two models are compared in some detail.

Before the reader gets to section 6, however, there are two additional sections. In one of them I examine welfare properties of the resulting equilibria (section 4). An important conclusion is that there is either too little agglomeration or the amount of agglomeration is just right. The former occurs at the asymmetric equilibria. It follows that whenever regions differ in size not as a result of inherent differences but rather as a result of endogenous economic pressures, it is desirable to make their inequality even more pronounced.

2

As mentioned at the beginning of this introduction, there are potentially many forces of agglomeration and dispersion. My model focuses on just two of them. In order to see whether it produces results that are peculiar to economies in which differentiated products that are costly to transport provide the incentive to agglomerate, I examine in section 5 an alternative model with no transport costs and with regional external economies of scale in the production of a homogenous product. The model produces equilibria that are similar to the equilibria of the model from section 2. This shows that there is nothing peculiar in my results.

Much of what is discuss in this paper applies to cities, regions and countries. For concreteness, however, I will refer to each location as a region, where a region is characterized by a supply of housing (although we could replace housing by any other good or service that is not traded across regions). The differentiated products can be produced anywhere, except that they are traded across regions at a cost. And finally, people move across regions in search of higher welfare. Clearly, this type of model applies also to cities, although in the study of cities one may want to specify in more detail their layouts. And it also applies to countries that allow free migration. These interpretations should be kept in mind in what follows.

2   Basic Model

In this section I develop a simple model of regions that focuses on the tradeoff between the supply of housing – which is not traded across regions and serves as a centrifugal force – and the manufacturing of differentiated industrial products with increasing returns to scale – which are footloose but costly to transport and serve as a centripetal force. Suppliers of manufactured products engage in monopolistic competition while suppliers of housing are competitive.

Each region has a fixed stock of housing and people can choose in which region to reside. When a person lives in a region she purchases in this region housing

services and all manufactured products; both locally produced and imported brands. A person works in the region in which she lives. As a result, labor supply in a region is determined by its population. And labor supply determines in turn the local output of differentiated products, including the number of brands. A more populated region is supplied with a better variety choice of locally manufactured goods while imported brands are costly to transport. This raises the local standard of living. On the other hand, housing costs are higher in a more populated region, which reduces the local standard of living. These two forces produce a tension between the desire to agglomerate and the desire to disperse. Individuals migrate from regions with a low standard of living to regions with a higher standard of living. Therefore in equilibrium the standard of living is the same in all populated regions.

My analysis is confined to the case of two regions and it focuses on the following questions:

- What determines the distribution of people across regions?

- Do lower transport costs lead to more or less agglomeration?

- Is the resulting degree of agglomeration efficient relative to the market structure?

In order to address these questions I now develop a formal model.

There are N identical individuals. They have a Cobb-Douglas utility function

$$u = h^\beta d^{1-\beta},\tag{1}$$

where h represents consumption of housing services and d represents a consumption index of differentiated products. This index is given by

$$d = \left[\int_0^n x(j)^\alpha dj\right]^{1/\alpha}, \quad 0 < \alpha < 1,\tag{2}$$

where n is the available number (measure) of brands and $x(j)$ represents consumption of brand j. As is well known, these preferences imply a constant price elasticity of

4

demand $\epsilon = 1/(1 - \alpha) > 1$ for each brand, where ϵ also represents the elasticity of substitution between them. As a result the price of a brand equals $1/\alpha$ times marginal costs.

Assume that it takes $a + x$ units of labor to manufacture x units of a single brand, where a represents fixed costs in terms of labor. Then in region i with a wage rate w_i the mill price of each locally produced brand equals

$$p_i = \frac{1}{\alpha} w_i . \tag{3}$$

In addition, free entry of manufacturers ensure zero profits on each brand, or

$$p_i = \left(\frac{a}{x} + 1 \right) w_i ,$$

which states that the price equals average costs. Together with (3) this free entry condition implies a constant output level for each brand, the same in both regions,

$$x = \frac{\alpha a}{1 - \alpha} . \tag{4}$$

The production function $a + x$ also implies that labor demand in region i equals $(a + x)n_i$, where n_i represents the number of brands that are manufactured in region i. Since labor demand equals labor supply in each region, then in view of (4) a region that is populated with N_i individuals, each one endowed with one unit of labor, manufactures the following number of brands:

$$n_i = \frac{1 - \alpha}{a} N_i . \tag{5}$$

2.1 Relative prices

Now assume that a resident of region i pays the mill price p_i for every locally manufactured brand and the price tp_j, $t > 1$, for a brand imported from region $j \neq i$. I will use the interpretation of 'melting iceberg' transport costs. Namely, in order to consume one unit of an imported brand an individual has to buy from the supplier t units of the good. Next, let E_i represent aggregate income of the residents in region

i. Then it follows from the preferences, the structure of transport costs, and from (4), that the condition that aggregate demand for a brand produced in region 1 equals its supply can be represented by

$$\frac{\alpha a}{1-\alpha} = \frac{p_1^{-\epsilon}}{n_1 p_1^{1-\epsilon} + n_2 \left(tp_2\right)^{1-\epsilon}}(1-\beta)E_1 + \frac{t\left(tp_1\right)^{-\epsilon}}{n_1 \left(tp_1\right)^{1-\epsilon} + n_2 p_2^{1-\epsilon}}(1-\beta)E_2 , \qquad (6)$$

where the left hand side represents supply and the right hand side represents demand. A similar condition holds in the other region.

Each individual spends a fraction β of his income on housing (due to the Cobb-Douglas preferences). Therefore the aggregate value of housing services equals βE, where $E = \sum_i E_i$. On the other hand, aggregate spending equals aggregate income, which is composed of labor income $\sum_i w_i N_i$ plus income from housing βE. It follows that aggregate income from housing equals $\beta \left(\sum_i w_i N_i\right)/(1-\beta)$. I assume that the housing stocks are equally owned by all individuals. Therefore income from housing by residents of region i equals the fraction N_i/N of the total income from housing. As a result, spending by residents of region i equals

$$E_i = w_i N_i + \frac{\beta N_i}{(1-\beta)N} \sum_k w_k N_k . \qquad (7)$$

Substituting this equation together with (3) and (5) into (6), we obtain the reduced form equilibrium condition

$$1 = \frac{fq^{1-\epsilon}}{fq^{1-\epsilon} + (1-f)\,t^{1-\epsilon}} \left[1 - \beta + \beta\left(f + \frac{1-f}{q}\right)\right]$$

$$+ \frac{(1-f)\,(tq)^{1-\epsilon}}{f\,(tq)^{1-\epsilon} + 1 - f} \left[\frac{1-\beta}{q} + \beta\left(f + \frac{1-f}{q}\right)\right] , \qquad (8)$$

where $f = N_1/N$ represents the fraction of individuals that reside in region 1 and $q = p_1/p_2 = w_1/w_2$ represents the relative mill price of a brand in region 1 as well as the relative wage in region 1.

The right hand side of (8) declines in q. Therefore this equation provides a unique solution to relative prices for each allocation of the population across regions; i.e., q is a *function* of f. Inspection of (8) reveals the following properties:

6

(a) The relative price depends on the fraction of people that live in each region and, importantly, it does not depend on the population's size or the supply of housing.

(b) When half of the population lives in each region the relative price equals one, independently of transport costs.

The first property is self evident. The second property can be inspected by substituting $f = 1/2$ and $q = 1$ on the right hand side of (8). It stems from the fact that equally populated regions supply the same number of brands and the same quantity of each. In addition, income per capita is the same in both regions if and only if the relative price equals one.[1] When income per capita is equalized, the composition of demand for differentiated products in terms of local versus imported goods is the same in each region. Therefore we have both symmetric demands and symmetric supplies, ensuring market clearing. As for housing, the residents of a region spend a fraction β of their income on housing. In this event the aggregate housing budget is the same in both regions. If the supply of housing is also the same in both, then housing prices are equalized. Otherwise the price of housing is higher in the region with a lower stock of housing, but the value of the housing stock is the same in both.

2.2 Relative utilities

Region i's housing stock provides a fixed flow of services H_i. As a result consumption of housing per resident equals H_i/N_i. In addition, spending per resident on manufactured products equals $(1 - \beta)E_i/N_i$. Therefore the utility level of an individual that resides in region i equals (see (1) and (2))

$$u_i = \left(\frac{H_i}{N_i}\right)^\beta \left[\frac{(1-\beta)E_i}{N_i P_{di}}\right]^{1-\beta}, \tag{9}$$

---

[1] Recall that income per capita from the ownership of housing is always equal while labor income per capita depends on the wage rate, and relative wages are also represented by q.

where P_{di} is the price index of differentiated products in region i, given by

$$P_{di} = \left[n_i p_i^{1-\epsilon} + n_j \left(t p_j \right)^{1-\epsilon} \right]^{\frac{1}{1-\epsilon}}, \quad j \neq i.$$

Using the pricing condition (3), the equilibrium number of brands (5), and the levels of spending (7), we compute from (9) the relative utility level $v = u_1/u_2$,

$$v = \left(\frac{H_1}{H_2} \frac{1-f}{f} \right)^\beta \left[\frac{(1-\beta)q + \beta(fq+1-f)}{1-\beta+\beta(fq+1-f)} \right]^{1-\beta} \left[\frac{fq^{1-\epsilon} + (1-f)t^{1-\epsilon}}{f\,(tq)^{1-\epsilon} + 1 - f} \right]^{\frac{1-\beta}{\epsilon-1}}. \quad (10)$$

Solving from (8) the relative price for each allocation of people f and substituting the result into (10) enables us to trace out relative utility levels for each f. This relationship is useful for the study of equilibria[2].

3 Region Size

In order to identify equilibrium configurations, we now study the relationship between relative utility levels and the fraction of people that live in region 1. The equilibrium compositions of region size consist of those values of f that ensure equal utility levels in each region, or a higher utility level in a single region that contains the entire population.

To begin with, consider the limiting case with no transport costs; i.e., $t \rightarrow 1$. In this case the relative price q equals one for all values of f (see (8)). As a result, relative utilities are given by (see (10))

$$v = \left(\frac{H_1}{H_2} \frac{1-f}{f} \right)^\beta.$$

It follows that the relative utility in a region declines with the relative size of its population. This relationship is exhibited in Figure 1 for the case of equal housing stocks in both regions. Here there exists a unique equilibrium configuration, described by point A, at which half of the population lives in each region. At this allocation the

---

[2]Krugman (1991) uses a similar approach.

utility level is the same everywhere. With any other distribution of individuals across regions the utility level is higher in the smaller region. I assume that individuals that live in a region with a lower utility level gradually migrate to the region with the higher utility level. This adjustment process will be used to study the stability of equilibria. It follows that whenever regions differ in population size individuals migrate from the larger to the smaller region, as indicated by the arrows. We conclude that point A describes a unique globally stable equilibrium. Finally, if regions have different stocks of housing, the unique stable equilibrium consists of an allocation of people that is proportional to each region's housing stock.

To understand this result, observe that in the absence of transport costs an individual has access to all brands of the differentiated product at prices that do not depend on the region in which he chooses to live. His income is also the same in either one of the regions. Therefore he migrates to the region with the lower housing costs. But the region with disproportionately fewer residents has the lower price of housing. The result is that people migrate from the region with a disproportionately larger population to the region with a disproportionately smaller population, until each region's population becomes proportional to its housing stock.

The next thing to observe from (8) is that as the fraction of people living in the first region approaches zero, the relative price q approaches $t^{-(\epsilon-1)/\epsilon}$, which is finite for finite transport costs. In view of this fact (10) implies that the relative utility of residents of region one approaches infinity as the fraction of people residing in region one approaches zero. This means that for all finite levels of transport costs the v curve is asymptotic to the vertical axis, as shown in Figure 1. My simulations show that the curve is also downward slopping for all levels of transport costs whenever the elasticity of substitution between brands ϵ is large or the expenditure share on housing β is large, so that $\beta\epsilon > 1$. In this event there exists a unique stable equilibrium in

which a region's population is a function of its relative housing stock.[3] On the other hand, whenever the elasticity of substitution or the expenditure share on housing are small, so that $\beta\epsilon < 1$, a region's relative welfare rises with its relative population for close to equal distributions of the population across regions. Figure 2 depicts an extreme outcome, when transport costs are infinite (and the housing stock is the same in each region). In this case there exist three equilibria. One in which half of the population lives in each region and two in which all people live in one region, either region 1 or 2. As can be seen from the figure, however, when the population is not the same in both regions, welfare is higher in the larger region. As a result people migrate from the smaller to the larger region, as indicated by the arrows in Figure 2. The end result is that everyone lives in the same region. It follows that the two equilibria in which a single region is occupied are locally stable while the equilibrium in which both regions are occupied is unstable. Moreover, whichever region has more people to begin with grows in size until it absorbs the entire population.

Our discussion of the case of unbounded transport costs has revealed the important roles of the elasticity of substitution and the intensity of preferences for housing. We have seen that whenever they are large both regions are occupied in equilibrium, while only one region is occupied when they are small. To understand this result, observe that with unbounded transport costs an individual that resides in region i consumes only brands that are manufactured in that region. If his elasticity of substitution between brands is high, the individual cares little about the available variety choice. Therefore he is not particularly attracted to a densely populated region with many brands of the differentiated product. But he is attracted to a

---

[3]Equality of utility levels implies the equilibrium relationship

$$\frac{N_1}{N_2} = \left(\frac{H_1}{H_2}\right)^{\frac{\beta(\epsilon-1)}{\beta\epsilon-1}}.$$

Since the exponent on the right hand side is larger than one, it follows that whenever the housing stocks differ across regions the equilibrium distribution of people is skewer than the distribution of housing.

region with low housing costs, which happens to be thinly populated. And the latter attraction is more important the more he prefers housing over differentiated products (the larger β). These pressures lead to an equilibrium in which the entire housing stock is used in equilibrium, which requires both regions to be occupied. On the other hand, individuals with a low elasticity of substitution between brands are willing to pay a premium for a large variety choice. They are attracted to a densely populated region with many brands of the differentiated product. This attraction is even more powerful when they care little about housing, because in densely populated regions housing costs are high. Individuals of this type tend to bunch together, producing large agglomerations. And in equilibrium all of them end up living in the same region.

We have seen that whenever brands of the differentiated product are highly substitutable for each other and the demand for housing is high (i.e., $\beta\epsilon > 1$), the relative utility in a region declines with its population when either transport costs are nil or infinite. In both cases there exists a unique globally stable equilibrium in which the population of a region is related to its housing stock, as depicted in Figure 1. Moreover, my simulations show that the relative utility in a region declines with its relative population for all intermediate values of transport costs. As a result there exists in this case a unique globally stable equilibrium for every level of transport costs, and both regions are occupied in these equilibria. When the housing stock is the same in both regions half of the population lives in each one of them.[4]

Next observe that whenever brands of the differentiated product are poorly substitutable for each other and the demand for housing is low (i.e., $\beta\epsilon < 1$), the relative utility in a region declines with its population when transport costs are negligible and rises with its population when transport costs are prohibitive. In the former case there exists a unique stable equilibrium in which both regions are occupied while in the latter case in a stable equilibrium the entire population lives in one region. These

---

[4]Observe that whenever $f = 1/2$ and $q = 1$ the equilibrium condition for the relative price (8) is satisfied for all transport costs and (10) implies that the utility level is the same in both cities independently of transport costs.

extreme cases show that whenever $\beta\epsilon < 1$ the degree of agglomeration depends on transport costs; it is low for negligible transport costs and high for prohibitive transport costs. However, for the intermediate range of transport costs the relative utility of a region declines with its population when the region is very small, rises with its population when the region becomes bigger, and eventually declines again when a large fraction of the population lives in the region, as exhibited in Figure 3 (in this figure too the housing stocks are taken to be equal in both regions). Evidently, now there are three equilibria, indicated by points A, B and C, except that the middle equilibrium at point B is unstable (see the arrows of the adjustment process). An interesting property of the stable equilibria is that despite the fact that the same stock of housing is available in both regions the regions differ in size. Moreover, simulations show that the relative size of the larger region increases with transport costs. So this is a case in which the tradeoff between the centrifugal forces of housing and the centripetal forces of differentiated products is most significant. In this equilibrium the utility level in the larger region is the same as in the smaller, because residents of the larger region consume a better variety choice which is just sufficient to compensate them for the lower consumption of housing services.

Importantly, Figure 3 describes the generic structure for $\beta\epsilon < 1$. As transport costs rise point A shifts to the left and point C shifts to the right. In the limit, as transport costs rise to infinity, point A approaches the vertical axis and C approaches the vertical line through $f = 1$. Namely, higher transport costs lead to more unequal regions and in the limit everyone lives in the same region. My simulations show that the inequality in the size of regions rises rapidly with transport costs. As a result, for even moderate levels of transport costs the smaller region is very small relative to the large one.

Figure 4 describes the fraction of people that live in region 1 in stable equilibria, when both regions have the same stock of housing. When transport costs are low half of the population lives in each region. For the range $t \geq \underline{t}$ one region is more

12

populated than the other, and the relative size of the larger region is larger the higher the transport costs. In the limit, as transport costs approach infinity, the entire population lives in one region.

We have seen that as simple as this model of regional formations might be, it produces a rich set of equilibrium outcomes and it helps to shed light on the roles of a small number of critical parameters in shaping geographical structures. In the next section I turn to examine the extent to which these structures are efficient, and in particular whether it is possible to improve the resulting degree of agglomeration.

4 Welfare

One of two types of equilibria emerges in our two-region economy, as depicted by Figures 1 and 3 (Figure 2 is a limiting case of 3). Are these equilibria efficient relative to the existing market structure? Namely, taking as given the fact that housing services are competitively supplied while the market for differentiated products is characterized by monopolistic competition, is it possible to reallocate people across regions and have the winners compensate the losers in a way that raises the welfare level of all individuals? We will shortly see that some of these equilibria are efficient in this sense while others are not. For the purpose of this analysis, however, we first need to compute the utility level of a representative individual in each region for a given allocation of the population across them.

Substituting the pricing equation (3), the number of brands (5) and the expenditure levels (7) into the utility specification (9), we obtain the utility level in each region as a function of the fraction of people that live in region 1, f, and the relative price q:

$$u_1 = A \left(\frac{H_1}{f} \right)^\beta \left(\frac{(1-\beta)q + \beta(fq + 1 - f)}{[fq^{1-\epsilon} + (1-f)t^{1-\epsilon}]^{\frac{1}{1-\epsilon}}} \right)^{1-\beta} N^{\frac{1-\beta\epsilon}{\epsilon-1}}, \tag{11}$$

13

$$u_2 = A \left(\frac{H_2}{1-f} \right)^{\beta} \left(\frac{1 - \beta + \beta(fq + 1 - f)}{\left[f(tq)^{1-\epsilon} + 1 - f \right]^{\frac{1}{1-\epsilon}}} \right)^{1-\beta} N^{\frac{1-\beta\epsilon}{\epsilon-1}}, \tag{12}$$

where $A = \alpha^{1-\beta} \left(\frac{1-\alpha}{a} \right)^{(1-\beta)/(\epsilon-1)}$ and $N = N_1 + N_2$ represents the size of the entire population. By solving from (8) the relative price q for each allocation f and substituting the result into these equations, we can trace out the utility levels in each region as functions of the fraction of people that live in region 1. These curves shed light on the efficiency issue, as I will shortly describe.

Before we proceed to this discussion, however, observe that the welfare level in each region is proportional to $N^{(1-\beta\epsilon)/(\epsilon-1)}$, while the solution to the relative price does not depend on the size of the population. It follows that in this economy welfare rises with population size if and only if $\beta\epsilon < 1$.[5] Namely, when the demand for housing is low and brands of the differentiated product poorly substitute for reach other, the economy is better off with a larger population. A low elasticity of substitution represents a situation in which individuals highly value variety. Under these circumstances they prefer a larger population because a larger population raises variety choice. True, a larger population raises housing prices as well, but given that these individuals place little weight on housing and much on variety choice, the eventual decline in the consumption of housing services as a result of the population's expansion is more than compensated for by the larger variety choice. When $\beta\epsilon > 1$, however, housing is important enough relative to variety to cause general distress when the population expands.

First consider the case in which transport costs are low or people have a strong preference for housing and a weak preference for variety (i.e., $\beta\epsilon > 1$). Also suppose that the regions have equal housing stocks. Then there exists a unique stable equilibrium in which half of the population lives in each region, as described in Figure 1. Moreover, under these circumstances the utility in each region declines with its

---

[5]Recall that this same condition determined whether prohibitive transport costs lead to an equilibrium with a single occupied city.

relative size, as depicted in Figure 5. Therefore a reallocation of people away from the equilibrium point raises the utility level of those that end up living in the smaller region and reduces the utility level of those that end up living in the larger region. Evidently, such reallocations are not Pareto improving. Next suppose that we tax individuals in the smaller region and subsidize individuals in the larger region, using lump-sum taxes and subsidies to equalize their well being. At what value of f will this common utility level be highest? Simulations show that it is highest at the equilibrium point, and it declines the more unequal regions are in size. The reason is that since in the absence of a tax-transfer scheme the fewer individuals that live in the smaller region are better off, they need to be highly taxed in order to compensate the many individuals that live in the larger region in order to equalize utility levels. This tax turns out to be too high to make all of them better off. We conclude that in this case the equilibrium allocation is efficient.

Next consider the case of intermediate transport costs and $\beta\epsilon < 1$, which yields a relative utility curve that is depicted in Figure 3. The resulting utility levels as functions of the fraction of people that live in region 1 are depicted in Figure 6 (the housing stocks are again taken to be the same in both regions). Point A in this figure corresponds to the equilibrium point A in Figure 3 and similarly for points B and C. As argued before, only A and C are stable. Evidently, if the economy is at a stable equilibrium and people are reallocated from the larger to the smaller region, then the well being of everyone rises. This shows that the stable equilibria are not efficient. Since a skewer distribution of region size raises everyone's welfare, we conclude that *market forces do not generate enough agglomeration.*

Now refer back to Figure 4, which describes stable equilibrium allocations for different levels of transport costs. Our welfare analysis has shown that whenever transport costs are relatively low; i.e., $t \leq \underline{t}$, then regions are of equal size and this is an efficient degree of agglomeration. But whenever transport costs are higher than $\underline{t}$ and finite, it is efficient to increase the population in the larger region and reduce it in

the smaller one. Therefore in this range the market provides too little agglomeration.

5 External Economies

Regional structures are shaped by many forces. Urban economists have emphasized a variety of centrifugal forces (in addition to housing) that generate congestion, such as intra-regional transport costs (e.g. Muth (1961) and Alonso (1964)), as well as a variety of centripetal forces that generate regional economies of scale, such as the supply of local public goods (e.g. Stiglitz (1977)). My model has focused on the tradeoff between two such forces: the centrifugal force of available housing and the centripetal force of available variety. Do our results depend in a critical way on the particular variables that were chosen to represent these forces? The answer is 'no'. Each one of them can be replaced with other forces that perform similar functions in the shaping of regions. In order to demonstrate these possibilities I develop in this section an alternative model, in which a homogeneous product that is produced with external economies of scale replaces the differentiated products. The equilibria of this model are similar to the equilibria of the original model, as demonstrated below. Both models share an important feature; they have economies of scale that are not internalized. In this section the lack of internalization will be self evident. In the previous sections it resulted from the fact that available variety was like a public good, and individuals did not take account of their migration decisions on the supply of variety.

Preferences are given by (1), where as before h represents consumption of housing services. The supply of housing is fixed in every region. This time, however, d stands for the consumption of a homogeneous product that is manufactured with labor and supplied competitively. A business firm that manufactures the homogeneous product in region i has a production function $d = A_i l$, where l is the firm's employment of labor. Therefore the marginal product of labor equals A_i in region i and the regional

16

wage rate w_i equals pA_i, where p represents the price of the product. This price is the same in both locations because there are no transport costs.

Now let there be regional external economies of scale, as a result of which $A_i = D_i^{\delta/(1+\delta)}$, $\delta > 0$, where D_i represents aggregate output of the homogeneous product in region i. Then

$$D_i = N_i^{1+\delta}$$

and

$$w_i = pN_i^{\delta} . \tag{13}$$

I maintain the original assumption about the ownership of housing (i.e., the housing stock in each region is equally owned by all individuals). As a result, regional expenditure levels are given by (7). Substituting (13) into (7) we can now compute the relative utility level $v = u_1/u_2$:

$$v = \left(\frac{H_1}{H_2}\frac{1-f}{f}\right)^{\beta} \left(\frac{(1-\beta)f^{\delta} + \beta\left[f^{1+\delta} + (1-f)^{1+\delta}\right]}{(1-\beta)(1-f)^{\delta} + \beta\left[f^{1+\delta} + (1-f)^{1+\delta}\right]}\right)^{1-\beta} . \tag{14}$$

This equation describes a direct relationship between the fraction of people that reside in region 1, f, and the relative utility of the residents in that region.

Two important properties emerge immediately from (14):

(a) When the external economies are small (i.e., δ is close to zero) or people spend a large share of their budget on housing (i.e., β is close to one), the relative utility of a region declines with its relative population size. This case is similar to the case depicted in Figure 1. As a result there exists a unique stable equilibrium in which both regions are occupied. And the population is equally divided between the regions whenever the housing stocks are of equal size.

(b) When the external economies are large (i.e., δ large) or people spend a small share of their budget on housing (i.e., β close to zero), the relative utility of a region rises with its relative population size. This case is similar to the case

17

depicted in Figure 2. In this event there exist two stable equilibria, with the entire population living in one region in each one of them.

Similar equilibria were identified in the original model. In fact, these equilibria are driven by similar forces. A strong preference for housing leads to an equilibrium in which housing is utilized up to its capacity and people allocate across regions according to the availability of housing. This is true in both models. On the other hand, large localized external economies lead to a concentration of people in one location. This ensures the highest real wage rate. In the model with external economies this feature is self evident and δ represents the local degree of economies of scale. In the original model, however, the elasticity of substitution across brands drives the local economies of scale; a lower elasticity of substitution generates stronger economies of scale in the supply of *variety*. So here lies the analogy between these alternative specifications.

What happens for intermediate values of preferences for housing and external economies? There do exist values (such as $\delta = 0.5$ and $\beta = 0.25$) for which the plot of (14) looks similar to Figure 3. In this event in a stable equilibrium regions differ in size even when they have the same supply of housing. Moreover, in such cases there is too little agglomeration, because the regional utility levels as functions of f are similar to Figure 6. Namely, it is efficient to concentrate the entire population in one region. Unfortunately, in this allocation people have an incentive to migrate to the empty region. Therefore complete agglomeration is not an equilibrium outcome.

This completes our discussion of the analogy between the models.

6 A Comparison with Krugman

Krugman (1991) was the first to construct a model with differentiated products and transport costs that can address the tradeoff between a centrifugal and a centripetal force in the shaping of regional structures. My assumptions about the differentiated

product have been borrowed from his work. He also assumed that the other commodity is homogeneous, but unlike my model of section 2 he assumed instead that the homogenous product is freely traded across regions. In addition he assumed that the ownership of the homogeneous product is concentrated in the hands of people who do not supply labor and that these people do not move across regions so that their consumption takes place in the region in which they own the homogeneous product. Clearly, these assumptions are not suitable to describe housing and he labeled the homogenous product 'agriculture'. There appear therefore to be two main differences between Krugman's model and mine:

(a) He assumed that the homogeneous product is freely traded across regions while I have assumed that it is not.

(b) He assumed that the income derived from the homogenous product that is located in region i is entirely spent in region i while I have assumed that this income generates demand in each region in proportion to the number of people that reside in it.

These differences in assumptions produce differences in results. Krugman has shown that whenever transport costs of differentiated products are small enough the relative utility of residents in region 1 rises with the fraction of people that live in this region, as depicted in Figure 2. In this case there exist two stable equilibria, with all the population living in one region in each one of them. It can be shown that whenever $\beta\epsilon < 1$ the relative utility level is also rising for large transport costs. On the other hand, whenever $\beta\epsilon > 1$ and transport costs are large enough the relative utility curve declines, as depicted in Figure 1. As a result there exists a unique stable equilibrium in which both regions are occupied. When the homogeneous product is

in equal supply in both regions half of the population lives in each one of them.[6] These relationships between the structure of equilibria and the level of transport costs are just the opposite from the results in section 3 above. Namely, while in Krugman's model low transport costs lead to agglomeration and high transport costs lead to dispersion, in my model low transport costs lead to dispersion and high transport costs lead to agglomeration. Given that there exist two differences between the models, one wonders which one of them drives the differences in results.

It can be shown that the difference in assumptions about the regional distribution of purchasing power that derives from the ownership of the homogeneous product do not cause the above described differences in results. By this I mean to say that if instead of equal ownership of the housing stocks by mobile workers we were to assume in the model of section 2 that there exist three groups of people: (i) worker that are mobile across regions; (ii) owners of housing in region 1 that live and consume in region 1; and (iii) owners of housing in region 2 that live and consume in region 2; then the same type of equilibria that exist in the model of section 2 would emerge in

---

[6]Using our notation, the equilibrium conditions in Krugman's model can be represented by the following three equations:

$$\frac{1}{1-\beta} = \frac{q^{1-\epsilon}}{fq^{1-\epsilon}+(1-f)\,t^{1-\epsilon}}\left(f + \frac{q_{H2}H_1}{\alpha qN}\right) + \frac{(tq)^{1-\epsilon}}{f\,(tq)^{1-\epsilon}+1-f}\left(\frac{1-f}{q} + \frac{q_{H2}H_2}{\alpha qN}\right),$$

$$\frac{1}{1-\beta} = \frac{t^{1-\epsilon}}{fq^{1-\epsilon}+(1-f)\,t^{1-\epsilon}}\left(qf + \frac{q_{H2}H_1}{\alpha N}\right) + \frac{1}{f\,(tq)^{1-\epsilon}+1-f}\left(1-f + \frac{q_{H2}H_2}{\alpha N}\right),$$

$$v = q\left[\frac{fq^{1-\epsilon}+(1-f)t^{1-\epsilon}}{f\,(tq)^{1-\epsilon}+1-f}\right]^{\frac{1-\beta}{\epsilon-1}},$$

where $\alpha = (\epsilon - 1)/\epsilon$ and $q_{H2} = p_H/p_2$. The price of H equals p_H in both locations due to the free tradeability of this good. The first two equations describe market clearing in each region while the third describes the relative utility of the mobile workers. Recall that the owners of H live and consume in the region in which they own the good. The first two equations provide solutions for q and q_{H2} as functions of the fraction of mobile workers that live in region 1, f. Substituting the resulting solution for the relative price q into the third equation we obtain the relative utility v as a function of f.

the modified model.[7]

I conclude that the fact that Krugman assumed that the homogenous product is freely traded across regions while I assume that it is not is responsible for the differences in results. Moreover, under the assumption of a freely traded homogenous product his additional assumption that some demand is tied to each location is essential for the presence of a centrifugal force. Namely, while in my model dispersion is driven by region-specific supplies (of nontraded housing services) in his model dispersion is driven by region-specific demands (of owners of the homogeneous product). Can Krugman's model produce an intermediate case in which regions with an equal supply of the homogeneous product differ in size?[8] Figure 7 presents a plot of the relative utility in his model for intermediate values of transport costs and $\beta\epsilon > 1$. In this example the amount of H is the same in both regions. As we see, there are three stable equilibria; either both regions are of equal size or all the population lives in one region. There do exist two additional equilibria at points A and C, in which regions differ in size. But those equilibria are unstable.

For the case of equal regional endowments of the homogeneous product my simulations show that in Krugman's model there exist two equilibria for all values of transport costs, with the entire population living in one region in each, whenever

---

[7]It can be shown that with the proposed modification of the structure of demand the equilibrium condition (8) for the relative price q would be replaced with

$$1 = \frac{fq^{1-\epsilon}}{fq^{1-\epsilon} + (1-f)t^{1-\epsilon}} + \frac{(1-f)t^{1-\epsilon}q^{-\epsilon}}{f(tq)^{1-\epsilon} + 1 - f}$$

and the equation for the relative utility level (10) would be replaced with

$$v = \left(\frac{H_1}{H_2}\frac{1-f}{f}\right)^{\beta} \left[\frac{fq^{1-\epsilon} + (1-f)t^{1-\epsilon}}{f(tq)^{1-\epsilon} + 1 - f}\right]^{\frac{1-\beta}{\epsilon-1}} q^{1-\beta}.$$

It follows that v declines with f for small levels of transport costs. It also declines with f for large transport costs whenever $\beta\epsilon > 1$. Finally, simulations show that for intermediate values of transport costs and $\beta\epsilon < 1$ the relative utility curve takes on the shape that is depicted in Figure 3.

[8]Krugman (1991) raises this possibility in footnote 3 but does not pursue it.

$\beta\epsilon < 1$. On the other hand, for $\beta\epsilon > 1$ the relationship between equilibrium regional structures and transport costs is as depicted in Figure 8. Namely, for low transport costs $t < \underline{t}_K$ there exist two stable equilibria with extreme agglomerations. For high transport costs $t > \bar{t}_K$ there exists a single stable equilibrium with an equal size of regions. Finally, for intermediate values of transport costs there exist three stable equilibria: two with extreme agglomerations and the third with an equal size of regions.

To close this section, let us examine the intuition behind Krugman's results in order to see why they differ from mine. For this purpose it suffices to focus on the extreme cases. When transport costs are nil every allocation of mobile works across regions constitutes an equilibrium in his model. Because under these circumstances every product can be bought for the same price at each location and wages are the same everywhere. As a result workers are indifferent as to where they live no matter how many of them reside in each region (i.e., the v curve is flat at 1). Now suppose that the same amount of the homogenous product is supplied in each location and half of the workers live in each region, while transport costs are small but positive. In this case too workers have no incentive to switch locations. But suppose that nevertheless some of them move from region 1 to 2. Then variety choice is larger in region 2 and more workers will be attracted to that region. Manufacturers of the differentiated product will shift production to region 2 in order to serve a larger market. With sufficiently low transport costs the demand for differentiated products by the owners of the homogenous product in region 1 is not sufficient to raise wages of workers in region 1 so as to compensate them for the larger variety choice available in region 2. As a result the cumulative process of migration will continue until all workers reside in region 2.

We have thus seen why low transport costs cause the relative utility of mobile workers in a region to rise with the fraction workers that live in it. What happens when transport costs are large, and for concreteness suppose that they are prohibitive?

22

Let us start again from an allocation with equally populated regions and equal well being levels of workers in both locations. Now suppose that some workers move from region 1 to 2. Then again there is more variety in region 2 which makes it attractive to live there. Except that his time manufacturers of differentiated products in region 2 cannot serve the owners of the homogeneous product in region 1. As a result they have an incentive to locate some production in region 1. Is the demand by owners of the homogeneous product in region 1 sufficiently high to induce production of differentiated products in that region at sufficiently high wages that will keep production workers in region 1? To keep workers in region 1 the wage rate has to be high enough to compensate them for the lower variety choice in region 1. It follows that the answer depends on two things: first, how large is the income level of the owners of the homogenous product; and second, how important is variety choice. If people spend a small fraction of their income on the homogeneous product the income level of its owners in region 1 will be small and they will not generate enough demand to compensate workers for a lower variety choice. If in addition the elasticity of substitution across products is small, workers require a high compensation in wages in order to forgo a better variety choice in region 2. Therefore whenever $\beta\epsilon < 1$ the utility of workers is higher in the region that has more of them (the v curve slops upwards), which leads to a cumulative process of migration from the smaller to the larger region. On the other hand, whenever the demand for the homogeneous product is high, the demand level by its owners in region 1 is large. And if in addition the elasticity of substitution across brands is large, workers request little compensation for loss of variety choice. In this event – which happens when $\beta\epsilon > 1$ – workers are better off in the smaller region (the v curve slops downward). As a result, workers migrate to the smaller region until a symmetric equilibrium is attained.

7 Concluding Comments

By providing a new framework for the analysis of regional structures, Krugman's work on economic geography has generated renewed interest in regional economics. His approach emphasizes the centripetal force of a manufacturing sector that supplies costly transported differentiated products. And it builds on a centrifugal force that consists of regional demand levels that are driven by immobile individuals. The latter is particularly suitable for societies in which agriculture plays a major role, and in which farmers are tied to their land.

I have examined in this paper an alternative framework, in which agriculture is replaced with housing. A prominent feature of housing services is that they are traded within a region but not across regions. As a result, locally supplied housing services produce the major centripetal force. This framework seems to be closer to standard urban economic models.

Importantly, while in Krugman's framework low transport costs lead to extreme agglomerations and higher transport costs may reduce the degree of agglomeration, in my framework low transport costs lead to little agglomeration while higher transport costs may lead to larger agglomerations. Since each one of these alternative frameworks is suitable in different circumstances, it is important to understand the economic forces that drive their results. I have tried to clarify them. Finally, I have shown that in economies of this type, if anything, free markets provide too little agglomeration.

These findings result from a sharp focus on a small number of features, which limits their role in explaining urban and regional structures. It is therefore necessary to broaden the scope of such investigations in order to obtain a better understanding of these problems. Standard elements of regional and urban economics, such as spacial layouts, commuting costs, and congestion, are natural candidates to be incorporated in the analysis.

References

[1] Alonso, William, *Location and Land Use* (Cambridge, MA: Harvard University Press), 1964.

[2] Krugman, Paul, "Increasing Returns and Economic Geography," *Journal of Political Economy*, Vol. 99, no. 3, pp. 483-499.

[3] Muth, Richard F., "The Spacial Structure of the Housing Market," *Papers and Proceedings of the Regional Science Association*, 7, pp. 207-220.

[4] Stiglitz, Joseph, "The Theory of Local Public Goods," in Feldstein, M.S. and Inman, R.P. (eds.), *The Economics of Public Services* (London: Macmillan Press), 1977.

[5] Tolley, George and Crihfield, John, "City Size and Place as Policy Issues," in Mills, E.S. (ed.), *Handbook of Regional and Urban Economics*, Vol. II (Amsterdam: Elsevier Science Publishers B.V.), 1987.

Figure 1
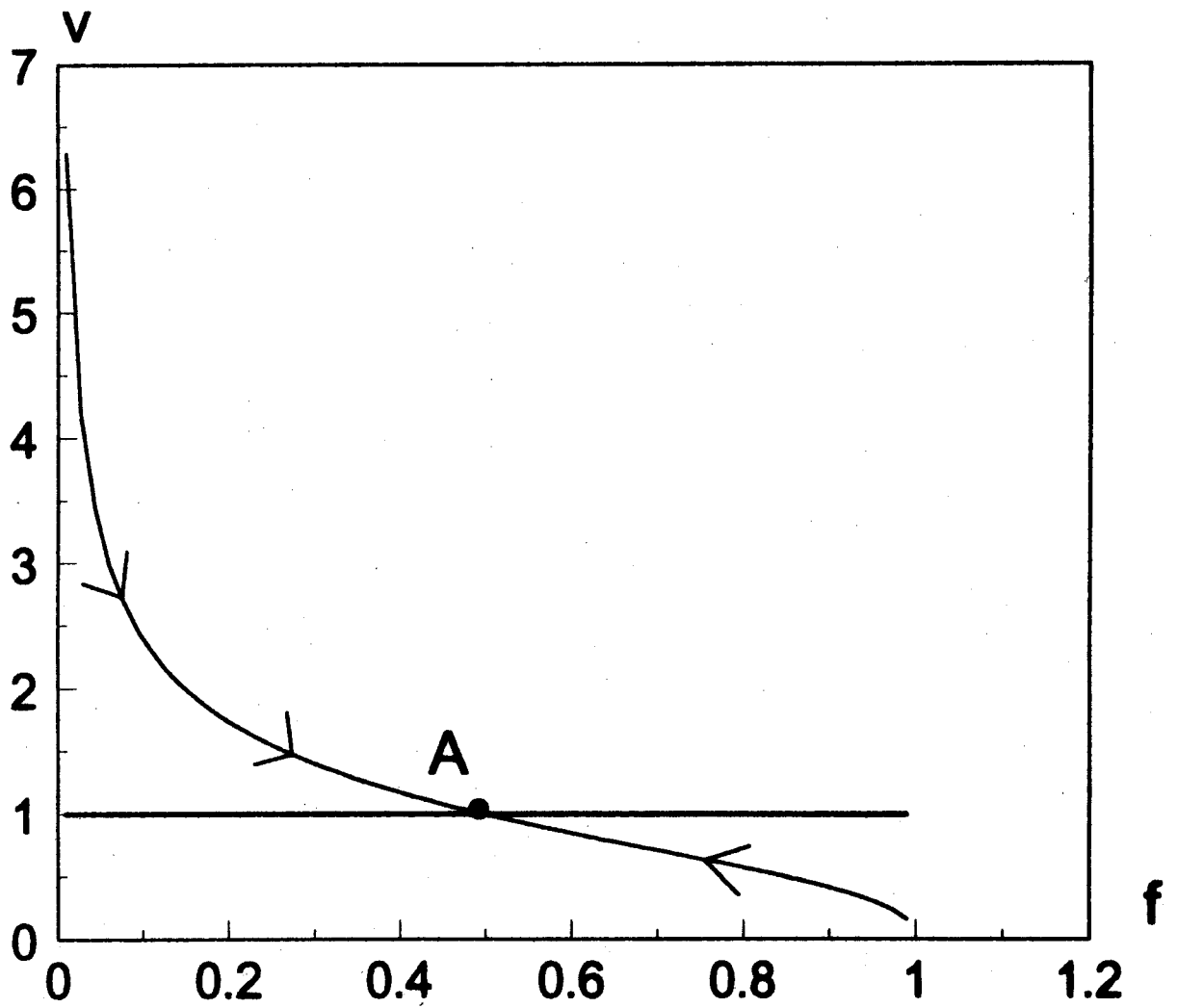
t close to 1 or βε > 1
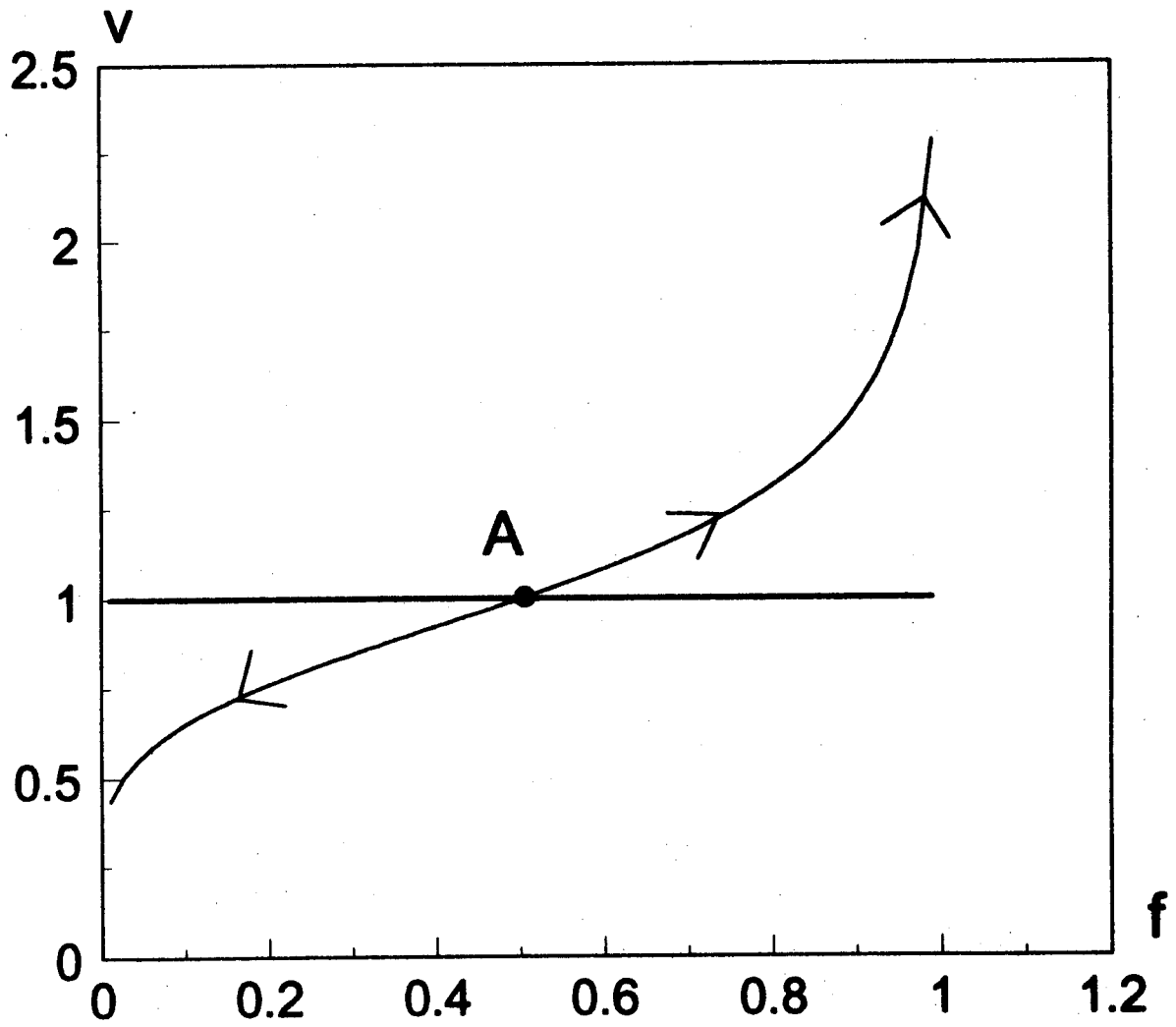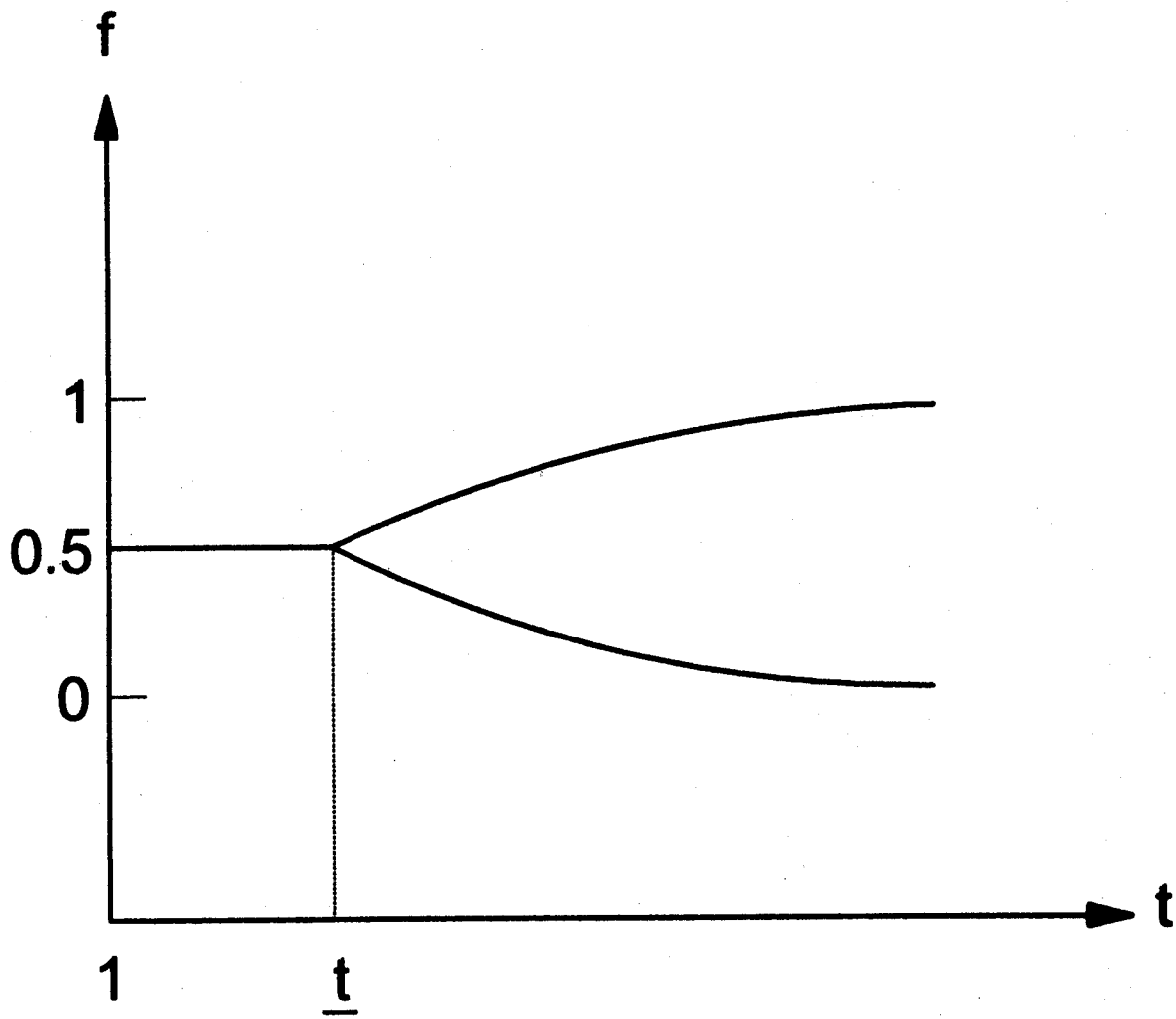
Figure 2

t unbounded and $\beta\varepsilon < 1$

Figure 3
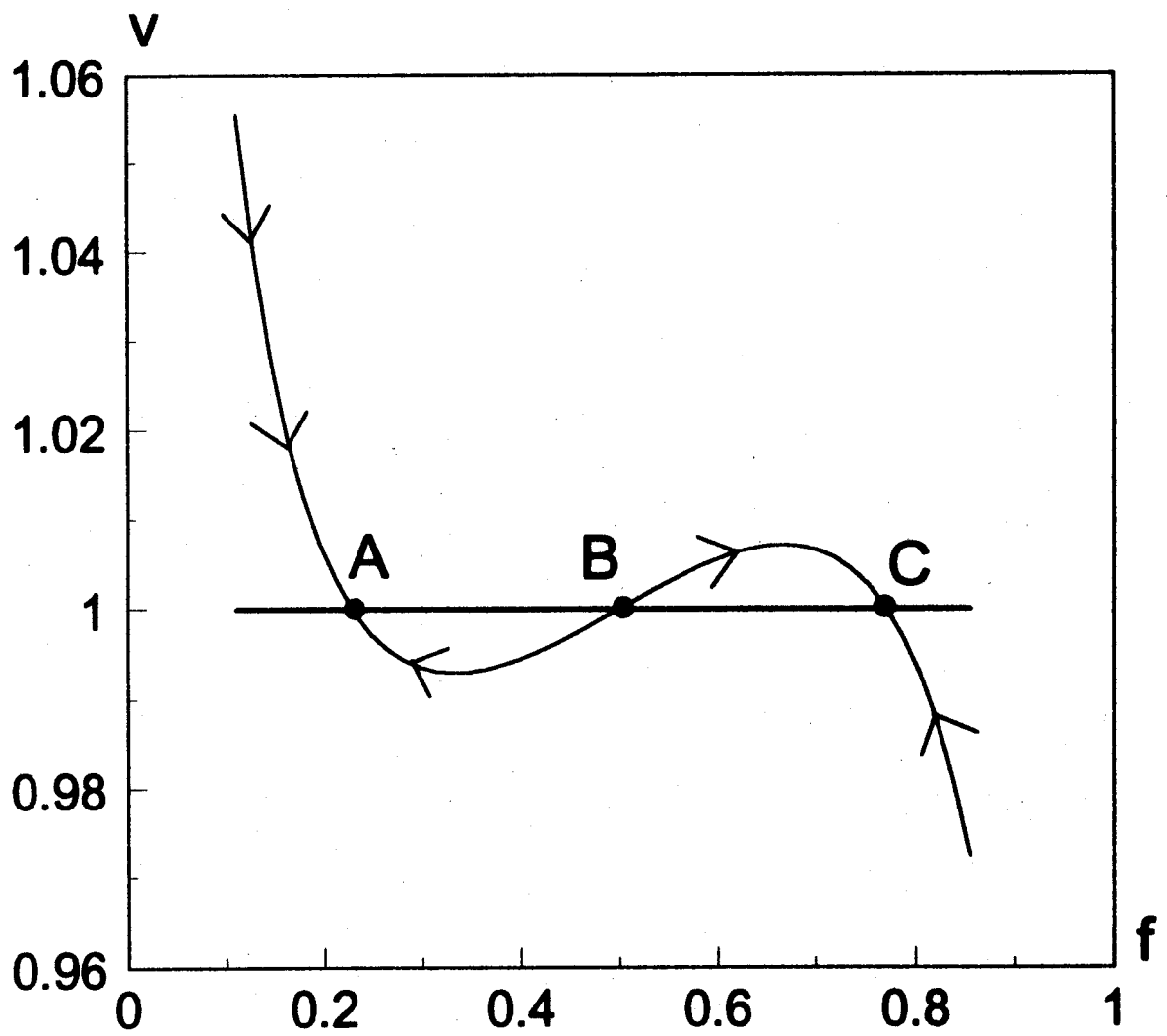
$\varepsilon=2$, $\beta=0.4$, $\beta\varepsilon < 1$

$t=6$

Figure 4
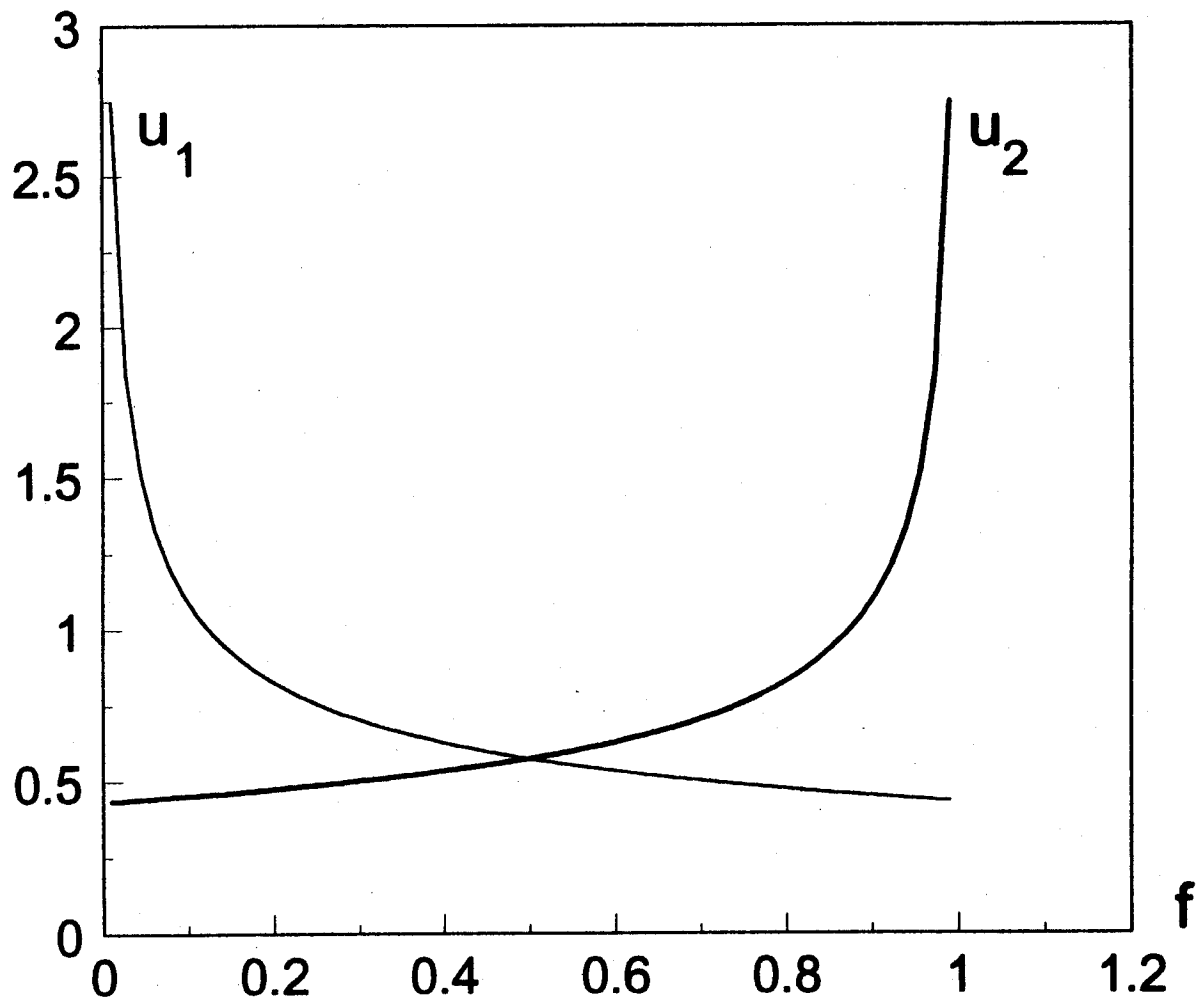

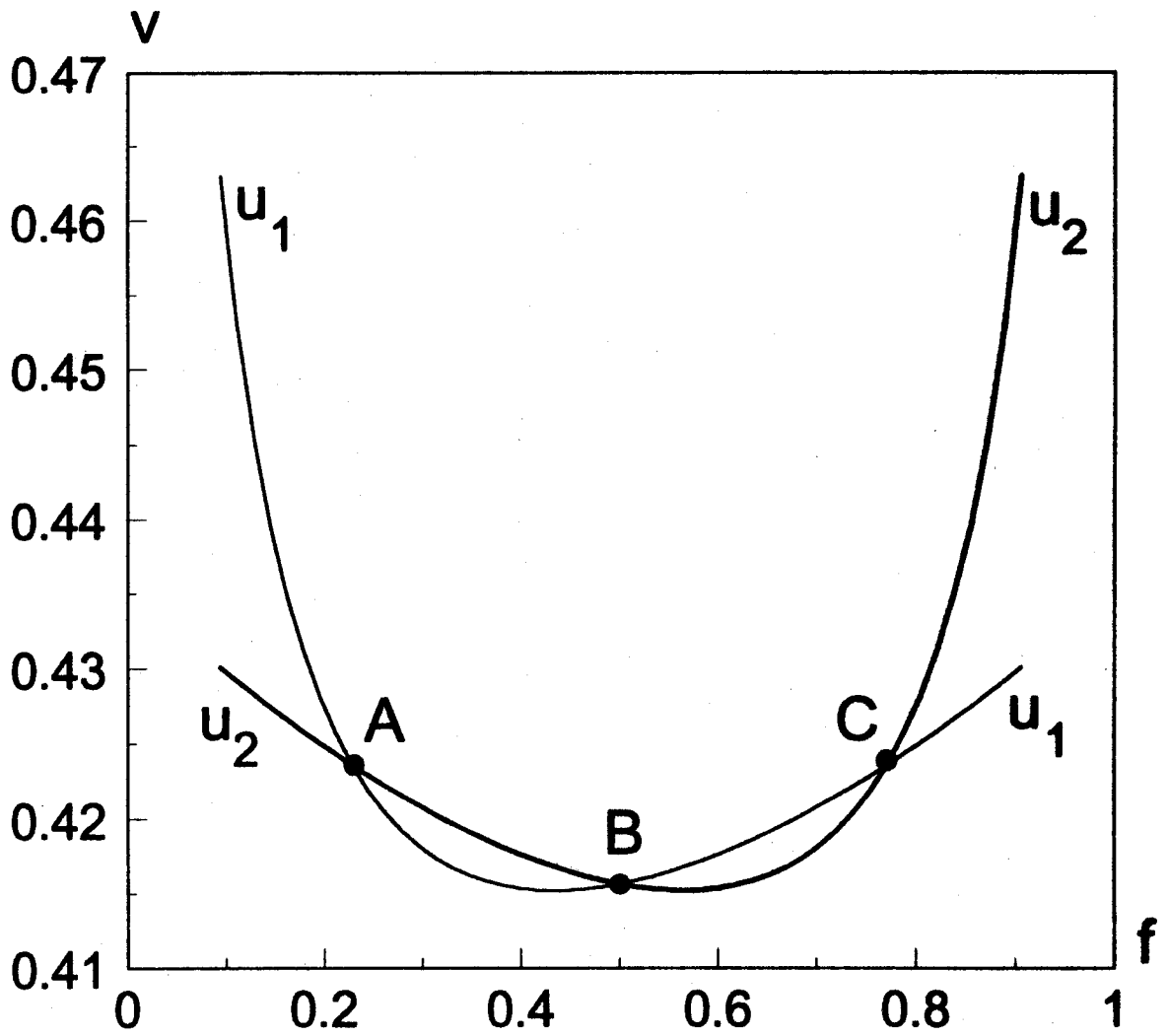$\beta\varepsilon < 1$

Figure 5

t close to 1 or $\beta\varepsilon > 1$

Figure 6

$\varepsilon = 2$, $\beta = 0.4$, $\beta\varepsilon < 1$

$t = 6$

Figure 7

$\varepsilon=4$, $\beta=0.7$, $\beta\varepsilon > 1$

t=1.666

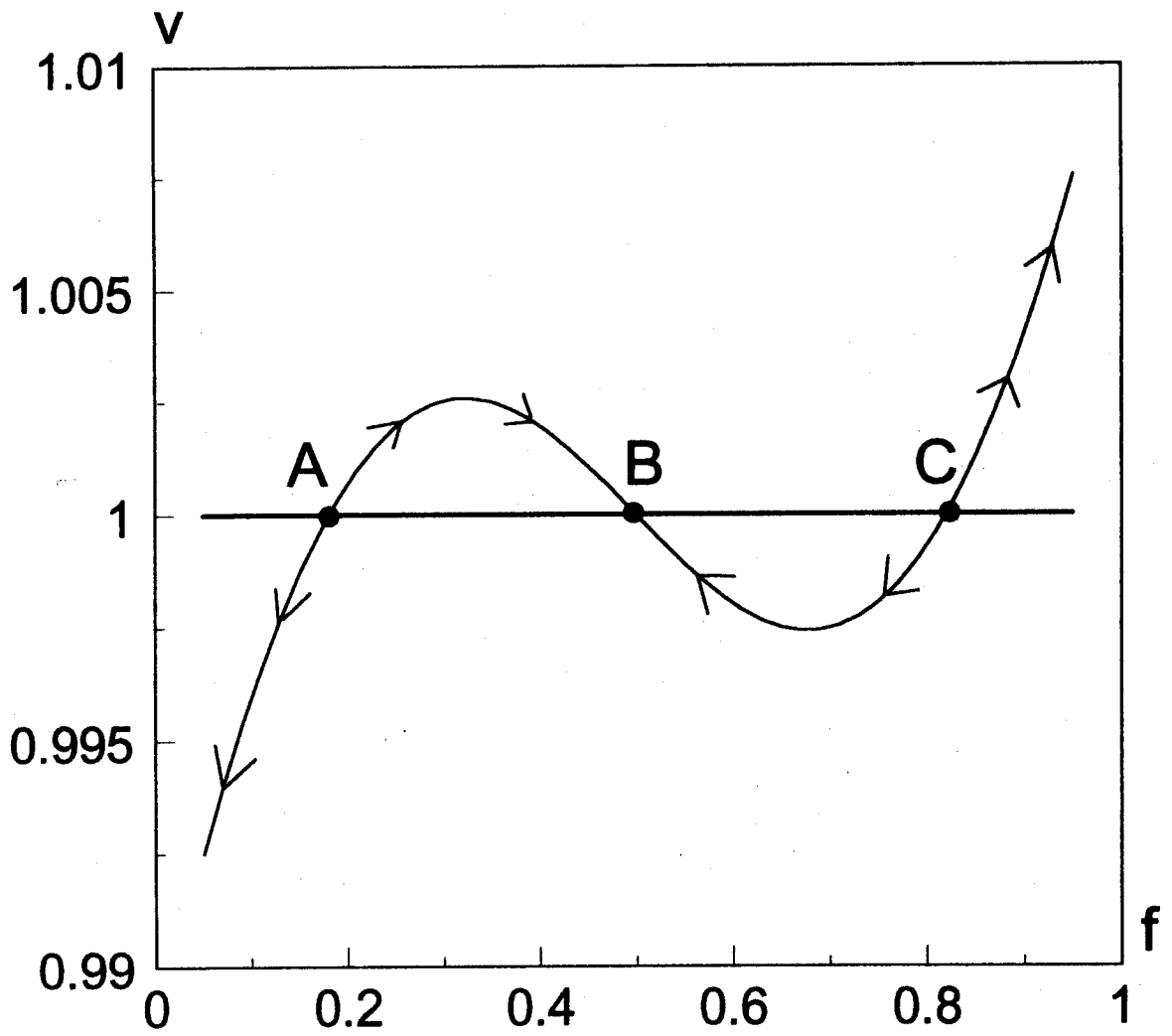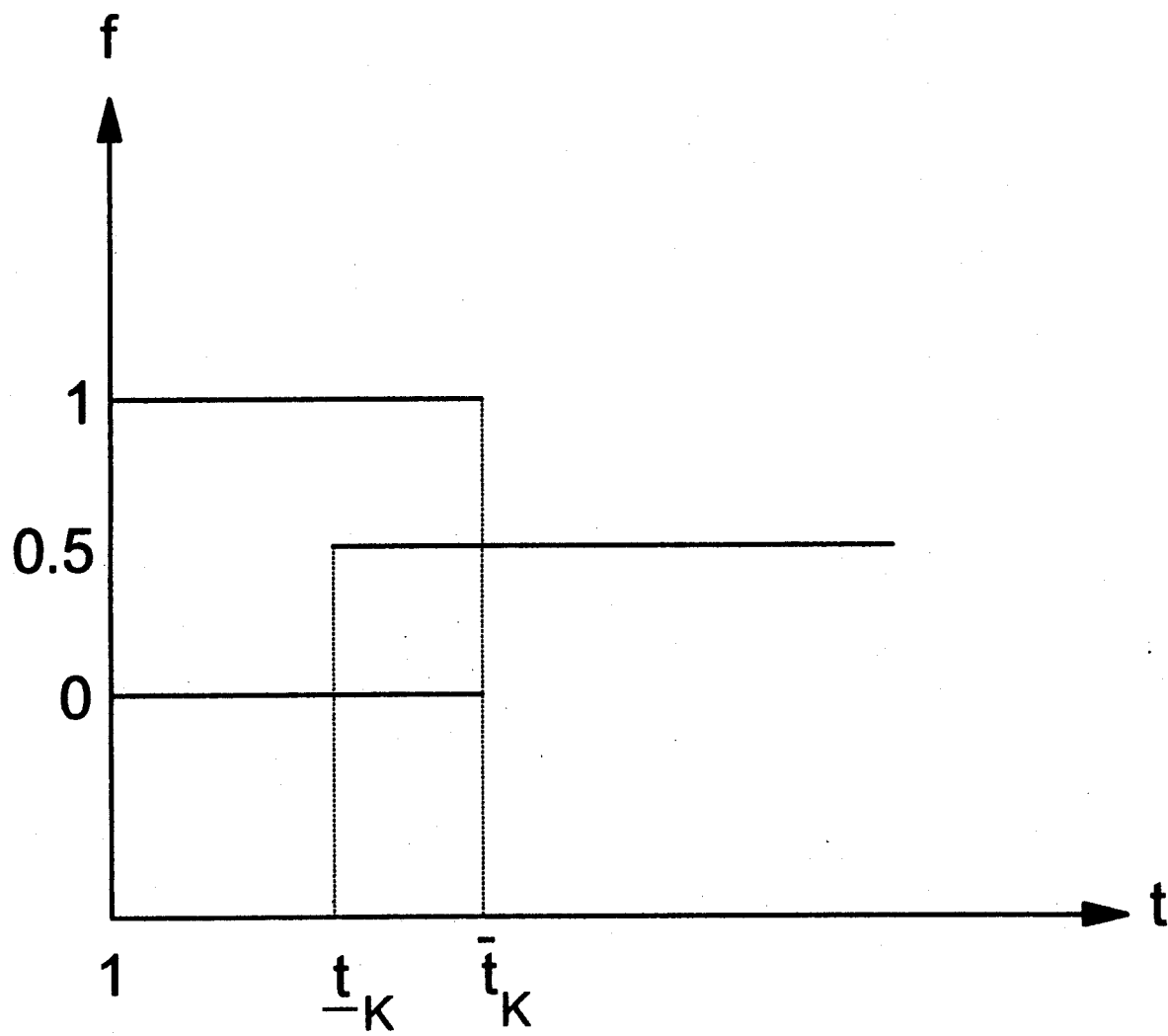Figure 8

$\beta\varepsilon > 1$