## 1. INTRODUCTION

The concept of agglomeration economies, first considered in a systematic (though rather restrictive) manner by Weber (1909), has proven to be an important feature in the analysis of industrial location, whether this is of a theoretical or empirical nature. In both cases considerable reliance has been placed on the three categories of agglomeration economy proposed by Ohlin (1933, p.40). These were described by Hoover (1937, pp.90-91) as scale economies, localisation economies and urbanisation economies, and were later considered in some detail by Isard (1956). Leaving aside the concerns of McCann (1995), who argued that attention should be focused on the cost issues underlying agglomeration economies, such a tripartite classification is incomplete in several respects, and therefore represents at best a partial summary. It has recently been argued that the agglomeration economies enjoyed by a firm can be divided into those based on internal economies and those based on external economies, and also that each kind of economy can be interpreted in terms of scale, scope or complexity (Parr, 2000). A classification organised around these distinctions subsumes the Ohlin classification, and also permits a sharpening of his categories. The concern here is to explore certain implications of this classification, and to examine a number of issues that have probably not received the attention which these deserve. It will be argued that such issues need to be addressed if the concept of agglomeration is to be employed effectively in the analysis of industrial location. Prior to this, however, a short discussion of the various types of agglomeration economy is presented.

## 2. TYPES OF AGGLOMERATION ECONOMIES

Agglomeration economies are regarded here as referring to the cost savings to the firm which result from the concentration of production at a given location, either on the part of the individual firm or by firms in general. The classification referred to above is based on the premise that agglomeration economies derive from economies which are internal to the firm, as well as and from externalities. Consideration is first given to internal economies and to the agglomeration economies based on these.

*Internal Economies*

We start with economies of horizontal integration, better known as internal economies of scale. These refer to the fact that the unit cost of a production (beyond some minimum scale) is a decreasing function of output. By contrast, economies of lateral integration (or internal economies of scope) refer to the diversity of production. Such economies occur when the firm's joint output of two or more products results in a

1

lower total cost than would be the case if the products were produced by separate single-product firms (Panzar and Willig, 1981). Economies of scope are usually held to result from the use of an input which would be underemployed with single-product output, but which could be used more efficiently if shared in the output of several products. In the case of economies of vertical integration (or internal economies of complexity) the cost savings to the firm are related to its engaging in the various stages or processes of production, rather than simply producing an end product. Such a structure of production results in a lower total cost for the end product than would be possible if the different stages were undertaken by separate specialist firms. Economies of complexity include such advantages as improved managerial oversight and superior levels of co-ordination.

None of these internal economies to the firm necessarily involves the concentration of economic activity. When, however, the economies underlying these various forms of integration are spatially constrained (i.e. require the co-location of the relevant activity or activities), we have the bases for various types of internally-based agglomeration economy. This is most obvious in the case of horizontal integration when the economies of scale involve production, e.g. a plate-glass plant. However, economies of lateral integration, involving the scope dimension, may similarly be spatially constrained, e.g. a firm operating a plant or commonly-located plants, producing electricity generators and pumping machinery, with specialised design expertise representing the shared (but indivisible and immobile) input. Spatial constraints may also be present with economies of vertical integration, involving the dimension of complexity, e.g. an iron and steel works or a petrochemical plant. Important here would be energy savings and the possibilities for enhanced quality control. These various spatially-constrained internal economies are summarised on the left side of Table 1. It is again emphasised that within a given firm the existence of one of these three types of internal economy does not automatically give rise to an agglomeration economy.

Table 1.        Underlying Bases for Agglomeration Economies to the Firm

| Dimension | Spatially-constrained internal economies | Spatially-constrained external economies |
|---|---|---|
| scale | economies of horizontal integration | localisation economies |
| scope | economies of lateral integration | urbanisation economies |
| complexity | economies of vertical integration | activity-complex economies |

*External Economies*

There exist other types of agglomeration economy, and these stem from economies which are external to the firm. Such externalities are beyond the control of the individual firm and typically result from the existence and/or collective action of other firms. Again, consideration is given to the dimensions of scale, scope and complexity. External economies of scale refer to those cost savings to the individual firm that depend on the scale of the industry to which the firm belongs. These economies are based on the existence of specialised servicing activities, the possibilities for co-operative research and development activity, and the advantage of industry-wide marketing. By contrast, external economies of scope refer to cost savings to the individual firm, which are dependent on the existence of firms in other industries. Here the economies are based on the shared use of inputs with other firms, either on a bi-lateral or multi-lateral basis (ownership of the shared input is generally in the hands of one firm but the externality affects all firms in the sharing arrangement). A third externality is concerned with external economies of complexity. These result from the fact that a firm is linked in input-output terms to firms in other industries, to form an identifiable production ensemble. The cost savings in this case derive from efficient information flows and the ability to co-ordinate its activities with other firms, particularly with respect to the avoidance of input-supply problems.

Each of these external economies may exist in a variety of spatial settings, including ones in which the relevant activities have a dispersed pattern of location, and examples are not difficult to identify. On the other hand, it is sometimes the case that certain of these external economies require a spatial concentration of the relevant activities. Robinson (1956) was emphasising this point when he spoke about "immobile external economies" which were specific to a particular location. And it is probably not overstating the point to argue that external economies only have their full expression in a spatial setting. Under these circumstances the various external economies form the bases for three further types of agglomeration economy, which are summarised on the right side of Table 1. We consider first spatially-constrained external economies of

3

scale, or what Marshall (1892) termed "localisation economies", these involving the concentration of like firms, i.e. firms in the same industry. Such a localisation of firms permits the emergence of pools of skilled labour, lower freight rates on inputs as well as outputs, access to specialist services, and the possibility of information spillovers. Localisation economies, which are external to the firm, are internal to the industry, being a function of the scale of the industry at the localisation. Agglomeration economies of this type were historically common in the shoe industry and various branches of the textile industry, and nowadays make their appearance in particular parts of the electronics industry.

In the case of spatially-constrained external economies of scope, more commonly known as "urbanisation economies", the concentration of economic activity involves unlike but unrelated firms. This not only facilitates the sharing of specialised inputs among diverse firms, but (more importantly) also permits these firms to share public utilities, transportation services and other elements of the common infrastructure, as well as specialised business services, all of which are typically provided by a third party (the state or the market). Urbanisation economies are external to the firm though internal to the urban concentration, and are a function of the scope (diversity) of the concentration, this being closely correlated with its size. Such economies continue to be important for firms in a wide range of industries, particularly where the firms are relatively small and/or new, with locations in industrial parks and inner-city districts of metropolitan areas (Goldstein and Goldberg, 1984; Lichtenberg, 1960; Vernon, 1960).

Finally, we have spatially-constrained external economies of complexity, termed here "activity-complex economies". These are based on the concentration of unlike firms which are related to each other in terms of backward and/or forward linkages, where the overall pattern of linkages is sequential (e.g. a garment complex) or convergent (e.g. a shipbuilding or aerospace complex). For the constituent firm within the complex, proximity provides the advantages of transportation-cost savings, efficient flows of materials among stages, and lower inventory costs. There may also be the advantage of specialist firms supplying goods and services required by firms in different parts of the complex, an interaction which Florence (1961) referred to as "diagonal linkage". Activity-complex economies are again external to the firm but in this case internal to the activity complex, and are usually a function of the extent or elaborateness of the complex. Such economies are not generally treated as a separate type, although the advantage of doing so will become evident.

It can be seen that this classification of agglomeration economies contains an important parallelism of structure. For each agglomeration economy based on an internal economy, there is a corresponding agglomeration economy based on an external economy. Whereas the former structure involves the integrated firm at a given location, gaining internal economies of scale, scope or complexity, the latter structure comprises a non-integrated structure of independent firms, with the common location permitting such firms to benefit from the corresponding external economies of scale, scope or complexity. The former structure thus reflects an organisational or management approach to the allocation of resources, while the latter structure depends on a market-based allocation, this being generally dependent on a relatively low level of transactions costs for the relevant firms, both in an *ex ante* and an *ex post* sense. To a large extent the two contrasting structures may be seen as alternatives or even substitutes.

## 3.    THE IDENTIFICATION OF AGGLOMERATION ECONOMIES

From what was argued above, the impression might be gained that there are clear differences among the various types of agglomeration economy. This tends to be true for agglomeration economies which are internal to the firm, although exceptions certainly exist. Taking a broader view of agglomeration economies, however, occasions arise when their identification depends on some extraneous or contextual factor. Three examples serve to illustrate the point. The first relates to industry organisation. We may consider the case of an urban centre which has over the years attracted several firms in the machine-tool industry, these serving the same product market. As a consequence of this the firms concerned may have the advantage of a skilled workforce, and may also enjoy the fruits of co-operative activity with respect to product development and marketing, for example. Such advantages to these like firms would usually be identified as agglomeration economies of the localisation type or spatially-constrained external economies of scale. If, for whatever reason, these firms decide to form a single corporate entity, the erstwhile externalities now become internalised, so that the agglomeration economies are therefore in the nature of economies of horizontal integration or spatially-constrained internal economies of scale. In the short run, at least, little has changed, so that in this instance the identification of agglomeration economies turns largely on the question of ownership.

A second example relates to the question of disaggregation.  We consider the case of the textile industry. This is typically organised into various stages (spinning,

weaving, dyeing, printing, finishing, etc.), which exist in a sequence. If textile firms in an urban centre are specialising in different stages of the industry and are known to be benefiting from the advantage of local commercial/financial services specific to the textile industry, how are the relevant agglomeration economies to be identified? If the industry is considered *en bloc*, we would probably want to regard the advantages of agglomeration as localisation economies or spatially-constrained external economies of scale. On the other hand, the industry may be considered in terms of its constituent parts, in which case the agglomeration economies would be identified differently. If the different parts of the industry were related in a production sequence, the agglomeration economies would be regarded as activity-complex economies or spatially-constrained external economies of complexity. If, however, the various parts of the industry were wholly unrelated (so that there existed, in effect, a set of different industries), the agglomeration economies would be in the nature of urbanisation economies or spatially-constrained external economies of scope. The identification of agglomeration economies in this case clearly depends on the extent of disaggregation.

The third example concerns the sectoral structure of the local economy and its long-run development. Suppose that at a particular point in time the sole export base of an urban centre involves firms in the same branch of the food-processing industry, and that these enjoy the advantage of specialist supply from a number of firms concerned with machinery maintenance and repair. The firms in the food-processing industry could again be said to have the benefit of localisation economies. Suppose further that over the next ten years another industry develops in this urban centre and is also dependent on the same firms in the machinery-maintenance industry. The manner in which the existing agglomeration economies are now identified will depend on the nature of the new industry. If it represents a backward-linkage from the food-processing industry (for example, the chemical industry which produces food additives), then the agglomeration economy relating to machinery maintenance, which was originally regarded as a localisation economy for the food-processing industry, must now be treated as an activity-complex economy for the sequentially-linked chemical/food-processing complex (the activity-complex economy here is based on diagonal linkages from machinery-maintenance firms to both elements of the complex). If, on the other hand, the new industry is unrelated to the food-processing industry and represents, say, the pulp and paper industry, the agglomeration economy for the food-processing industry, involving machinery maintenance, must now be identified as an

urbanisation economy, which is common to both the food-processing and pulp and paper industries. In this case, therefore, the identification of agglomeration economies is a function of the (changing) composition of economic activity within the urban centre.

The difficulties of identification contained in these three examples do not vitiate the definitions of agglomeration economies considered earlier, but rather draw attention to the need for careful consideration of the context in which the various agglomeration economies exist. Such difficulties of identification, which are not to be underestimated, probably caused at least one author to despair of the practical possibility of distinguishing between localisation and urbanisation economies (Evans, 1989, p.41), and may have forced others to use the less specific term "external economies" (Lichtenberg, 1960; Vernon, 1960). This question of identification assumes a critical importance when we come to consider the co-existence of agglomeration economies of different types, either for a given firm or a given urban concentration (Parr, 2000).

## 4    THE EXTENT OF AN AGGLOMERATION ECONOMY FOR THE FIRM

We know that in a general sense a particular agglomeration economy represents an advantage to a firm which results from the concentration of economic activity at a given location. But what is the extent (the value) of this advantage? Furthermore, with respect to what alternative location or locational pattern is this advantage being measured? In the substantial literature on agglomeration economies, questions such as these frequently go unanswered. It is possible to treat separately those agglomeration economies which are based on internal economies and those which are externality based, although it will be obvious by now that a firm may benefit from agglomeration economies of both kinds. In each case the agglomeration economy is expressed as a cost saving per unit of output.

*The Case of Agglomeration Economies Based on Internal Economies*

Agglomeration economies which are based on internal economies are considered initially in gross terms. For an optimally-located firm it is a fairly straightforward matter to determine the cost savings deriving from spatially-constrained internal economies of scale, involving a large-scale plant (we assume as a yardstick  some minimum scale of plant). It is also relatively easy to measure those cost savings related to spatially-constrained internal economies of scope or complexity, involving in either case a single-plant operation or a multi-plant structure at a common location. The extent of an individual agglomeration economy is examined using the notion of a "reference configuration". This is defined as what would represent the profit-maximising set of

7

locations, if we excluded from the analysis those costs upon which the individual agglomeration economy is based. We may reasonably assume that the reference configuration does not correspond to the agglomeration in question, and that the reference configuration is different for each agglomeration economy. We are thus assuming that the various agglomeration economies are wholly independent of one another in the sense that the decision to exploit one does not affect the decision to exploit another. For an individual agglomeration economy the gross savings represent the difference between all costs at the reference configuration (including that set of costs on which the agglomeration economy is based) and these same costs at the optimal location, i.e. the firm's production agglomeration, comprising a single plant or set of plants.

However, an agglomeration economy is usually in the nature of a net saving, i.e. a saving over and above all the relevant expenses incurred in securing it. We should therefore speak of a "net agglomeration economy". Thus the value of an individual net agglomeration economy is the gross savings or cost advantage to the firm resulting from concentration, adjusted downward by any higher transportation (assembly and/or distribution) costs and by any lost revenues arising from this concentration. In that situation where there are actually revenue gains, these augment the extent of the agglomeration economy (although the concern of this paper is with agglomeration economies on the cost side, the issue of revenue cannot be excluded from the calculation). If, in the case of a particular agglomeration economy, its net value was negative, this would imply that in terms of the particular dimension of production (scale, scope or complexity) the firm was not optimally located and that the relevant reference configuration should have been selected. This problem does not arise, of course, since it is assumed that the firm is optimally located at the production agglomeration.

It is possible for a firm, some or all of whose production facilities are concentrated at a particular location, to benefit from more than one internally-based agglomeration economy. For example, a firm may have the advantage of spatially-constrained economies of scale in production, but may also benefit from spatially-constrained economies of complexity, resulting from the multi-stage or multi process nature of the firm's production. The concern is now with "net overall agglomeration economies" enjoyed by the firm at the relevant agglomeration. Since the reference configuration is assumed to be different for each agglomeration economy, so that the individual

agglomeration economies are separable (i.e. exist independently of one another), the value of the net overall agglomeration economies is given by the sum of the individual net agglomeration economies.

*The Case of Externality-Based Agglomeration Economies*

Up to this point the concern has been with those agglomeration economies which are under the control of a firm. In certain cases the firm's decision to exploit one or more agglomeration economies may simply result in the existence of a company town. In other cases, however, the firm (or one of its constituent plants) may wish to take advnatage of agglomeration economies which are based on externalities (for convenience we simply refer to the firm from now on). Attention thus turns to the firm at a particular agglomeration. We assume that the representative firm, as part of the agglomeration of economic activity, contributes only marginally to the existence of an individual agglomeration economy, and in this way avoid the problem of locational interdependence among firms. The question as to whether a firm is better located at an agglomeration rather than elsewhere is a familiar one in location theory. It was dealt with by Weber (1909) and Palander (1935) by means of isodapane analysis, and can readily be considered in terms of the Isard (1956) input-substation framework. These essentially *ex ante* approaches placed considerable emphasis on the point of minimum transportation costs. Here, however, a less restrictive view is taken and the question will be approached in *ex post* terms. Three factors need to be made explicit: first, the agglomeration economy under consideration; second, the particular agglomeration at which the firm is located (assumed to be the optimal location); and third, the "benchmark location". This is defined as the location that would represent the profit-maximising location, if we excluded those costs on which the agglomeration economy depended. Such a quasi-optimal location might represent the point of minimum transportation costs à la Weber or the point of maximum revenue, or could simply be some other location. However, the rule is imposed that the benchmark location cannot coincide with the agglomeration at which the firm is located. In order to avoid such an eventuality, the next-best quasi-optimal location (situated beyond some minimum distance from the agglomeration) is automatically selected as the benchmark location.

Attention is focused first on the gross cost savings, these representing the costs at the benchmark location (all cost elements now included) *minus* the costs of the agglomeration. The gross cost savings associated with a particular agglomeration economy may well be in the nature of a residual: certain elements of a given cost may

be higher at the agglomeration than at the benchmark location, although in overall terms the given cost is lower. Examples include high costs for specialised labor (in the localisation case), substantial location rents and congestion (in the urbanisation case), and co-ordination problems (in the activity-complex case). Weber (1909) regarded such increased costs as belonging to his class of "deglomerative factors". As with agglomeration economies based on internl economies, the extent of an externality-based agglomeration economy must be considered in net terms. This involves taking the gross cost savings, and subtracting from these any additional transportation costs and also any revenue losses, resulting from location at the agglomeration rather than at the benchmark. Transportation-cost savings and/or revenue gains would be added to gross savings, thus enhancing the value of the agglomeration economy (for the sake of convenience, these possibilities are now excluded from the discussion).

This approach outlined above is satisfactory, if the firm is known to benefit from only one type of agglomeration economy, but a firm may enjoy more than one agglomeration economy at a given agglomeration. In contrast to the case of more than one internally-based agglomeration economy, however, the various externality-based agglomeration economies are not independent of one another, since there is, of necessity, only one benchmark location. This represents what would be the profit-maximising location, if we excluded from consideration these cost elements on which the two or more externality-based agglomeration economies depended. Overall gross savings represent the costs at the benchmark location (all costs now considered) *minus* the costs at the agglomeration. Overall net agglomeration economies are determined by taking the gross savings, and subtracting from these any additional transportation costs and lost revenue resulting from a location at the agglomeration rather than at the benchmark location. The value of overall net agglomeration economies will always be non-negative, given the assumption that the firm is optimally located at the agglomeration. Even though the agglomeration economies are being considered collectively, it is possible to provide an imputed value for each individual agglomeration economy. This involves taking the value of gross cost savings attributable to each agglomeration economy (when considered separately), next deriving their individual shares, and then applying these to the value of the net agglomeration economies. As an example, assume that the gross cost savings per unit of output are as follows: for localisation economies, £50; for urbanisation economies, £30; and for activity-complex economies, £20. The respective shares are therefore 50 per cent, 30

per cent and 20 per cent. When these shares are applied to the value of overall net agglomeration (pre-calculated to be £90 per unit of output), the individual net agglomeration economies have the following imputed values: localisation economies, £45; urbanisation economies, £27; activity-complex economies, £18.

An interesting (and by no means unusual) case arises when certain agglomeration economies facing a firm are positive, while others are negative, with net overall agglomeration economies being positive. For example, localisation economies and activity-complex economies may be positive, while urbanisation economies are negative. But the firm is not in a position to "cherry pick": either it is located at the agglomeration, accepting whatever disadvantages are present (in the knowledge that particular advantages more than offset these), or it is located elsewhere. We are assuming the former case, however. The value of overall net agglomeration economies, as well as the imputed value of each net agglomeration economy (diseconomy), is derived by a procedure similar to that outlined above. As an example, we consider the situation where gross cost savings (per unit of output) associated with each agglomeration economy are as follows: for localisation economies, £40; for activity-complex economies, £30; and for urbanisation economies, –£20. The relative shares are therefore 80 per cent, 60 per cent and –40 per cent, respectively. Applying these shares to the value of net overall agglomeration economies (pre-calculated at £25 per unit of output), we obtain the following imputed values for each net agglomeration economy (diseconomy): localisation economies, £20; activity-complex economies, £15; urbanisation economies, –£10. This example provides a further illustration of the fact that agglomeration economies may sometimes only be realised if a cost is incurred. Note that by this procedure an individual net agglomeration economy will be negative, if the relevant gross-savings element is also negative.

*Reinforcing Agglomeration Economies*

It will be recalled that in calculating the value of an individual agglomeration economy (or overall agglomeration economies), the benchmark location is not allowed to coincide with the agglomeration in question. This requirement is accommodated by selecting the next-best quasi-optimal location as the benchmark location. Without this procedure the value of the individual agglomeration economy (or overall agglomeration economies) would assume a zero value. Let us consider the case where it is necessary to employ such a device. The situation is an interesting one, because although the firm may be optimally located at a particular agglomeration, the optimal nature of this

location is wholly independent of the existence of agglomeration economies. In other words, the firm benefits from one or several agglomeration economies at the agglomeration in question, but would still prefer a location there, even if such an advantage was not available. The agglomeration economy thus augments the existing locational advantage of the agglomeration. This represents the case of "reinforcing agglomeration economies", the significance of which is generally not remarked upon. Agglomeration economies possessing such a characteristic are probably more common than is supposed, and will be considered further in the next section.

## 5. THE ABSENCE OF AGGLOMERATION ECONOMIES AT AN AGGLOMERATION

It is sometimes the case that firms in the same industry are found at a particular location, yet appear to enjoy no localisation economies. Examples of such cases might include flour mills in the Buffalo area and steel works in and around Chicago. According to McCann (1995, p.572) such a phenomenon may simply be the result of the firms concerned having the same suppliers or relying on the same raw-material locations and also serving the same markets. Under these conditions it would not be surprising for the firms to have a common location, even though no localisation economies were present. Generalising the argument (i.e. without sole reference to like firms), McCann (1995, p.573) argued that "distance-transaction costs" (broadly speaking, transportation costs) can result in a firm being optimally located at the location of other firms, although its presence does not involve trading relationships with local firms, so that "such clustering is purely the incidental result of optimising behaviour". Taking a longer-run perspective, Coe and Townsend (1998, p.386) offered the somewhat different interpretation that "the spatial juxtaposition of industries may represent only inertia". McCann (1995) went on to cite a further example, involving a location which had lost its manufacturing export base, but from which the skilled unemployed had been reluctant to migrate. The fact that firms might subsequently be attracted to such a location would not reflect the possibility of agglomeration economies but rather the availability of relatively low-cost labour. This would, of course, be consistent with standard Weberian analysis, which would view such a locational tendency as labour orientation.

The distinction as to whether the agglomeration of economic activity is based on agglomeration economies or is largely independent of these is an important one, and one which certainly needs to be borne in mind in econometric analyses concerned with

the spatial association of economic activities. A certain degree of caution is therefore required. Two points are important in this connection. First, in all the cases cited above the very fact that the individual firm is located at an agglomeration rather than in a company town or at a greenfield site (both, say, 50km away) suggests that it might be benefiting from urbanisation economies, involving the shared access to infrastructure, utilities, public facilities, amenities etc. Admittedly, these may be of secondary or minor significance. Second, and more important perhaps, even if the co-location of similar and/or different firms is wholly independent of the existence of agglomeration economies, the firms concerned (or certain of their number) may nevertheless benefit from these, in which case we are speaking of reinforcing agglomeration economies, as discussed in Section 4. Simply because agglomeration economies are not decisive in a locational decision does not necessarily mean that these are absent.

A similar doubt about the existence of agglomeration economies appeared in a thoughtful commentary by Crampton and Evans (1992). Drawing on work by Nicholson *et al.* (1981), which was part of a study of the incubator hypothesis concerned with location of new engineering firms in inner London, they argued that "the majority sold to the printing or construction firms located there. There was no evidence …. that firms located there because of agglomeration economies. They did not set out to minimise the distance of their plant from services or suppliers. They did, on the other hand, locate there to maximise the size of their initial market" (Crampton and Evans, 1992, p.263). This conclusion is evidently based on their particular definition of agglomeration economies: one in which proximity to "services or suppliers" is recognised as a basis for agglomeration economies, while proximity to the firms comprising the market is not. But in this latter situation there exists the possibility of agglomeration economies of the activity-complex type, particularly with respect to the transfer of goods and services as well as the advantages of co-ordination between engineering firms and the printing or construction firms. While the maximisation of revenues may well have been the primary motivation for engineering firms to select a location at the market, this hardly negates the existence there of agglomeration economies, a possibility apparently not explored. The advantage of treating activity-complex economies as a separate type becomes apparent, as does the need to consider the possible existence of reinforcing agglomeration economies.

Thus, to the question of whether agglomeration economies may be absent from a particular agglomeration, the answer must be that while this is possible, it is probably

not likely. As we saw in the previous section, a location within an agglomeration may require the firm to incur certain additional costs (including agglomeration diseconomies), so that unless offsetting benefits in the form of agglomeration economies are also available, such a location is unlikely to be sustainable over the long term.

## 6. AGGLOMERATION ECONOMIES AND SPACE

Agglomeration economies, by their very nature, involve an obvious but important spatial aspect: concentration. However, the influence of space is present in at least three other respects. We consider first what might be termed "intra-urban space", and this is especially important for externality-based agglomeration economies. Agglomeration economies are frequently spoken about with reference to intra-urban space, and this can sometimes be misleading. Although agglomeration economies involve a location within an urban centre, it is often not at all clear whether this requires immediate proximity to other economic activity, or simply a location within the same urban centre. For many kinds of urbanisation economy a firm is able to gain access to these at virtually any location within the urban centre. Under such circumstances the urban centre can be treated as a point location, a practice which becomes increasingly questionable, the larger the territorial extent of the urban centre. For certain kinds of localisation or activity-complex economies, however, the individual firm may have to be very discriminating in its locational decision, since the distance-decay effect of a location away from similar or from linked economic activity may be considerable. This factor was especially important in earlier eras (Scott, 1982; Vernon, 1960).

A second type of space which is relevant to agglomeration economies is concerned with "regional space". Increasingly, we encounter the term "regional agglomeration economies", although this has several distinct meanings. Sometimes the term is employed simply as locational qualifier, drawing attention to the fact that agglomeration economies are available to firms at particular locations within a given region. At other times the term is used to emphasise the fact that agglomeration economies are inextricably tied to the process of regional economic development. In the competition for investment among regions the existence of powerful agglomeration economies can prove a decisive factor. Isard (1960, p.404) suggested the term "urbanization-regionalization economies" in order to emphasise this very point. A further usage of the term refers to the fact that the advantages customarily associated with externality-based agglomeration economies are actually available to the relevant firms throughout (or

1
4

over a significant part of) a given region. McCann and Fingleton (1996), for example, use the term "regional agglomeration economies" in connection with the computer industry of Central Scotland. This use of the term seems to be based on the view that the advantages which were traditionally dependent upon economic activity being organised around a well-defined localisation, urban concentration or activity complex (or some combination of these) can now be realised with a regional spatial structure which is more dispersed. Such a trend is becoming increasingly apparent within the so-called "new industrial districts" of Third Italy and Baden-Württemberg (Harrison, 1992; Scott, 1988).

But to describe the advantage to firms in this setting as "regional agglomeration economies" is misleading as well as mildly contradictory. In terms of the scheme outlined in Section 2, the economies being referred to represent external economies which are not dependent on immediate spatial proximity. However, the spatial constraint, though less severe, may not be entirely absent, so that the influence of space cannot be completely discounted. For this reason, and in the spirit of Robinson (1956), we may use the term "partially-mobile external economies" or "regionally-mobile external economies", rather than the less precise term "regional agglomeration economies". This characterisation seems in keeping with the view of Moulaert and Djellal (1995), who argued that agglomeration economies which have traditionally been associated with urban concentrations, need to be redefined in terms of alternative scales and spatial configurations. It is also consistent with the work of Capello (2000) on "urban network externalities". Furthermore, the characterisation goes some way to meeting the objections of Allen (1991) and Coe and Townsend (1998). The latter authors argued that the influence of agglomeration economies (which historically gave rise to what they unpleasingly referred to as "localized agglomerations") have come to be replaced by regional (i.e. region-wide) advantages, especially in the case of non-manufacturing activity.

A third kind of space relevant to agglomeration economies involves the territory of a city region or metropolis-based region. Such an entity consists of a metropolitan zone and a surrounding non-metropolitan zone. The latter is dominated economically by the metropolis, and contains a rural population as well as an urban population distributed across a network of centres, usually arranged within a hierarchy. These metropolis-based regions have become prominent in the spatial organisation of the economic system in most developed nations, although their origins date from the era of

industrialisation and urbanisation in the nineteenth century. During earlier times, agglomeration economies could only be realised if there was very close proximity of economic activity within the metropolis. This was largely due to the very high cost of intra-urban transportation, which was reflected in the relative compactness of cities (Fales and Moses, 1972). But this very compactness was the cause of high operating costs, resulting from location rents and congestion. Thus to gain agglomeration economies, firms had to incur substantial additional costs or diseconomies of agglomeration, a situation which prevailed well into the twentieth century. It was rendered less pronounced, however, first by improvements in intra-urban transportation, encouraging decentralisation or suburbanisation (Moses and Williamson, 1967), and somewhat later by corresponding improvements in interurban or intraregional transportation and communications. These have given rise to such trends as working from home, teleconferencing and changes in the spatial organisation of the firm, all of which have increased the tendency toward regional deconcentration (Mogridge and Parr, 1997).

Under these changed circumstances the firm that nowadays shuns the metropolis in order to avoid diseconomies of agglomeration (particularly those of the urbanisation type), is probably not required to forgo the benefits of a metropolitan location, or at least not all of these. To an increasing extent many of the advantages that were once to be considered agglomeration economies (because of the required proximity of economic activity) are now available well beyond the confines of the metropolis. Obvious examples include access to major ports, international airports and freight terminals, as well as the availability of a wide range of specialist financial and business services. For many firms the benefits of a metropolis no longer require a location within it or even in close proximity, but merely accessibility to it. What seems to be happening is that the incidence of diseconomies of agglomeration is largely confined to the metropolis, while the benefits generated within the metropolis are not nearly so locationally circumscribed as formerly. The contemporary metropolis in nations of the developed world is increasingly able to supply a wide array of manufactured and service inputs to firms located within its wider hinterland. We appear to be witnessing what Richardson (1995, p.146) described as "the dissipation of agglomeration economies or at least a major extension of their spatial range". But here again the problem of terminology intrudes, and a case exists for abandoning the term "agglomeration economies", when attempting to describe such an outcome. As an alternative, and in the manner of Christaller (1933)

and Lösch (1944), we may justifiably refer to the relevant activities as "central place functions", these being supplied from the metropolis to a reasonably well-defined (though not continuous) market area (Parr, 1999). Indeed, such a role can be expected to form an increasingly important part of the economic base of the large metropolis.

## 7. CONCLUDING COMMENTS

Despite its widespread use the concept of agglomeration economies is not a straightforward one, largely because it is related to a number of diverse economic phenomena. Furthermore, the concept is frequently used imprecisely. In this connection, Cappellin (1988) has pointed to the dangers of the concept being employed in a tautological sense and to the need for exploring the basis on which firms seek to reduce costs by selecting a common location, a concern re-iterated in somewhat different terms by McCann (1995). Attention has been focused here on the different types of agglomeration economy and, by implication, their co-existence. The proposed classification of agglomeration economies contained three salient features. First, it made a clear distinction between agglomeration economies which were based on internal economies to the firm, and those based on externalities, thus emphasising an important parallelism of structure. Second, in dealing with the former, the classification considered agglomeration economies based on internal economies of scope and of complexity, in addition to those based on the more familiar economies of scale. Third, with respect to agglomeration economies based on external economies, the classification interpreted localisation economies in terms of scale, and urbanisation economies in terms of scope, but also included a further category based on complexity. As with any system of classification, difficulties of identification may arise, and attention was drawn to the need to consider the broader context in which agglomeration economies exist, whether this is organisational, structural or temporal.

There are, in addition, a number of more general issues, which have tended to receive inadequate attention. Of particular importance in this connection is the residual nature of agglomeration economies, i.e. the fact that the benefits of agglomeration are usually only realised by incurring a cost. In those instances where this is not so, we have the case of reinforcing agglomeration economies, usually related to agglomeration economies of the urbanisation type. This has implications for another issue. The view has been expressed that the agglomeration of economic activity at a given location may simply be the result of coincidence or a reflection of the spatial organisation of production during some previous industrial era, rather than the presence of

agglomeration economies. This important question is clearly an essentially empirical one, which has to be resolved on a case-by-case basis. In all cases, however, the subtle influence of reinforcing agglomeration economies is not to be underestimated. Finally, attention was drawn to the spatial setting of agglomeration economies. By their very nature these are concerned with spatial concentration, but other spatial factors are involved. In the case of intra-urban space, for example, considerations relating to the physical extent of an agglomeration may be of critical importance. When we turn to regional space, we encounter the changing nature of agglomeration economies and the fact that the benefits once associated with metropolitan areas are slowly coming to be replaced with benefits which are more properly thought of in region-wide terms. The concept of agglomeration economies continues to have considerable potential for assisting our understanding of locational decisions and the spatial structure of the economy, but unless the concept is employed with precision and caution, it is in danger of losing some of its cutting edge and even becoming the source of confusion.

## References

Allen J, 1991 "Services and the UK economy: regionalisation and economic dislocation" *Transactions of the Institute of British Geographies* 17 292-305

Capellin R, 1998 "Transaction Costs and Urban Agglomeration" *Revue d'Economie Régionale et Urbaine* 2 261-278

Capello R, 2000 "The city network paradigm: measuring urban network externalities" *Urban Studies* 37 1925-1945

Christaller, W, 1933 *Die zentralen Orte in Süddeutschland*. Translated by C. Baskin, 1966, as *Central Places of Southern Germany* (Prentice-Hall, Englewood Cliffs, N.J.)

Coe N, Townsend A R, 1998 "Debunking the myth of localized agglomerations: the development of a regionalized service economy in South-East England" *Transactions of the Institute of British Geographers* 23 385-404

Crampton G, Evans A, 1992 "The economy of an agglomeration: the case of London" *Urban Studies* 29 259-271

Evans A, 1989 *Urban Economics: An Introduction* (Blackwell, Oxford)

Fales R, Moses L, 1972 "Land-use theory and spatial structure of the nineteenth century city" *Papers of the Regional Science Association* 28 49-80

Florence P S, 1961 *The Logic of British and American Industry* (Routledge and Kegan Paul, London)

Goldstein G S, Gronberg T J, 1984 "Economies of scope and economies of agglomeration" *Journal of Urban Economics* 16 91-104

Harrison, B, 1992 "Industrial districts: old wine in new bottles?" *Regional Studies* 26 469-483

Hoover E M, 1937 *Location Theory and the Shoe and Leather Industry* (Harvard University Press, Cambridge, Mass.)

Isard W, 1956 *Location and Space-Economy* (The Technology Press of Massachusetts, Institute of Technology, Cambridge, Mass.)

Isard W, 1960 *Methods of Regional Analysis* (The MIT Press, Cambridge, Mass.)

Lichtenberg R, 1960 *One Tenth of a Nation* (Harvard University Press, Cambridge, Mass.)

Lösch, A, 1944 *Die räumliche Ordnung der Wirtschaft*. Translated by W.H. Woglom and W.F. Stolper, 1954, as *The Economics of Location* (Yale University Press, New Haven, CT.)

McCann P, 1995 "Rethinking the economics of location and agglomeration" *Urban Studies* 32 563-577

McCann P, Fingleton B, 1996 "Regional agglomeration impact of just-in-time input linkages: evidence from the Scottish elections industry" *Scottish Journal of Political Economy* 43 493-518

Marshall A, 1892 *The Economics of Industry* (Macmillan, London)

Mogridge M, Parr J B, 1997 "Metropolis or region: on the structure and development of London" *Regional Studies*, 31, 97-115

Moses L, Williamson H F, 1967 "The location of economic activity in cities" *American Economic Review* 57 211-222

Moulaert F, Djellal F, 1995 "Information technology consulting firms: economies of agglomeration from a wide area perspective" *Urban Studies* 32 105-122

Nicholson B M, Brinkley I, Evans A, 1981 "The role of inner cities and the development of manufacturing" *Urban Studies* 18 57-71

Ohlin B, 1933 *Interregional and Internal Trade* (Harvard University Press, Cambridge, Mass.)

Palander T, 1935 *Beiträge zur Standortstheorie* (Almqvist & Wiksells, Upsala)

Panzar J C, Willig R D, 1981 "Economies of Scope" *American Economic Review* 71 268-272

Parr J B, 1999 "Decentralization and deconcentration: the metropolitan area in its wider setting", in *Urban Change in the United States and Western Europe: Corporation Analysis and Policy* Eds. A Summers, L Senn (The Urban Institute Press, Washington, DC) pp 217-242

Parr J B, 2000 "Missing elements in the analysis of agglomeration" paper presented at the Sixth World Congress, Regional Science Association International, Lugano, Switzerland

Richardson H W, 1998 "Economies and diseconomies of agglomeration", in *Urban Agglomeration and Economic Growth* Ed. H Giersch (Springer-Verlag, Berlin) pp 123-155

Robinson E A G, 1956 *The Structure of Competitive Industry* (Cambridge University Press, Cambridge, UK)

Scott A J, 1982 "Locational patterns and dynamics of industrial activity in the modern metropolis" *Urban Studies* 19 111-142

Scott A J, 1988 *New Industrial Spaces: Flexible Production Organisation and Regional Development in North America and Western Europe* (Pion, London)

Vernon R, 1960 *Metropolis 1985* (Harvard University Press, Cambridge, Mass.)

Weber A, 1909 *Über den Standort der Industrien* (Mohr, Tübingen). Translated by C. J. Friedrich, 1929, as *Alfred Weber's Theory of the Location of Industries* (University of Chicago Press, Chicago)