An Investigation of Market Areas and Supply Areas: An Integrated Framework

Daisuke Nakamura

Doctor of Philosophy

Department of Urban Studies
Faculty of Law, Business and Social Sciences

October 2006

For my family

Acknowledgements

This is a thesis submitted in part requirement for a Doctor of Philosophy at the University of Glasgow. The author wishes to thank his parents and family, as well as the academic and support staff in the Department of Urban Studies at the University of Glasgow. He also wishes to record his thanks to the staff of the Faculty of Law, Business and Social Sciences. Responsibility for the contents of this thesis remains solely with the author.

DAISUKE NAKAMURA

Glasgow, Scotland
October 2006

Abstract

The concern of the thesis is to clarify the structural relevance between market areas and supply areas through the investigation of firm location under the given conditions of market demand, deposit of inputs and technologies for production. Conventional economic analysis studies a solid interaction between input and output, through the structure of the production function, by means of the duality theory in the input-output framework. This corresponds to the framework of market areas and supply areas in location theory. However, the existing market-area analysis and supply-area analysis focus examination on an independent framework, and a series of approaches has not been sufficiently developed. Although the integrated framework of both types of area would be treated as an extended version of the duality theory, the framework would not be complete unless the analysis took additional spatial factors into consideration. These factors are suggested to be parts of spatially unconstrained and constrained internal and external economies. The spatially constrained types of economies are called agglomeration economies and these, together with spatially unconstrained types of economies, constitute the neglected factors in existing market-area analysis and supply-area analysis. As agglomeration economies have a trade-off interaction with transportation costs, an analysis of assembly and distribution transportation costs is also required. This research clarifies these neglected factors and considers them with the duality theory, applying the input-output framework to both types of area analysis. This alternative approach not only demonstrates the effects of market area change on the spatial structure of supply area and vice versa, but also investigates the incentives governing the determination of the firm location.

Table of Contents

Lists of Figures

Chapter 1. Introduction

This research will clarify the structural relationship between firm location, market areas and supply areas. To be precise, this analysis concerns how inputs are obtained from supply areas to the assembly plant, how processing is engaged, and how the product is distributed to market areas under given conditions of spatial and economic organisations. In the past, market areas have been analysed in terms of spatial competition, with demand conditions, technology and factor prices given. By contrast, supply areas have been examined with respect to spatial competition of inputs with the given structures of assembly cost, technology and the demand conditions of output. Although both types of area have been studied in various types of approach, neither market-area analysis nor supply-area analysis has dealt with the location of production. In order to investigate the optimal firm location with respect to a given market-area and supply-area structures, every independent framework needs to be integrated in a single framework. It is possible to analyse the relationship between market areas and supply areas by combining both types of area framework. As a major concern of each type of area analysis is solely spatial competition and formation, it should be noted that these approaches all assume firms to be located at the centre of an area. However, this assumption may cause problems when a series of economic activities is taken into account.

From the standpoint of a producer, every plant has supply areas to obtain inputs from suppliers, and market areas to distribute output. As both types of area analysis assume the plant to be at the centre of the area, this firm must logically be located at the centre of the market area and supply area. However, this hypothesis cannot be applied in general, and is particularly unsuitable for manufacturing firms. This logical problem is caused by the fragmented approach of established location analysis. As a result, this thesis will attempt to integrate market areas and supply areas by means of the theory of firm location. Treatment of an integrated methodology will be prefaced by a consideration of the input-output framework which is relevant to the duality theory of conventional economic analysis. This theory states that the unknown cost function is derived from the given production function and structure of factor cost, and that the unknown production function is derived from the given cost function and structure of factor cost. From these relations, the definition can be established that the unknown structure of factor cost is derived from the given cost and production functions. In this way, the duality theory can be restated in terms of an input-output framework: *The unknown cost function is derived from the given structure of factor cost and production function. By contrast, the unknown structure of factor cost is derived from the given cost and production functions.* It is possible to apply this alternative framework to location analysis as follows: *The spatial configurations of market areas are derived from the given spatial production function and supply-area configurations, in addition to the condition of market-area organisation. Likewise, the spatial configurations of supply areas are derived from the given spatial production function and market-area configurations, in addition to the condition of supply-area organisation.* This is a spatial version of duality theory under the particular assumptions and will be termed the "spatial duality theory (SDT)".

In addition, the structure of the production function should be extended. The conventional production function normally contains economies of scale which are parts of internal economies. In this thesis, it is also necessary to introduce the whole notion of spatially unconstrained and constrained internal and external economies. These factors in technological parts will be combined with the production function and will be termed the "spatial production function". In addition, non-technological parts will be inserted in the structure of the spatial factor cost curve. It will be clear during the analysis that the market areas and supply areas can be linked through the spatial production and cost functions. This linkage will constitute an integrated framework analysis enabling us to observe the optimal firm location. The establishment of the integrated framework will be examined by comparative-static analysis in terms of market areas and supply areas. These generalised results will be applied to eight representative hypothetical examples and will be observed by the effects of various economic forces on the decision-making of the firm location. Finally, this research will clarify that the additional economic factors have a crucial role in the existing framework of market areas and supply areas. It will also indicate further avenues of research with respect to spatial market competitions, in the notion of cooperative and competing relationships between firms.

1.1. Background to the Research

There are three main core factors in this research; namely firm location, market areas and supply areas. The study of firm location was initially investigated by Alfred Weber (1909) in terms of the location of industries. Although the concept of distance had already been examined by Launhardt (1885), the factors governing the location of industries were not systematically formalised until Weber. His analysis is based on the Varignon frame which determines the centre of gravity of a triangle and he applied this method to location-triangle analysis. The location triangle is described by two distant raw-material deposits and one market at each apex of a triangle. In addition, there is a production plant processing the related inputs and output. The optimal firm location is found at the centre of gravity of this triangle relying on the ratios of transportation rates for inputs and outputs with respect to weight and distance. This framework was formally generalised by Moses (1958), applying the concept of the production function. It was further developed by Khalili, Mathur and Bodenhorn (1974), with a more generalised form of the production function, while Hwang and Mai (1992) examined the influence of consumer demand. However, it should be noted that they all dropped the notion of economic factors in terms of agglomeration economies due to simplifications of conditions of the analysis. In addition, these approaches assume supply points and market points, rather than supply areas and market areas. In this way, the extensions of Weber present certain limitations to further detailed analysis.

Area analysis was initially developed by Lösch (1938; 1954) in terms of market areas. Although he also referred to the notion of supply areas, supply-area analysis had few expansions and was not fully generalised in his literature. Market-area analysis was generalised by Mills and Lav (1964). Although they established a simplified alternative framework, a crucial problem exists in their analysis, rendering it invalid to Lösch's framework. The original Löschian analysis was further extended by Denike and Parr (1970). They investigated the proper definition of the spatial demand curve and detailed the spatial equilibrium framework under the free-entry condition. Market areas have also been analysed in terms of the urban system with respect to the economic law of market areas and the law of retail gravitation. The economic law of market areas was formally introduced by Fetter (1924) and generalised by Hyson and Hyson (1950). The law examines the economic territories of the market area in terms of price and freight-rate competition. Parr (1997; 2002b) applies these approaches to the Löschian central-place model by means of a framework provided by Launhardt (1885). There is also another economic law, named the law of retail gravitation. This was formally introduced by Reilly (1929; 1953). Since then, Hoover (1971) has generalised these retail relationships. This is concerned with spatial competition, and examines the market shares between two cities at a third location, the share being determined by the ratio of the two city populations and the distance to the third location.

While supply-area analysis is explicitly stated by Isard (1956) and Beckmann (1968), no generalisation was conducted until the investigation of Parr (1993a; 1993b). Further extensions have been attempted by Parr and Swales (1996; 1999) with respect to a wider industrial approach. They investigate the spatial structure of supply areas referring to given market competition and output level. In addition, they also examine supply-area configuration in terms of circular, hexagonal and truncated configurations. On this point, it is worth noting that additional factors will be required for the integrated framework analysis, namely spatially unconstrained and constrained internal and external economies. Although these economies were examined by Marshall (1890), Weber (1909), Hoover (1948), Isard (1956) and Evans (1972), each single approach was independently examined and the relevance of one to the other has not been clarified. These independent studies were systematically integrated by Parr (2002c), as the categorisation into six types of agglomeration economies. These additional factors will be a part of the spatial production function and factor cost. An integrated framework will be composed of market areas and supply areas in addition to these factors by way of the application of the duality theory, which was originally formalised by Shephard (1953). However, this approach has not yet been applied to location analysis.

1.2. Thesis Structure

The thesis will consist of eight chapters. Following the introduction in Chapter 1, market-area analysis and its relevant extensions will be examined in Chapter 2. In this chapter, the definition, conditions of the analysis, spatial competition and configuration of market areas will be examined. Likewise, supply-area analysis and its relevant extensions will be studied in Chapter 3. This chapter will introduce the existing literature on supply-area analysis and explore further prospective extensions and limitations of the analysis. The similarities and dissimilarities in these two types of area will be investigated in Chapter 4. In addition, this chapter will further explore the additional economic factors which are required but not contained in the existing framework of both types of area. These additional factors will be introduced to both types of area in the following chapter. Chapter 5 will investigate the structural relationship between firm operation and internal and external economies. In addition, agglomeration economies and firm location will be investigated by means of the extended Weber location-triangle approach. Furthermore, firm location and spatial competition will be analysed, followed by the examination of firm location in terms of the three types of industry. This will refer to agglomeration economies and the Weber location-triangle approach.

With the completion of the above investigations, an integrated-framework analysis will be presented. In addition, the alternative duality theory will be initially examined, followed by the spatial duality theory *(SDT)* analysis in Chapter 6. These will be examined by means of comparative-static methods with respect to spatial configurations of market areas and supply areas. Furthermore, the analysis will be applied to various spatial economic patterns in Chapter 7 with eight examples. These hypothetical examples will also be analysed under several patterns of spatial competition. Finally, concluding comments will be provided in Chapter 8 identifying the necessity of neglecting economic factors in the existing analysis of market-area analysis and supply-area analysis. In addition, further avenues of research will be suggested.

Chapter 2. Market-Area Analysis

2.1. Introduction

This chapter will consist first of an examination of the analysis of market areas, referring to Lösch (1954), followed by further methodological expansions in terms of two types of economic law. Second, the structure of market areas with generalised spatial organisations will be examined. Finally, the limitations of independent market-area analysis and the problem of the single-framework approach will be provided.

2.2. An Overview of Market-Area Analysis

The concept of market areas was first formalised as a part of the economics of location in Lösch (1954). The definition of the market differs from the terminology of conventional economic theory with respect to the dimensional standpoint of view. In other words, while conventional economic analysis considers the market as a single economic point, market-area analysis treats the market as a two dimensional economic plain. Lösch (1938) claims that the ideal shape of the market area is formed as a regular hexagon under the given conditions of the relevant demand function. Lösch (1954) further expands the general conditions of the location equilibrium, highlighting several problems in the established literature prior to his own work. First, he criticises the study of Weber (1909) which initially introduced the notion of firm location to the economic literature. According to Lösch, Weber's approach becomes invalid if the theory is applied to rich countries. In addition, Weber's study is criticised for giving insufficient consideration to consumers. Second, Predöhl (1925) is criticised for failing to sufficiently investigate the evidence within the framework of individual economic units. Finally, Schneider (1935) is criticised with respect to the arrangement of dependent and independent variables in his work and specifically with regard to his treatment of the locations as assumed rather than subject to investigation. In the light of these criticisms, Lösch reconsiders the point that location equilibrium is obtained at the highest profit for a producer and the lowest cost for a consumer. He also omits several economic factors which might cause difficulties for detailed analysis and suggests the following assumptions: that there is a distribution of industrial raw materials over a wide plain, an evenly distributed agricultural population, and that every economic agent lives in a similar environment and has equal access to all industries and their production methods.

Under these assumptions, Lösch provides five conditions of his approach. *Condition 1* assumes that individuals have to maximise their location. This condition solely affects the shape of the area and maximises the profit of a given number of producers. As a result, other elements such as the size of the area are not taken into account at this stage. *Condition 2* considers that the entire space is always sufficiently occupied by numerous locations and implies that there are no administrative and geographical constraints in economic space. However, Lösch did not consider the following case: the space may be filled by circles, with an empty space between any three circles. In such a case, this condition becomes invalid. *Condition 3* defines the absence of abnormal profits. However, some exceptional cases are raised with this condition. Lösch exemplifies a case of an empty physical space which is insufficient to fill one unit of a new entrant. In addition, another case is exemplified where there is a spatially advantageous location but not enough demand from the market. Although Lösch did not extend these points further, the market-area formation relies on the shape of the demand curve *AR* and the cost curve *AC* of the producer. As illustrated in Figure 2-1 (below), these curves shift in various directions in long-run analysis. These effects can be generalised by comparative-static methods in terms of the nature of cost and revenue curves. The comparative-static results will also specify the relationship between price competition of producers and their share of market area according to the configurations and positions of demand and cost curves. These will be further investigated in later chapters. *Condition 4* considers that smaller sizes of area, supply, production and sales are preferred for any economic activity. This can be explained by the Bertrand price competition model where higher entry of firms to the market causes an eventual overload in price competition and unprofitable situations for all enterprises. This condition considers the distance between two objective locations as a dependent variable, and has, as an incentive, the aim of maximising the number of producers. Finally, *Condition 5* considers that the boundary of two neighbouring markets is an indifferent economic position. That is, the boundary line represents a position which has exactly the same circumstances as both sides of the markets.

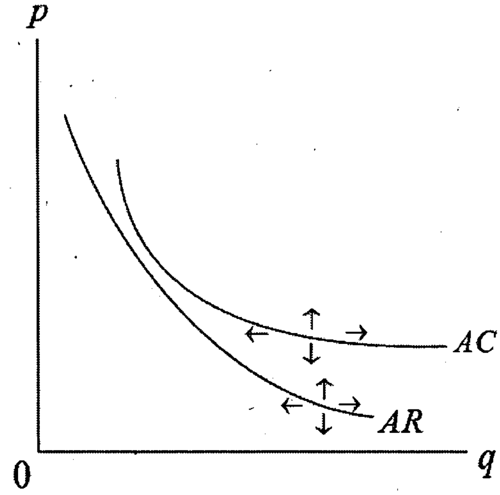


Figure 2-1. Shifts of average cost and demand curves

According to Lösch, the above five conditions are determined by the size of the market area, the limits of market areas, the conditions of production locations both within them and the entire area, and the freight-on-board $(f.o.b.)$ pricing system. It is noted that the optimal location for producers is not always coincident with that of consumers. Although this point has not been extended in his analysis, it may be possible to investigate it further with respect to the analysis of consumers' and producers' surplus. Based on these five conditions, the characteristics of market areas will be examined. Although Lösch does not explicitly examine the impact of producer's cost structure on the structure of market areas, the market area can be analysed in terms of the two opposing forces of size shrink and enlargement. The enlargement of market areas not only brings economies of scale for production, but also causes extra shipping costs. These can be explained by the nature of the average total cost curve ATC in Figure 2-2 (below) as the sum of an average cost curve AC and an average transportation cost curve AT_rC_i $(i = 1,2)$. For the transportation cost, there are two types of structures in location analysis. One is the freight-on-board $(f.o.b.)$ pricing system - where an individual consumer pays his freight cost in order to purchase distant selling commodities - while the other is the cost-insurance-freight $(c.i.f.)$ pricing system -- where producers bear the freight charges for shipment to consumers. If the transportation costs are assumed under $c.i.f.$, then they will reflect the additional average cost to firms as shown in the diagram of a change from ATC_1 to ATC_2, which is caused by a change from $ATrC_1$ to $ATrC_2$. This in turn indirectly affects the purchase price for consumers.

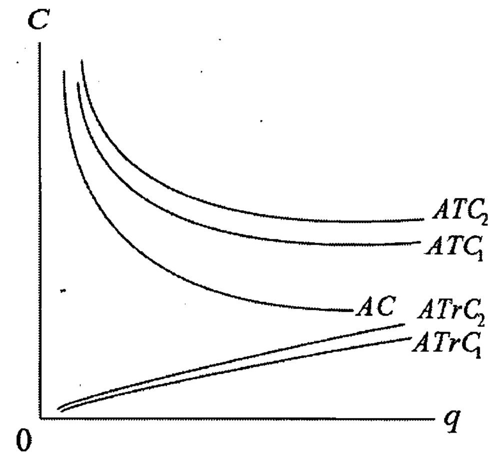


Figure 2-2. Average cost curve AC, transportation curve $ATrC$ and average total cost curve ATC

By contrast, if transportation costs are assumed under the $f.o.b.$ pricing system, the change in transportation rate cannot be observed in the above diagram and this can be seen as the movement of price relevant curves. Löschian analysis generally considers this $f.o.b.$ pricing system. However, this cost increase is still observed in the enlargement and shrinking of market areas for the acquisition of inputs. As a relevant analysis of the structure of cost curves, the following note on economies of scale should be referred to at this stage. In terms of economies of scale, economies can be categorised into two types: large-scale production and specialisation. Large-scale production is feasible in both short-run and long-run conditions, as the movement shifts along an average cost curve. These economies achieve cost-saving production, as the output levels increase. By contrast, a specialised economy is feasible only in the long-run production operation as the presence of fixed cost does not allow the condition of existing production facilities to change in the short run. This change is expressed as the movement of the long-run average cost curve from *LAC* to LAC_A as shown in Figure 2-3 (below). For the shrinking of market areas to occur in this way, all above relations have the opposite force. However, these approaches have not been examined in the Löschian analysis as the production process is treated as a given parameter. These effects, namely those which the technological entities for production processes have on firm location, will be explored in later chapters.

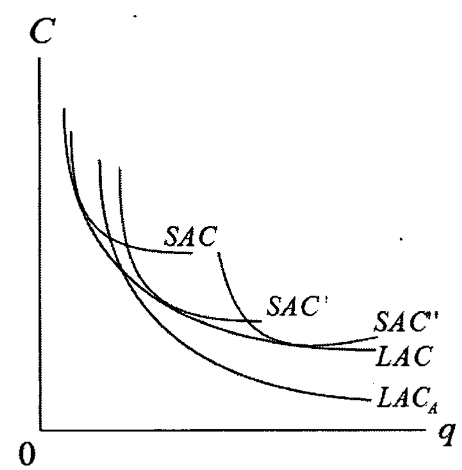


Figure 2-3. The short-run and long-run average cost curves

There are further assumptions made by Lösch (1954) in order to maintain equivalent spatial conditions across the plain. First, raw materials should be evenly and adequately distributed on the plain. Second, the economic plain should be allocated homogeneous areas. Third, economic behaviour should be shown by firm owners intending to produce manufactured goods over and above their own required levels. In addition, regularly distributed firms should be present. Finally, the demand curve should be the same for all individuals. Lösch also suggests that these five additional assumptions can be geometrically observed. Figure 2-4 (below) illustrates an individual demand curve AR for beer. If the price at the brewery is OP, the price increases from OP as the distance increases from the brewery. If the maximum reservation price is OF, no beer can be sold above the freight cost level $PF(=OF-OP)$.

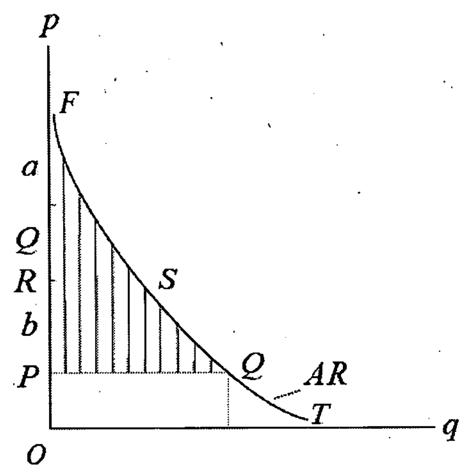


Figure 2-4. The market area from the demand curve for product as a function of distance (Source: Lösch, 1954: 106, modified)

The total sales in this district are illustrated in Figure 2-5 (below) as the rotation of the triangle PQF in the above diagram under the condition of a constant density of demand.

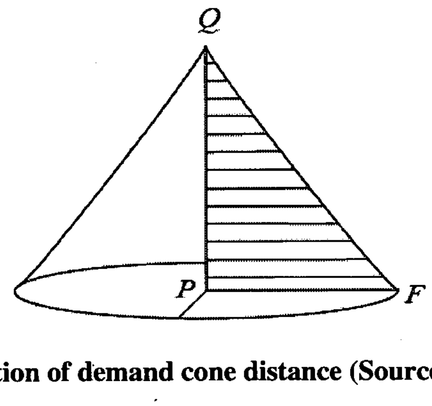


Figure 2-5. Derivation of demand cone distance (Source: Lösch, 1954: 106)

Using these relations, freight costs TrC per kilometre are calculated by the following equation to find the spatial relationship between price and demand:

$$TrC = (Value\ for\ PF) / (Freight\ rate\ per\ Km) \qquad (2\text{-}1)$$

The volume of a solid of revolution equals the area of the generating surface multiplied by the path of its centre of gravity. In order to derive the total demand, the following two assumptions are considered. First, the surface PQF has an area F. Second, the ordinate of its centre of gravity for P and its origin is denoted by y_0. From these assumptions, the centre of gravity revolves along the path $2\pi y_0$ and the area of the generating surface is given as $2\pi y_0 F$. Applying the formula of the centre of gravity, the area of the generating surface can be re-expressed. The centre of gravity is:

$$y_0 F = \int_0^R f(p+t)t\,dt \qquad (2\text{-}2)$$

Therefore,

$$2\pi \int_0^R f(p+t)t\,dt \qquad (2\text{-}3)$$

where R = maximum possible shipping cost, p = mill price at the brewery, t = shipping cost per unit between brewery and consumer and $f(p+t)$ = individual demand as a function of price at the place of consumption. Considering the population density as $b/2$, the total demand d as a function of $f.o.b.$ price p is defined as the following formula.

$$d = b\pi \int_0^R f(p+t)t\,dt \qquad (2\text{-}4)$$

The result of the calculated volume of the demand cone for various arbitrary brewery prices is drawn as the curve Δ in Figure 2-6 (below), where total demand is a function of brewery price. Although Lösch (1954) shows that the demand curve is illustrated as concave to the origin, Denike and Parr (1970) prove that the curves must be convex to the origin by solving the first and second order conditions of the aggregate demand function based on the assumption in Lösch.

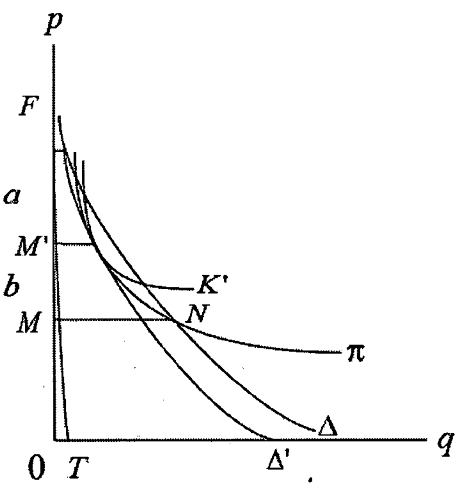


Figure 2-6. The cost curves (Source: Lösch, 1954, 106; and Denike and Parr, 1970)

In the above diagram, the curve FT expresses the consumer individual demand curve. The planning curve π represents the envelope of the short-run average cost curves for plants of various sizes. If the production cost on the curve π and the demand curve Δ do not intersect, either shipping costs are too high or the advantages of large-scale production are too small. The longest market radius or shipping distance is the same as the radius of the demand curve where volume is $(2MN/B)$ and is equivalent to MF. If the short-run average cost curve is given as K', spatial equilibrium is achieved where the demand curve is given by Δ' and abnormal profits disappear. In this case, the maximum market-area radius becomes $M'F$.

In terms of the shape of market areas, Lösch expresses the shape of the region in the following formula, where the volume of the portion of cone cut off by a plain parallel to the axis of rotation at the distance ρ:

$$V = \frac{H}{3}\left(R^2 \arccos\frac{\rho}{R} - 2\rho\sqrt{R^2 - \rho^2} + 2.302632\frac{\rho^3}{R}\log\frac{R + \sqrt{R^2 - \rho^2}}{R} \right) \tag{2-5}$$

The remaining symbols are defined as follows: R = radius of base of original cone (largest possible shipping costs PF); r = the radius of the circle circumscribed about the hexagonal area of base removal $(r < R)$; and H for the height of the demand cone (individual demand at site of factory PQ). The population density is assumed to be 1 for reasons of simplicity.

Lösch states that the advantages of a hexagonal shape are not explained by the ratio of perimeter to area but by the ratio of cone to area. In other words, while the maximisation problem of a honeycomb for bees in terms of space considers a hexagonal column, the maximisation problem of revenue in market areas need to examine a hexagonal pyramid. There is no development in terms of the number of producers in Lösch (1954). However, the argument that incomplete utilisation of the capacity of a region increases the number of firms is provided. This initially causes a producer's price to increase as the economies of scale for production are cancelled out. However, there exists an opposite economic force that smaller scales of production achieve freight cost savings. As a result, consumers either obtain certain benefits from reducing sizes or cause losses. The anticipated result - that the disadvantages of price increases are eliminated by the freight cost savings - should be further investigated through the analysis of cost and demand functions. The interactions between several spatial configurations will be examined in further depth in later chapters.

Lösch (1954) proves that the hexagonal market area is the ideal shape of region when compared with circular, triangular, or square economic regions, in terms of the utilisation of space. For the circular form, the demand of a smaller circle is more advantageous than the larger form as increasing shipment costs reduce the average benefits of the area. Although the circular form is the greatest in terms of the demand of the curtailed sales areas, the hexagon is better in terms of eliminating empty corners. Considering the above economic situations, the unused corners of the circle are utilised. This is shown in Figure 2-7 (below) in the shaded areas by which the shape becomes a hexagon from a circle. In addition, the utilisation of space can also be beneficial to consumers from the standpoint of consumer exclusions.

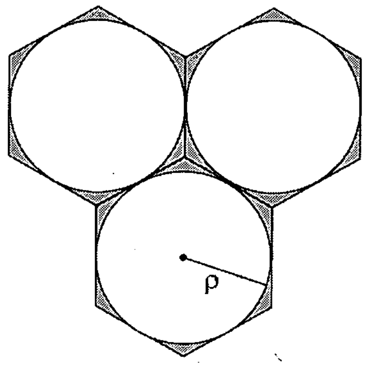


Figure 2-7. Circular and hexagonal market areas

According to the calculation in Lösch (1954), demand in a hexagonal market area is greater than in a square, a circle, or an equilateral triangle of the same area, by 2.4%, 10%, and 12% respectively. These orders of demand proportion are held as long as the demand curve is assumed to be linear. The more elastic demand at the boundary of the region enhances the advantages of the hexagonal shape. As a result, the advantage of the regular hexagon is utilised when a regional shape becomes larger and more rounded, when the demand curve at the boundary becomes more elastic, and when reduced shipping costs are available. The formation of the inscribed circle radius of the hexagonal market area, shown as ρ in the above diagram, depends not only on the condition of the cost curve but also on consumer demand. As consumers are assumed to be equally and continuously distributed, ρ could have any value. However, the number of possible values for ρ is limited if the population is equally but not continuously distributed. The discontinuous distribution of population should be examined from this point of view.

Regarding a situation of the economic activity between separated locations, Lösch provides the following example of farmers as producers and consumers. While farmers who are producers can shorten the distance between farm buildings and fields (with the exception of mixed cropping), farmers as consumers must have a distance between the centre of the village and the field, between the town and the farm buildings, or between the town and the field. In particular, farmers in uncultivated areas who are consumers and also producers, receive help from neighbours by means of cooperatively owned machines, water power, electricity, coal, artificial fertilisers, and selling their products. Although Lösch did not expand the argument further in any depth, its net effect is worth investigating and will be attempted in Chapter 5 with respect to the notion of spatially constrained internal and external economies.

Finally, there are several natural or economic conditions which act as additional constraints in Lösch's analysis. First, the existence of mountains, woods, rich in lakes, and widely occupied farm areas make districts wider. Second, it is difficult to provide extensive fertilisation, churches, government offices, and daily requirements cheaply a small market area. On the other hand, a large market area requires large roads for heavier road traffic. Third, higher freight costs per mile tend to discourage visits to distantly dispersed villages. It is apparent that the actual cost and time have more recently shortened as the quality and cost of roads has been largely improved. While smaller scales increase production costs, this disadvantage is offset by these improvements. Moreover, the invention of the telephone and the radio has helped scattered settlements. The interlacing of markets also increases the presence of villages.

This section has examined the core elements of the analysis of market areas, mainly referring to the approach of Lösch (1954). This approach will be extended in further depth in later chapters with the other approaches of market-area analysis which will be introduced in the following sections.

2.3. The Economic Law of Market Areas

The economic law of market areas was initially introduced by Fetter (1924) in a work which predates that of Lösch (1938; 1954). This law investigates the economic relationship between territorial market boundaries. Fetter describes the economic law of market areas as the law of market-district limits, or the law of market tributary territory. While it brings more precision and clarity into the field of economic studies, it is difficult to define the idea of the extent of the market tributary territory for levels of market prices and freight rates. This section will introduce the theoretical basis of analysis developed by Fetter (1924) and generalised later by Hyson and Hyson (1950).

The basic concept of the economic law of market areas is that there are numbers of buyers and sellers trading as in real economic competitions. There are two forces in a market, namely the centrifugal force from a market and the centripetal force towards the market. For instance, the former could be a manufacturing centre in relation to consumers, and the latter, large exchanges in relation to the scattered producing forms. Middlemen's markets such as stock exchanges and jobbing centres depend on an aspect of groups of trades. It is a necessary condition for the coexistence of two closely related markets to satisfy the following equation.

$$p_A - p_B < tr_{AB} \qquad (2\text{-}6)$$

where p_A and p_B represents market price at cities A and B respectively, and tr_{AB} shows the freight cost between cities A and B. If the necessary condition is not achieved, a protective tariff on goods may be applied in order to survive in the market.

The general economic law of market areas is formulated by applying the above coexisting conditions. The boundary line between the tributary territories of two geographical markets competing for similar good is formed as a hyperbolic curve, as shown in Figure 2-8 (below). At each point on this line, the difference between the freights of the two markets appears equal to the difference between the market prices. It follows that the freight-rate difference and the price difference are unequal on other side of this line. The ratio of prices in the two markets determines the location of the boundary line: the lower the relative price, the larger the tributary area; the higher the base price, the narrower the territory providing the freight rates which are remained constant. If the freight rates decrease with distance, the boundary curves are still symmetrical but the precise shape should be re-examined. It should also be noted that water transportation or topological obstacles change the shape of the boundary line. The formulation can be shown by the isodapane of the Weber analysis although no such attempt has been previously observed. This extension will be conducted in Chapter 5. In this way, the economic law of market areas considers not only the general law of demand and supply, but also transportation costs which vary with distance.

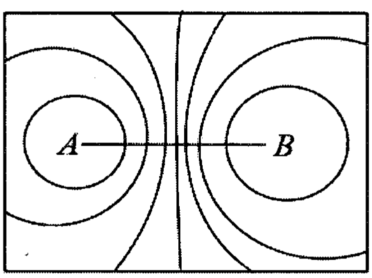


Figure 2-8. Hyperbolic curves and relations of two markets A and B (referred to: Hyson and Hyson, 1950)

Hyson and Hyson (1950) have derived the condition for indifference between two given fixed markets from external consuming points. Their generalisation of the economic law assumes the following three conditions: first, the commodity is standardised; second, whole economic agents have complete knowledge of the market condition; third, freight charges with distance are equal everywhere. With these assumptions, the market prices of the commodity at the markets A and B are set as p_A and p_B respectively. In addition, t_A is assumed to be the freight rate per unit per mile between a location P and the market A, while t_B is the freight rate between P and B. As it is theoretically difficult to examine competing modes of transportation, these freight rates are treated as constant. The indifferent location of two markets for a consumer is then given as the point at which the following equation is satisfied:

$$p_A + t_A D_{AP} = p_B + t_B D_{BP} \tag{2-7}$$

Here, p_A and p_B represent the market price at location A and B, t_A and t_B show the freight rates of A and B, and D_{AP} and D_{BP} express the distance to P from markets A and B respectively. It follows from the above equation that the boundary line between the territorial tributary to markets A and B can be rewritten as:

$$D_{AP} - \frac{t_B}{t_A} D_{BP} = \frac{p_B - p_A}{t_A} \tag{2-8}$$

The theoretical relation of each parameter is interpreted by the following three points: the curve becomes one branch of a hyperbola when $t_A = t_B$ and becomes a circle when $t_A \neq t_B$ and $p_A = p_B$; the curve is regenerated into a straight line when $t_A = t_B$ and also $p_A = p_B$; the size of the tributary area is determined by the relative price at the two markets, the ratio of freight rates, and the ratio of difference in price to freight rates. The results in each case determine the location of the boundary line, and the economic law of market areas is generally stated as follows: that the boundary line between tributary territories is a hyper-circle, and that the freight cost difference equals the price difference between two locations on the hyper circle while it is different on either side of line.

Parr (1997) further examines the law and applies this to the analysis of Launhardt (1885) with respect to five possible outcomes. These are shown in Figure 2-9 (below).

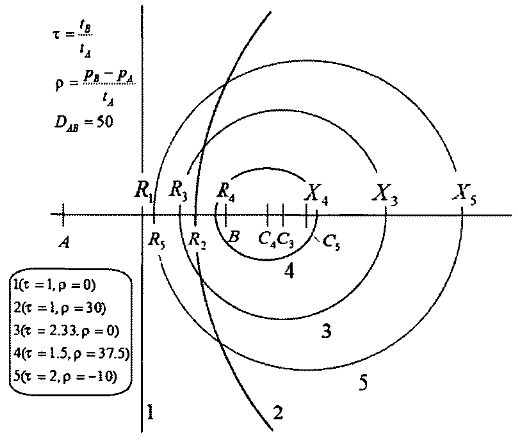


Figure 2-9. The economic law of market areas with five cases (Source: Parr, 1997)

In a case where prices and transportation rates are constants, the transportation-rate ratio t_B / t_A can be expressed as τ $(\tau > 0)$ and the ratio of the price differential to τ is replaced by ρ:

$$D_{AP} - \tau D_{BP} = \rho \qquad (2\text{-}9)$$

The perpendicular bisector case (*Case 1*) is where $\tau = 1$ and $\rho = 0$. When $\tau = 1$ but $0 < \rho < D_{AB}$ (*Case 2*), centre A has a price advantage and the boundary becomes a hyperbola being concave to B and closer to B. When $\tau = 1$, centre A has a transportation-rate advantage and the boundary surrounds centre B. If $\rho = 0$ (*Case 3*), the boundary becomes a circle. If $0 < \rho < D_{AB}$ (*Case 4*), both price and transportation-rate advantages shape the boundary as a horizontally elongated oval. If $-\tau D_{AB} < \rho < 0$ (*Case 5*), centre B has a price advantage and the boundary becomes a vertically elongated oval. Parr (1997) further examines *Cases 3, 4* and *5* with respect to Launhardt (1885). Launhardt analyses market competition with non-local goods and suggests that the expansion of the market can be restricted due to the existence of other shipping points for goods. As shown in Figure 2-10 (below), there are two origins of goods, A and B.

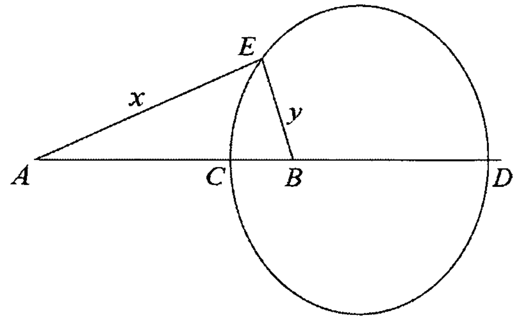


Figure 2-10. Two centres and distances (Source: Launhardt, 1885: 157)

It is assumed that the prices of the same value quantity are p_1 and p_2 respectively. Likewise, freight rates of the same value quantity are f_1 and f_2 and the distance between A and E or B and E is denoted as x and y. If both goods have equal price for the same value quantity, these relationships are expressed as:

$$p_1 + f_1 x = p_2 + f_2 y \quad \textbf{(2-10)}$$

Under certain conditions, an ellipse-family circle is derived with given length l as:

$$CD = \frac{2f_2}{f_2^2 - f_1^2}(p_1 - p_2 + f_1 l) \quad \textbf{(2-11)}$$

In the case $p_1 = p_2$, the ellipse changes to a circle holding the same origin. As a result of competition, the relevant market is formed as a polygon. The boundary is also a territorial boundary. As is shown in Figure 2-11 (below), line l connects two neighbouring markets A and B:

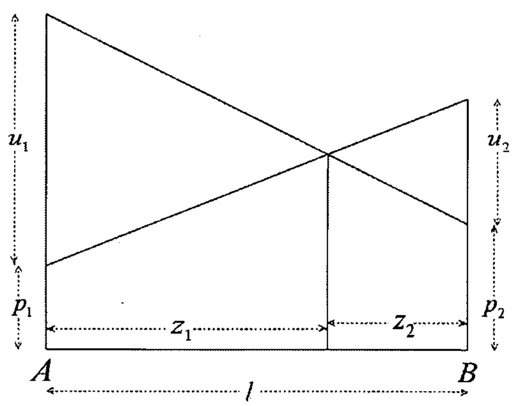


Figure 2-11. Two origins, prices and costs (Source: Launhardt, 1885: 159)

The distance between the market place A and the boundary as z_1 is:

$$z_1 = \frac{p_2 - p_1 + f_2 l}{f_1 + f_2} \tag{2-12}$$

At the point A, the local price difference u_1 between foreign and domestic goods becomes:

$$u_1 = p_2 - p_1 + f_2 l \tag{2-13}$$

By inserting domestic price difference u_1, z_1 is expressed as:

$$z_1 = \frac{u_1}{f_1 + f_2} \tag{2-14}$$

This implies that reducing freight rates leads to an expansion of cheaper goods into the market. Parr (1997) applies the equations in Launhardt to the economic law of market areas by replacing p_1, p_2, f_1, f_2 and l by p_A, p_B, t_A, t_B and D_{AB}. The horizontal diameter of the boundary D_{RX} is then derived from Equation (2-11) with fixed D_{AB} as:

$$D_{RX} = \frac{2t_B}{t_B^2 - t_A^2}\left(p_A - p_B + t_A D_{AB}\right) \tag{2-15}$$

where $\tau > 1$ and $-\tau D_{AB} < \rho < D_{AB}$. Similarly, the distance D_{AR} between centre A and the boundary is derived from Equation (2-12) as:

$$D_{AR} = \frac{p_B - p_A + t_B D_{AB}}{t_A + t_B} \tag{2-16}$$

Likewise the distance D_{BR} between centre B and the boundary is given as:

$$D_{BR} = \frac{p_A - p_B + t_A D_{AB}}{t_A + t_B} \tag{2-17}$$

The condition $D_{AR} > D_{BR}$ is satisfied in *Cases 3* and *4* but satisfied in *Case 5* only accordance with the following expressions:

$$\frac{2\left(p_A - p_B\right)}{t_A D_{AB}} + 1 < \tau < \infty \qquad \rho < 0 \tag{2-18}$$

$$\frac{D_{AB}(t_A - t_B)}{2t_A} < \rho < \infty \qquad \tau > 1 \qquad (2\text{-}19)$$

The horizontal distance D_{BC} between centre B and origin of the diameter of the relevant boundary C is derived from these equations as:

$$D_{BC} = \frac{t_A(p_A - p_B) + t_A^2 D_{AB}}{t_B^2 - t_A^2} \qquad (2\text{-}20)$$

where $\tau > 1$ and $-\tau D_{AB} < \rho < D_{AB}$. The opposite relationship, where centre A is enclosed by the boundary line, can also be obtained.

2.4. The Law of Retail Gravitation

While the previous section has examined the economic law of market areas, there is another approach towards market-area analysis, namely 'the law of retail gravitation'. Reilly (1929) introduced the law of retail gravitation in order to investigate retail relationships in Texas. In this approach, the economic factors are: the different sizes of the cities and towns, income class differences, and the different population sizes of the centre of the cities and towns, and the distances between cities and towns. The size difference among neighbouring cities or towns specifies what kind of retail structure is preferable, and the income class differences indicate consumers' behaviour regarding their purchases. Retail markets are divided into three parts in relation to the classification of different-sized population centres, namely primary, secondary and tertiary retail markets. On the basis of these considerations, the law of retail gravitation is expressed as the following formula:

$$\frac{B_A}{B_B} = \left(\frac{P_A}{P_B}\right)^N \left(\frac{D_B}{D_A}\right)^n \qquad (2\text{-}21)$$

where B_A and B_B denote the business which cities A and B draw from intermediate town T. In addition, P_A and P_B express the populations of cities A and B, and D_A and D_B denote the distances from cities A and B to intermediate town T. The unknown variables N and n are specified as follows: if N is treated as the first power of the population, the variable n can be solved by the following procedure using the above formula:

$$\frac{B_A}{B_B} = \left(\frac{P_A}{P_B}\right)\left(\frac{D_B}{D_A}\right)^n \qquad (2\text{-}22)$$

$$\left(\frac{D_B}{D_A}\right)^n = \frac{B_A}{B_B}\left(\frac{P_B}{P_A}\right) \qquad (2\text{-}23)$$

$$n\log\left(\frac{D_B}{D_A}\right) = \log\left(\frac{B_A}{B_B}\frac{P_B}{P_A}\right) \qquad (2\text{-}24)$$

$$n = \log\left(\frac{\frac{B_A}{B_B}\frac{P_B}{P_A}}{\frac{D_B}{D_A}}\right) \qquad (2\text{-}25)$$

Applying these results to empirical studies, the appropriate supply of commodities in retail stores in cities and towns can be found. The empirical results of n show that business opportunities decline at a rate approximately proportional to that of the square of the distance. This analysis has also been applied to mail-order selling, house-to-house selling, and localised retail-store selling to examine the advantages and disadvantages of each. In addition, Reilly (1953) has further analysed the law of retail gravitation, focusing particularly on the coexistence in a competitive market of large-city and small-town retailers.

The law of retail gravitation was formalised by Hoover (1971). The market area boundary is the point

$$\frac{D_A^2}{D_B^2} = \frac{P_A}{P_B} \qquad (2\text{-}26)$$

where P_A and P_B represent the population of cities A and B, and D_A and D_B show the distance between the boundary and cities A and B respectively. Hoover assumes that there are two cities which have a distance w, and that one city is m times larger than the other city. According to his analysis, the boundary is a circle of radius $w\sqrt{m}/(m-1)$

follows: if N is treated as the first power of the population, the variable n can be solved by the following procedure using the above formula:

$$\frac{B_A}{B_B}=\left(\frac{P_A}{P_B}\right)\left(\frac{D_B}{D_A}\right)^n \qquad (2\text{-}22)$$

$$\left(\frac{D_B}{D_A}\right)^n=\frac{B_A}{B_B}\left(\frac{P_B}{P_A}\right) \qquad (2\text{-}23)$$

$$n\log\left(\frac{D_B}{D_A}\right)=\log\left(\frac{B_A}{B_B}\frac{P_B}{P_A}\right) \qquad (2\text{-}24)$$

$$n=\log\left(\frac{\frac{B_A}{B_B}\frac{P_B}{P_A}}{\frac{D_B}{D_A}}\right) \qquad (2\text{-}25)$$

Applying these results to empirical studies, the appropriate supply of commodities in retail stores in cities and towns can be found. The empirical results of n show that business opportunities decline at a rate approximately proportional to that of the square of the distance. This analysis has also been applied to mail-order selling, house-to-house selling, and localised retail-store selling to examine the advantages and disadvantages of each. In addition, Reilly (1953) has further analysed the law of retail gravitation, focusing particularly on the coexistence in a competitive market of large-city and small-town retailers.

The law of retail gravitation was formalised by Hoover (1971). The market area boundary is the point

$$\frac{D_A^2}{D_B^2}=\frac{P_A}{P_B} \qquad (2\text{-}26)$$

where P_A and P_B represent the population of cities A and B, and D_A and D_B show the distance between the boundary and cities A and B respectively. Hoover assumes that there are two cities which have a distance w, and that one city is m times larger than the other city. According to his analysis, the boundary is a circle of radius $w\sqrt{m}/(m-1)$

and the centre is $w/(m-1)$ miles away from the smaller city. Empirically, larger towns are observed as preferable for a rural family living midway between two cities for convenience purchases. This is especially the case as they not only obtain ubiquitous goods but also differentiated products, namely cinema tickets, clothing, binoculars, and washing machine parts. In addition, larger towns tend to offer an economy of time and money through the availability of a greater range and variety of products. By contrast, small towns are normally located away from the main road making it necessary to travel further to access them. Given the mathematical condition $P_A = mP_B$ as assumed earlier, the relationship between D_A and D_B, which respectively represent the distance between the boundary and cities A and B, is shown in the following equation:

$$\frac{D_A^2}{D_B^2} = \frac{mP_B}{P_B} \quad (2\text{-}27)$$

$$\frac{D_A^2}{D_B^2} = m \quad (2\text{-}28)$$

$$D_A^2 = mD_B^2 \quad (2\text{-}29)$$

As a result:

$$D_A = \sqrt{m}D_B \quad (2\text{-}30)$$

The above equation shows that the size differential between two cities affects the distance between the market boundary and each city by the power of 0.5 under certain economic assumptions based on the law of retail gravitation.

Parr (1997) develops the law further to solve the share or volume of trade at a specific location by considering the size of each centre A, B and its distance from the given point with adjustment parameters of size and distance differentials. The adjusting parameters take $0.9 < \alpha < 1.1$ for size and $1.5 < \beta < 2.5$ for distance according to Reilly's study (1929). The simplest case is where two centres have equal shares $R_{AT} = R_{BT}$ and Equation (2-21) becomes:

$$\left(\frac{D_{AP}}{D_{BP}}\right)^{\beta} = \left(\frac{Z_A}{Z_B}\right)^{\alpha} \quad (2\text{-}31)$$

where Z_A and Z_B = populations of cities A and B. The boundary is shown as the perpendicular bisector in Figure 2-12 (below).

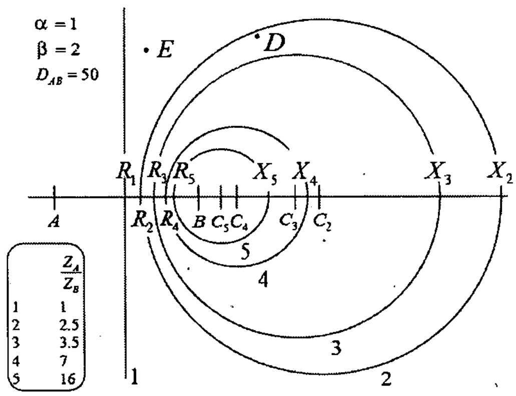


Figure 2-12. The law of retail gravitation with five cases (Source: Parr, 1997)

If Z_A is greater than Z_B, the boundary becomes a circular shape which surrounds centre B. These circles are the market areas of the centre B, and the outside of the circle is the market area of centre A. In the above diagram, the point D represents the majority of purchases made at the centre B. Likewise, the point E represents the majority of purchases made at the centre A. For $Z_A > Z_B$, each diameter of the circle D_{RX} is expressed with fixed distance D_{AB} between centres A and B as:

$$D_{RX} = \frac{2D_{AB}\left(\frac{Z_A}{Z_B}\right)^{\frac{\alpha}{\beta}}}{\left(\frac{Z_A}{Z_B}\right)^{\frac{2\alpha}{\beta}} - 1} \quad (2\text{-}32)$$

The distance D_{AR} becomes:

$$D_{AR} = \frac{D_{AB}\left(\frac{Z_A}{Z_B}\right)^{\frac{\alpha}{\beta}}}{1 + \left(\frac{Z_A}{Z_B}\right)^{\frac{\alpha}{\beta}}} \quad (2\text{-}33)$$

Likewise, the distance D_{BR} is:

$$D_{BR} = \frac{D_{AB}}{1 + \left(\frac{Z_A}{Z_B} \right)^{\frac{\alpha}{\beta}}} \qquad (2\text{-}34)$$

The distance D_{BC} between centres B and C is:

$$D_{BC} = \frac{D_{AB}}{\left(\frac{Z_A}{Z_B} \right)^{\frac{2\alpha}{\beta}} - 1} \qquad (2\text{-}35)$$

2.5. The Relationship between the Two Economic Laws

The similarities of these laws are that they both measure the force of gravity toward the city centre, consumer demand is treated as one of the essential economic factors, and both laws divide the examination into different types of goods. The differences between these laws are first that the economic law of market areas sets its objective function as cost factors while the law of retail gravitations sets it as potential demand factors. Second, there are differences in the theoretical methods applied: the economic law of market areas applies an isodapane analysis and the law of retail gravitation applies a ratio analysis. Finally, there are differences concerning the advantageous condition of the city: for the economic law of market areas it is cost and price circumstances, while for the law of retail gravitation it is the capability of specifying retailing types of goods. Parr (1995a) has compared and contrasted the two theories and explains these differences as follows. The first difference, he explains, lies in the derivation of the outcome; according to the economic law of market areas the outcome is derived from the price differentials between various centres, whereas the law of retail gravitation derives the outcome from the existence of size differences. The second difference is referred to as the division of trade between centres. While the economic law of market areas requires exclusive access to the area, the law of retail gravitation allows the area to be shared. The third difference, finally, is with respect to the preference of categorised goods according to the types of goods. While the economic law of market areas concerns a single good, the law of retail gravitation considers varieties of products.

Parr (1997) compares the two different approaches of the economic law of market areas and the law of retail gravitation, providing some theoretical modification. As is shown in the previous sections, the law of retail gravitation is based on empirical observation while the economic law of market areas is formulated on a theoretical basis. Both laws require the notion of population densities and examine a boundary line between two neighbouring centres in different manners. Whereas the law of retail gravitations treats retail trade in general, the economic law of market areas covers manufactured goods based on industrial bases. As examined in Parr (1997), the market area boundary line is exclusive to a single centre under the economic law of market areas but can be shared by two centres in the law of retail gravitation. In addition, the economic law of market areas considers the entire volume of sales in particular goods whereas the law of retail gravitation refers to a part of sales with respect to the choice of the consumer. An additional two restrictions can be applied to the law of retail gravitation to enable a simultaneous examination with the economic law of market areas: that the boundary has exclusive trade, which is assumed under the economic law of market areas, and that consumers are able to purchase from either centre. Under these conditions, the two laws coincide with *Case 1* of the economic law of market areas and the condition $(Z_A / Z_B) = 1$ in the law of retail gravitation. Interesting cases are those of *Case 3* of the economic law of market areas and $D_{AB} > D_{BR}$ with $Z_A > Z_B$ of the law of retail gravitation. These two cases form a circular shape which surrounds centre B. From this coincidental result, the following law is generated: the condition of the economic law of market areas where $p_A = p_B$ can be equal to the ratio of the centre sizes in the law of retail gravitation with respect to the inverse ratio of the transportation rates. This can be represented by

$$\tau = \left(\frac{Z_A}{Z_B} \right)^{\frac{\alpha}{\beta}} \tag{2-36}$$

where $(t_B / t_A =)$ $\tau > 1$, $Z_A > Z_B$. From this analysis, it becomes clear that the laws are similar insofar as they each examine two centres, and advantages depend on cost factors. The main difference is that different economic factors are applied to specify the boundary.

2.6. Distribution Costs and the Optimal Market-Area Radius

As examined in a previous section, the analysis of Lösch (1954) solves the optimal size and shape of the spatial structure. However, he assumes that the economic space is an entirely homogeneous plain and that there are no other economic forces which attract particular economic areas. By contrast, the two economic laws examine the more complex territorial boundaries between two market areas but do not directly derive the optimal market-area radius. This section will further explore the derivation process of the optimal market-area size in relation to the work of Mills and Lav (1964). The distance and distribution costs for triangular, square, hexagonal and circular market areas will also be considered. As examined by Lösch (1954), for the assumption of distribution cost, it is more common to apply the *f.o.b.* pricing system to the location analysis.

Launhardt (1885) investigates the general idea of freight rates as a dependent variable of the physical distance. Hotelling (1929) applies freight rates to the spatial economic competition and Schneider (1935) introduces price formation and price policy under the consideration of the geographical distribution of producer and consumer. Moreover, Hoover (1937) investigates the market area boundary through the analysis of spatial price discrimination. Mills and Lav (1964) generalise a model of the market area under a free entry condition of the market, examining a location model in a single industry with uniform and undifferentiated spaces. This is an alternative analysis of Lösch which assumes a space-filling hexagonal market equilibrium with profit maximisation behaviour of firms. This alternative model considers a single commodity which can be produced at any point under the same cost function, and assumes a constant average cost which is greater than the relevant constant marginal cost. The total unit cost of production output q is expressed with positive constant integers A and k as:

$$A + kq \tag{2-37}$$

The unit transportation cost per distance u is expressed as:

$$tu \tag{2-38}$$

Using these two relations under the same linear individual demand curve, the demand per consumer q_F is expressed with the *f.o.b.* price p as:

$$q_F = a - b(p + tu) \tag{2-39}$$

where $a =$ an intercept of vertical axis and $b =$ slope of the curve. Under the condition of the regular shape of the market area, the total sales Q are shown with the minimum distance u between the firm and any point of the market area within the maximum radius of the market area U as the following:

For regular s-shaped polygon:

$$Q = 2sD\int_0^{\frac{\pi}{s}}\left\{\int_0^{\frac{U}{\cos\theta}}[a - b(p + tu)]u\,du\right\}d\theta \tag{2-40}$$

For a circular market area:

$$Q_c = D\int_0^{2\pi}\left\{\int_0^{U}[a - b(p + tu)]u\,du\right\}d\theta \tag{2-41}$$

As the total profit is $\Pi = pq - A - kq$, the total profit for triangular Π_T, square Π_S, hexagonal Π_H, and circular Π_C market areas are expressed as the following formula:

$$\Pi_T = (6DU^2)\left(\frac{a\sqrt{3}}{2} - \frac{bp\sqrt{3}}{2} - 0.7969btU\right)(p - k) - A \tag{2-42}$$

$$\Pi_S = \left(8DU^2\right)\left(\frac{a}{2} - \frac{bp}{2} - 0.3848btU\right)(p-k) - A \qquad (2\text{-}43)$$

$$\Pi_H = \left(12DU^2\right)\left(\frac{a}{2\sqrt{3}} - \frac{bp}{2\sqrt{3}} - 0.2027btU\right)(p-k) - A \qquad (2\text{-}44)$$

$$\Pi_C = \left(2\pi DU^2\right)\left(\frac{a}{2} - \frac{bp}{2} - \frac{btU}{3}\right)(p-k) - A \qquad (2\text{-}45)$$

These results can be applied to more convenient mathematical treatments than the generalised equation provided by Lösch (1954). As Denike and Parr (1970) investigate, however, this generalisation does not include implicit functions of price, and potential problems can be found in formation of market-area shapes. Mills and Lav concluded that the optimal spatial structure must be a regular dodecagon shape. However, as long as the spatial competition exists, market areas should not have any corners apart from the regular hexagon. In addition, some of their calculus shows inappropriate assumptions and results. Thus this analysis will not refer to these points. In order to clarify this point, the relationship between implicit price function and market-area structure will be examined in the following section.

2.7. Demand Cone, Demand Curve and Distribution Cost

This section will examine the relationship between market-area analysis and output price, applying the four-dimensional diagram in Parr (2002b).

Figure 2-13 (below) shows the relationship between cost curve, demand curve, and demand cone. In the figure, *Phase* (I) shows the consumer demand curve and *Phase* (II) represents the *f.o.b.* distribution cost. The demand cone is illustrated in *Phase* (III), which is derived by *Phases* (I) and (II). In *Phase* (IV), a straight line is drawn in order to connect *Phases* (I) and (III) by 45° reflection line.

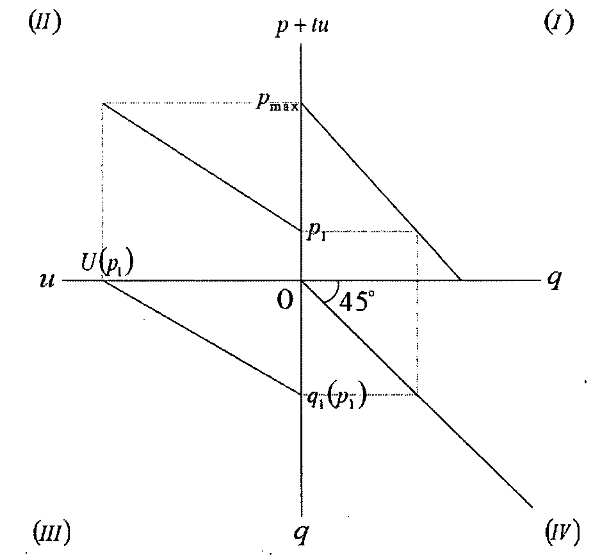


Figure 2-13. Output q and radius u at price p_1 (Referred to Parr, 2002b: 35)

Figure 2-14 (below) illustrates how the demand cone shifts through changes in output price:

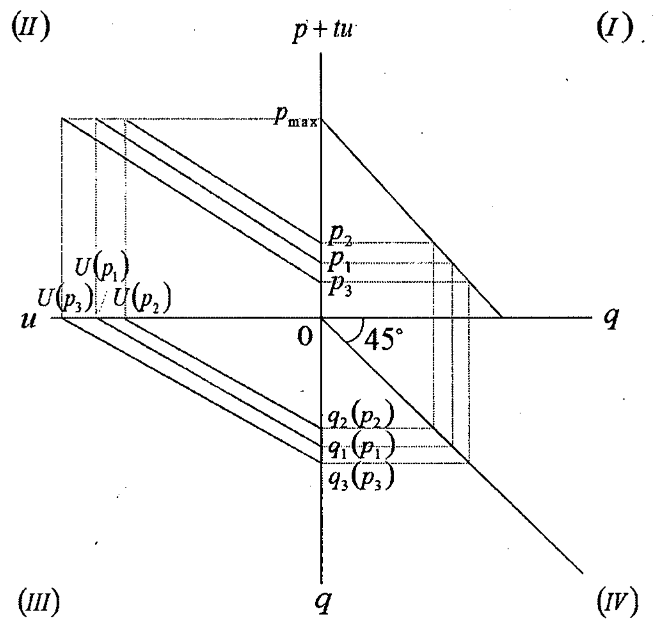


Figure 2-14. Output q and radius u at various prices (Referred to Parr, 2002b: 35)

This case shows how the demand cone is affected by changes in output price p. As shown in the above diagram, the demand cone shifts parallel when the output price level changes. As a result, the maximum market area radius U has relations to the output price p and output level q as:

$$\frac{\partial U}{\partial p} < 0 \qquad \textbf{(2-46)}$$

$$\frac{U_i(p_i)}{q_i(p_i)} = \frac{U_j(p_j)}{q_j(p_j)} \qquad i = 1, \ldots, n; \; j = 1, \ldots, m; \; i \neq j \qquad \textbf{(2-47)}$$

The above Equation (2-47) represents a discriminant if the demand cone is changed by parallel or by slope.

Figure 2-15 (below) depicts how the demand cone shifts according to the change in transportation rate. This case can be seen when technical improvement or regression is observed on the distribution transportation system.

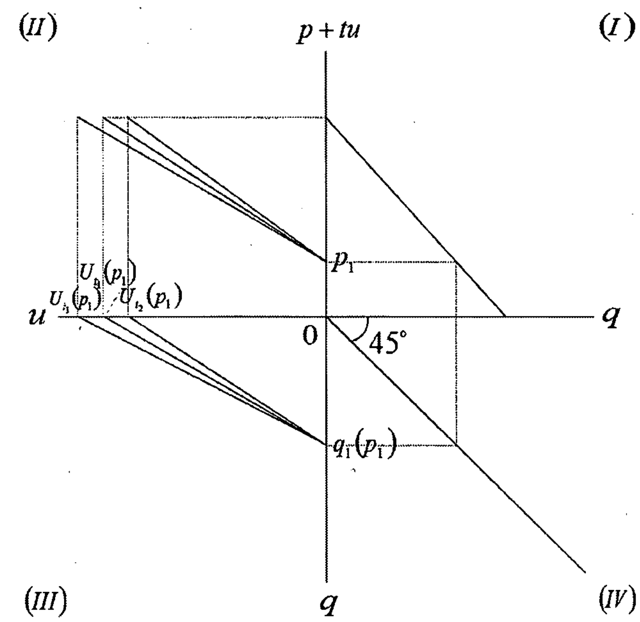


Figure 2-15. Output q and radius u at various transportation rates (Referred to Parr, 2002b: 35)

In this case, as shown in the above diagram, a change in distribution transportation rate t affects the slope of demand cone holding base point

$q_1(p_1)$. As a result, the maximum market-area radius U has relations to the distribution transportation rate t, output price p and output q as:

$$\frac{\partial U}{\partial t} < 0 \qquad (2\text{-}48)$$

$$\frac{U_i(p_i)}{q_i(p_i)} \neq \frac{U_j(p_i)}{q_i(p_i)} \qquad i, j = 1,\ldots,n\,;\; i \neq j \qquad (2\text{-}49)$$

Figure 2-16 (below) shows how the demand cone shifts through changes in the demand curve:

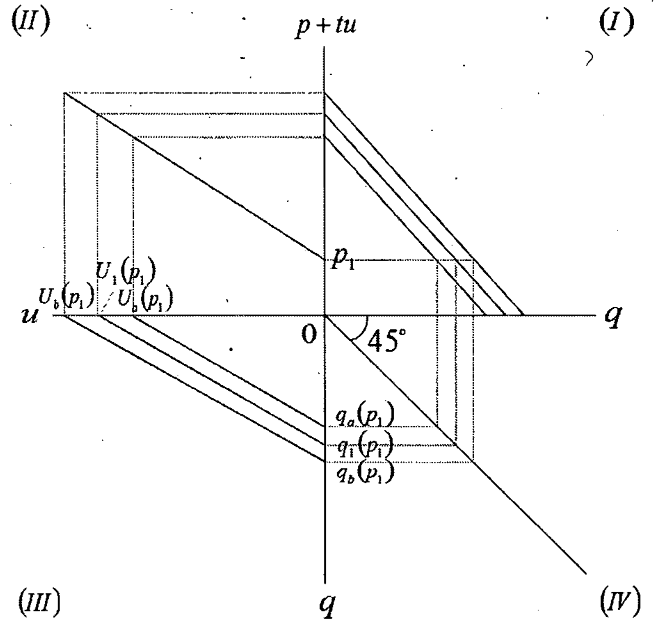


Figure 2-16. Output q and radius u at parallel shifts of demand curve (Referred to Parr, 2002b: 35)

As shown in the above diagram, a parallel shift in the demand curve causes a parallel shift in the demand cone. As a result, the maximum market-area radius U has relations to the average revenue AR, output price p and output q as:

$$\frac{\partial U}{\partial AR} > 0 \qquad \text{in terms of parallel shifts} \qquad (2\text{-}50)$$

$$\frac{U_i(p_i)}{q_i(p_i)} = \frac{U_j(p_i)}{q_j(p_i)} \qquad i = 1, \ldots, n; \, j = 1, \ldots, m; \; i \neq j \quad \textbf{(2-51)}$$

Figure 2-17 (below) illustrates how the demand cone shifts according to changes in the slope of demand curve. This can be seen when consumers' goods-preference changes.

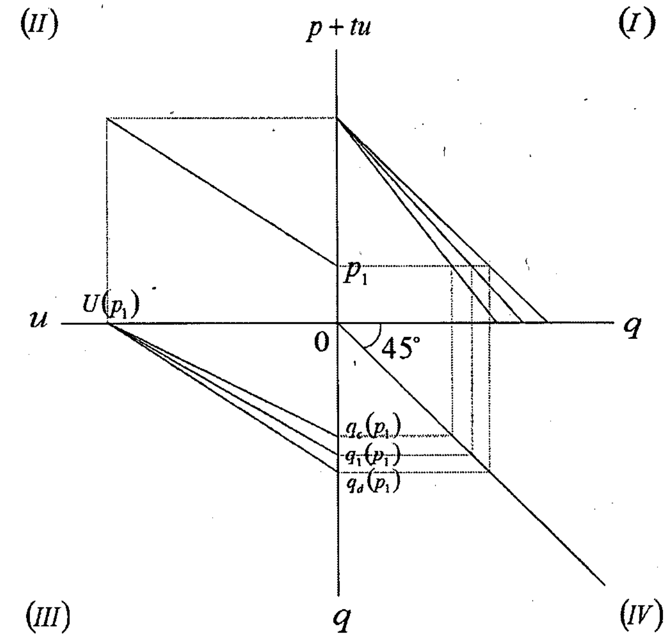


Figure 2-17. Output q and radius u at several slopes of the demand curve

(Referred to Parr, 2002b: 35)

As shown in the above diagram, a change of slope in the demand curve causes the demand cone to shift, holding the maximum market-area radius constant. As a result, the maximum market-area radius U has relations to the average revenue AR, output price p and output level q as:

$$\frac{\partial U}{\partial AR} = 0 \qquad \text{in terms of slope shifts} \quad \textbf{(2-52)}$$

$$\frac{U_i(p_i)}{q_i(p_i)} \neq \frac{U_i(p_i)}{q_j(p_i)} \qquad i = 1, \ldots, n; \, j = 1, \ldots, m; \; i \neq j \quad \textbf{(2-53)}$$

To sum up, it becomes clear that the demand cone can be changed by either a change in the output price, the transportation rate or the shape of the demand curve. In addition, two variables, namely quantity of output q and market-area radius u, are both functions of the output price p. The volume of the demand cone is given as the aggregate demand Q_1 at price p_1 as:

$$Q_1 = D\int_0^{2\pi}\left[\int_0^{U(p_1)} f(p_1 + tu)udu\right]d\theta \qquad (2\text{-}54)$$

where D represents population density and U shows the maximum market-area radius. The formulation (2-43) in Mills and Lav (1964) has a more simplified form, however this alternative formulation (2-54) maintains a more proper formation in terms of implicit price function.

2.8. The Limitations of Market Areas

In this chapter, the analysis of market areas in Lösch (1954), the economic law of market areas, the law of retail gravitation and the spatial competition of market areas have been studied. This section will examine the limitations of these analyses of market areas.

2.8.1. Non Market-Oriented Economic Activity

Market-area analysis is applicable where the relevant economic activity is market-oriented. Market-oriented goods and services form a centre of distribution. The centre of the market area is a point of distribution and production of these goods and services, which are sensitive to consumer behaviour. These goods and services include various ranges of manufacturing, retailing and financial services. Well-known examples of manufacturing include flour milling, bakeries and milk bottling plants. These products are consumed by households and directly are affected by the relevant demand curve of consumers. Based on the relevant demand curve and price setting levels, the market area is formed, and the optimal size, shape and number of firms are derived. There are several other orientations which are not market-oriented. These are raw-material oriented, port oriented and energy-saving oriented cases. Representative examples of these cases include: coal-mining for raw-material orientation, automobile assembly for port-orientation, and the steel industry for energy orientation. These are less sensitive to consumers than market-oriented products. In addition, market areas are not defined, as distribution of these outputs tends not to rely on market orientation. Furthermore, industries of semi-assembled goods such as automobile assembly and the electronics industries have particular contracts with downstream firms, and supply sites can be indicated as a set of points but not an area. In this way, it is clear that market-area analysis has certain limitations when the industry is not market oriented.

2.8.2. *Spatial Exclusivity and Product Differentiation*

Market-area analysis assumes spatial exclusivity of market areas. The assumption of spatial exclusivity implies that an area is not shared by any others located outside of that area. This assumption is sustained as long as identical products are uniformly distributed over the plain. If there is product differentiation, market areas can no longer be exclusive to individual consumers' choice. In addition, the uniform distribution pattern can be changed due to the appearance of portions of overlapping areas. Furthermore, this product differentiation also allows price discrimination such as with Bertrand price competition. The Bertrand price model results in negative profit for all relevant firms if the commodities are assumed to homogeneous. However, this equilibrium becomes different if the commodities are product differentiated. The alternative equilibrium can be found in conventional economic analysis by the Bertrand-Nash equilibrium. In this way, the existing market-area analysis is required in order to modify the spatial competition model if the conditions of spatial exclusivity and product differentiation are changed.

2.8.3. *Inclusion of Multi-Stage Production Process*

There is a case that goods and services are not purchased from households, but from manufacturing firms as semi-assembled goods. In this case, suppliers expand their market area from the centre of distribution according to their demand conditions. There is also the opposite point of view belonging to the purchasers. They observe from the other side a supply area with respect to the given conditions of their production scale. As a result, there are two different types of areas, a market area for the supplier A and a supply area for the purchaser B as shown in Figure 2-18 (below).

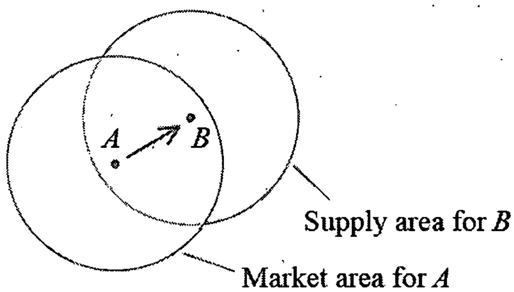


Figure 2-18. Multi-stage production process

In the situation where multi-stage production processes exists, the analysis of the market area also needs to analyse supply area conditions due to the fact that the relevant demand condition is generated by the condition of the supply areas of purchasers. The existing static market-area analysis has difficulties in managing these types of dynamic framework. Under the existing analysis, market and supply areas are examined independently and the relationship between these two different approaches has not been fully investigated.

2.8.4. *The Presence of Administrative and Spatial Constraints*

It is apparent that almost every economic activity encounters administrative and geographical constraints. Market-area analysis assumes that the entire space is opened to all individuals. If the entire space is not opened evenly and is divided by national boundaries or other geographical conditions, the assumption of continuity cannot survive and market-area analysis will have limitations in finding the optimal spatial structure. Administrative and geographical conditions cause these additional constraints to spatial economic activity with respect to the availability of land. As a result, it is necessary for firms to take into account spatial restrictions as an additional cost burden of the use of land. As they do so, they can either reduce the amount of product output in order to adjust their production scale to the feasible output levels, or increases the amount of exports to other areas in order to maintain the original level of production scale. In the former case, a reduction of the production level causes costs to increase if the production processing is taken under the economies of scale. In the latter case, an increase in exports may cause costs to increase as export duties and other transaction costs to access areas beyond the boundary are incurred. The additional consideration of restricted land use affects the assumption of the uniform cost structure of market-area analysis. The use of land is one of the various factors of administrative and spatial constraints. As Lösch (1954) states, the economic equilibrium is not always the equilibrium of nature, and proper ideas of laws should be enacted in order to adjust the equilibrium level between the economic and natural standpoint of view. In order to evaluate the equilibrium of market areas properly, these absent factors should be included in the analysis. However, this cannot be demonstrated in a straightforward manner and this is one of the limitations of market-area analysis.

2.8.5. Uneven Distributions of Inputs

Market-area analysis assumes that inputs are evenly distributed and avoids the requirement to consider the problems associated with the distant transportation of inputs. However, there are many cases where inputs are dispersed on the plain and therefore the consideration of distant transportation is required. The assumption of evenly distributed input can be sustained if relevant transportation costs of inputs are negligible. However, if the transportation rate reaches a high level, this assumption cannot be survived because uneven distribution inputs cause certain cost changes as shown in Figure 2-19 (below). In the diagram, TrC_L represents a lower transportation-rate curve, TrC_H shows a higher transportation-rate curve, and d expresses distance. It is obvious that the higher transportation rate is more sensitive to increased cost of shipping.

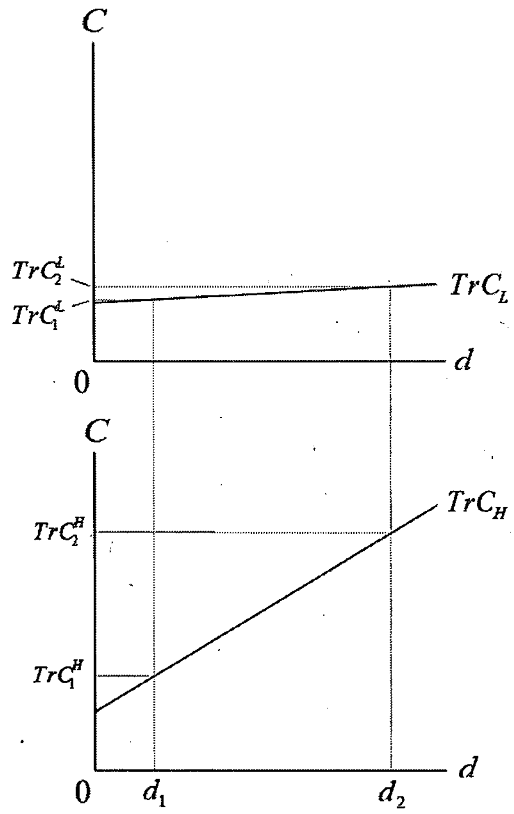


Figure 2-19. Lower and higher transportation rates

As illustrated in the above diagram, if the transportation rate is set at a low level, the cost differential between distances d_1 and d_2 is only the amount of $TrC_2^L - TrC_1^L$. However, the same circumstance for a higher level of transportation rate causes the considerable level of cost change $TrC_2^H - TrC_1^H$. In this way, higher levels of transportation rates may certainly influence the outcome of location problems. The examination of location and transportation costs requires the investigation of agglomeration economies as these have a trade-off interaction with transportation costs. Although agglomeration economies have been examined in location theory, market-area analysis cannot directly introduce these economies, owing to the fact that they are observed in the analysis of a production process which market-area analysis treats as a given fixed economic factor, but not as a dependent variable. As a result, these further approaches are beyond the scope of the analysis of market areas.

2.9. The Limitations of Independent Analysis of Market Areas

The previous section examines the limitations of market-area analysis. This section will indicate an extensive framework of the existing market-area analysis. While market areas have been investigated in depth for decades, the sequence of production process and production inputs has been treated as a given constant factor in the analysis of the market areas for the purpose of theoretical simplification. However, these examined economic factors should be integrated in order to analyse the spatial economic structure more precisely. One of the difficulties concerns the formation of production function as its function is defined by the related inputs of the producer. In other words, the supply areas of their production should also be investigated in order to derive the production function. Moreover, during the production process, there are certain influences of economies relating to the producer's production scale. Unless the argument contains these additional economic elements, further detailed analysis of market areas will not be achieved. The following chapters will investigate these economic factors of market areas and all the required elements will be combined in later chapters.

2.10. Conclusion

This chapter first examines Lösch (1954), commenting on some points which are not explicit in his analysis and laying the foundations by which they can be extended in later chapters. Second, the economic law of market areas and the law of retail gravitation, alternative approaches to the analysis of market areas, are introduced. In addition, the theoretical similarities and differences of these two laws are examined. Third, spatial competition and organisation of market areas are analysed. Finally, the theoretical limitations of market-area analysis are discussed and the related implication regarding the necessity of re-examining the analysis is made. Regarding these further extensions, a statement in the argument of location decision in Lösch (1954) should be re-examined; namely that the location of farm buildings and fields with the consideration of accessibility to the centre of villages can be solved by taking into account the relationship between the optimal firm location and market areas. This will be further analysed in later chapters.

Chapter 3. Supply-Area Analysis

Supply-area analysis investigates how individual firms obtain their inputs, such as raw materials and labour, in order to achieve the optimal levels of production under given spatial economic conditions. In pure economic theory, these inputs commonly refer to raw materials, labour and capital. By contrast, the element of capital has a more complex structure in location theory as it involves the investigation of inter-regional or international trade. Thus, it is generally excluded from the analysis or assumed to constant for reasons of simplicity. This chapter will first attempt an overview of supply areas. Second, it will refer to Parr (1993a; 1993b) and Parr and Swales (1996; 1999), who have investigated equilibrium level of inputs and supply-area configurations under certain conditions of spatial competition for products. Third, the properties of raw materials in the framework of supply-area analysis will be detailed with respect to the limited availability of inputs. In addition, the attributes of assembly costs will be examined in terms of several possible types of transportation system. Finally, the limitations of supply-area analysis will be analysed, followed by an examination of the limitation of independent supply-area analysis.

3.1. An Overview of Supply-Area Analysis

This section will introduce an overview of supply-area analysis. While market-area analysis investigates the relationship between output and relevant market areas under the given conditions of demand and spatial competition, supply-area analysis examines the relationship between input and relevant suppliers under the given conditions of factor prices and assembly cost. While some literature also refers supply-area analysis to the production process itself, the analysis of the production process should be examined through production function. In location theory, production function is not sufficiently detailed and this results in the internal and external economies being insufficiently included. These aspects will be further examined later in this chapter.

Supply areas were initially investigated in a systematic way by Lösch (1938). He noted that geographical and cultural regions are artificial units without full economic relevance. This indicates that it is necessary, when investigating real spatial structure, to consider given geological conditions, administrative spatial allocation and other artificial or natural circumstances. Lösch derives the concept of supply areas from his analysis of the nature of economic regions. In order to illustrate actual economic circumstances, he generates individual supply areas for a given line of production, and permits the existence of empty areas. Figure 3-1 (below) illustrates these situations for bakeries, cotton gins and coal mines.

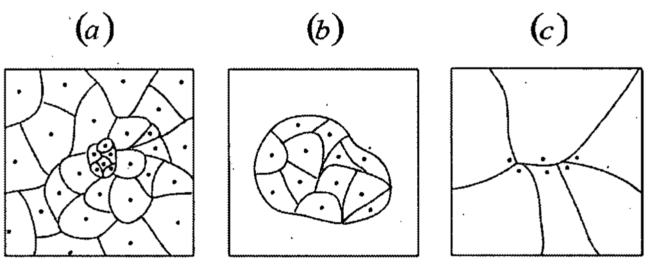


Figure 3-1. The nets of areas (Source: Lösch, 1938)

In the above figure, diagram (a) does not have any constraint of area, and diagram (c) shows single points of supply which does not refer to an area. The net limited case, diagram (b), represents an area of supply, suggesting the example of cotton gins. This example can be extended as follows. At the industrial location of cotton gins, raw cotton is used as input for producing textile goods. As the suppliers of raw cotton are dispersed across the plain, the required amount of raw cotton is collected radially from the location of the production plant. Although the extent of this radius is the exact concept of the supply area, Lösch did not extend this further in his analysis.

Beckmann (1968) has indicated the applicability of input-output analysis to the framework of market areas and supply areas. However, further investigation has not been attempted within this framework. Instead, Beckmann examines supply areas with respect to plant location as determined by the relationship between available economies of scale and transportation costs in addition, to obtain optimal supply-area size. This also takes into account a relative ratio of weight and bulk between inputs and output. He additionally considers the advantage of spatial proximity to the market or rural location.

3.2. Producer Supply Area within a Region under Spatial Competition

Parr (1993a) examines producers who are engaging productions within a region and compete for access to a dispersed commodity input. The producer supply area is investigated for a particular manufacturing activity through a free-entry model and the long-run equilibrium. The following assumptions are considered in his analysis. First, there is a bounded region and the commodity is uniformly dispersed. This commodity is used as a unit of factor for producing a unit of product. The factor is ubiquitous and has constant returns to scale. Second, producers pay transportation costs from suppliers to the production locations. There are economies of scale for this production. The uniform unit shipping cost of the product to the external region is paid by producers. Producers are price takers and face a constant price for the product. Finally, there are no internal and external economies. The relevant cost and revenue curves are illustrated in Figure 3-2 (below).

In the diagram, AC^A shows the average assembly cost curve and this includes uniform rate of transportation cost from the suppliers to the manufacturing plant. The horizontal curve AC^B represents the average commodity cost and AC^D depicts the average delivery cost. The U-shaped curve $ACAC$ represents the average total cost which is the combination of vertically added curves AC^A, AC^B and AC^D. This curve $ACAC$ is also called the planning curve for a certain scale of circular spatial configuration. The relevant marginal cost curve is illustrated as

$MCMC$. The curves AR_e and MR_e represent the average revenue and marginal revenue for each producer and these are equivalent to price level p.

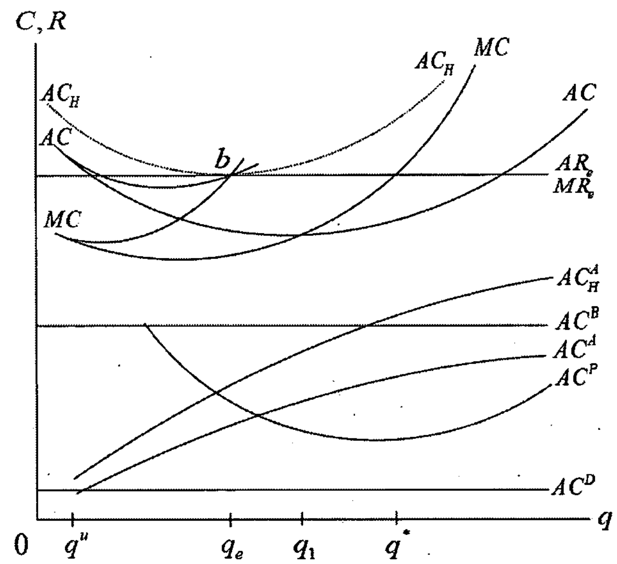


Figure 3-2. Individual cost and revenue curves (Source: Parr, 1993a)

Parr first demonstrates a non-competitive model of a producer within a region. In this case, this producer is a monopolist and the optimal quantity of output is determined at q^* where marginal cost equals marginal revenue for maximising his profit. Second, the model is extended to the long-run competitive equilibrium. Under conditions of an exogenously-determined product price, the average cost curve level increases due to the appearance of new entrants. However, location analysis has a different movement. As the accessibility to the commodity is restricted, spatial configuration cannot keep a circular shape and changes to a space-filling polygon. The shape eventually forms a regular hexagon. This increases the shape of the average assembly cost curve AC^A to AC_H^A, as the hexagonal shape cannot sustain the minimum cost in the case of the circular shape. Thus, the planning curve for the alternative hexagon becomes $AC_H AC_H$ and the equilibrium scale becomes q_e where the relevant marginal cost $AC_H b$ equals marginal revenue AR_e. The actual planning curve becomes ACb

and its marginal cost is MCb, as the spatial formation is circular until the production scale exceeds the inscribed circle of the hexagon. Figure 3-3 (below) represents a brief idea of three different types of planning curve.

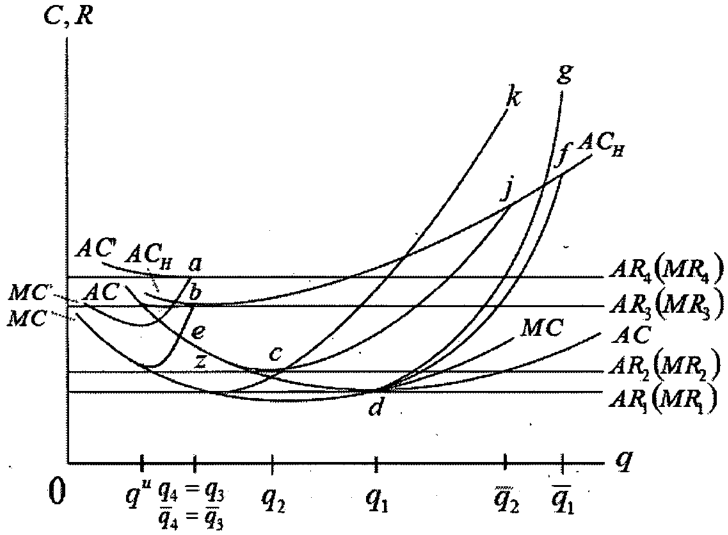


Figure 3-3. The optimal scale of outputs with different price levels (Source: Parr, 1993a)

At scale q_e, profit becomes zero due to the competition, and the relevant supply-area size L_e is q_e / μ. The modifying process of the planning curve due to spatial competition of supply areas can be explained by the movement from $ACAC$ to ACb. Although Parr did not fully compare this process with aspatial competition, the difference will be shown in Figure 3-4 (below). While spatial planning curve $ACAC$ changes to ACb, and the optimal scale of output q^* shifts to q_e in the free-entry spatial competition model, the curve moves to AC_2AC_2 and the optimal scale becomes q_2 as illustrated in the diagram. In the aspatial case, the movement of the curve $ACAC$ to AC_2AC_2 takes an upward and left-hand side direction, if the original minimum average cost is less than the average revenue. This can be clarified by noting that the additional entry of producers continues up to the output level q_2 and the reduced optimal scale of production causes less efficient production. This effect is observed in the left-hand side movement of the long-run average cost curve, in addition to the upward movement by the increased number of competitors.

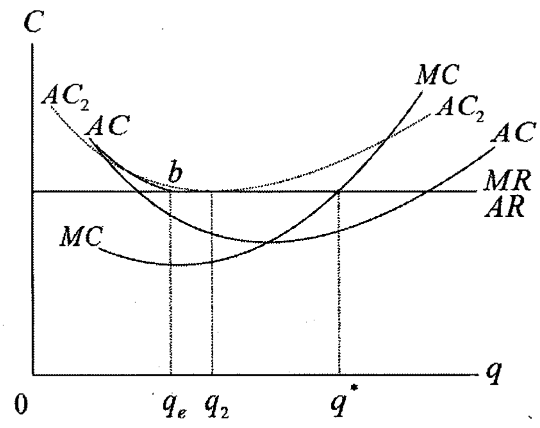


Figure 3-4. Spatial planning curve and aspatial average cost curve

Parr further extends the investigation to a more general analysis with respect to four different levels of price. The initial case is the same as the previous space-filling hexagonal model. In Figure 3-3 (shown earlier), this shows a price level $p = AR_3 = AR_e$ and the equilibrium scale q_3. The maximum feasible scale of output $\overline{q}_3$ is q_3 and the supply area is $\overline{L}_3$. The second case is $p = AR_4$, more generally expressed as $p > AR_3$. In this case, there exists a rent to each producer and the average commodity cost becomes $AC^B + AR_4 - AR_3$. This increases the relevant planning curve AC_Hb up to $AC'a$ and the maximum feasible scale of output is $\overline{q}_4$ which is equal to the equilibrium level q_4. The spatial configuration is hexagonal and the equilibrium supply-area size becomes L_4 which is equal to the maximum size $\overline{L}_4$. As the economies of scale achieve a minimum at this level, rents will be kept by each producer. The third case is $p = AR_1$ which has the equilibrium scale of output q_1 and supply-area size L_1. The relevant planning curve becomes $ACdf$ and the feasible maximum scale of output and supply-area size will be $\overline{q}_1$ and $\overline{L}_1$ respectively. However, producers maintain the smaller levels of output and supply-area size in order to take advantage of cost minimisation through maximising economies of scale. As a result, spatial configuration becomes circular and there exist areas of supply-area exclusion between these circles. The final case is $p = AR_2$ or more generally $AR_1 < p < AR_3$. In the case of $p = AR_2$, the equilibrium scale of output and supply-area size are q_2 and L_2, respectively. As the relevant planning curve is $ACcj$, the maximum feasible scale of output and supply-area size are $\overline{q}_2$ and $\overline{L}_2$. As with the previous case, producers keep smaller levels at q_2 and L_2 for having maximum economies of scale. In this case, the spatial configuration becomes something between circular and hexagonal; in other words, it becomes a truncated circle. Parr defines the deviation of the equilibrium scale of output from the maximised economies of scale and specifies the technical inefficiency in terms of price conditions: $p - AR_1$ for $AR_3 \geq p > AR_1$ and $AR_3 - AR_1$ for $p > AR_3$.

Three types of spatial configuration are finally compared and contrasted. In order to examine the minimum distance to the supply-area boundary n_p, Parr first defines the maximum feasible supply-area size $\overline{L}_p$. This is expressed by the ratio of the maximum feasible scale of output $\overline{q}_p$ and industrial output per square kilometre of the supply area μ.

$$\overline{L}_p = \frac{\overline{q}_p}{\mu} \qquad (3\text{-}1)$$

For the hexagonal case, the maximum feasible supply-area size is represented as the following expression:

$$\overline{L}_p = \frac{6n_p^2}{\sqrt{3}} \qquad (3\text{-}2)$$

The minimum distance to the supply-area boundary at price p as n_p can be solved with respect to the maximum feasible scale of output $\overline{q}_p$ and industrial output per square kilometre of supply area μ.

$$n_p = \frac{3^{\frac{1}{4}}\sqrt{\overline{q}_p}}{\sqrt{6\mu}} \qquad (3\text{-}3)$$

Similarly, the maximum distance to the supply-area boundary at price p as x_p is:

$$x_p = \frac{3^{\frac{1}{4}}\sqrt{\overline{q}_p}}{\sqrt{6\mu}\cos\theta_p} \qquad (3\text{-}4)$$

The length of any straight-line portion of a supply-area boundary at price p as t_p is:

$$t_p = \frac{2\left(3^{\frac{1}{4}}\sqrt{\overline{q}_p}\tan\theta_p\right)}{\sqrt{6\mu}} \qquad (3\text{-}5)$$

Parr also demonstrates the derivations of the equilibrium frequency of producers at price p as N_p with the notation of the extent of region ℓ:

$$N_p = \frac{\ell}{\overline{L}_p} \qquad (3\text{-}6)$$

or

$$N_p = \frac{\ell\mu}{\overline{q}_p} \qquad (3\text{-}7)$$

The equilibrium spacing between any pair of neighbouring producers at price p as s_p is twice the value of the minimum distance to the supply-area boundary at price p. As a result, this can be expressed as:

$$s_p = 2n_p = \frac{2\left(3^{\frac{1}{4}}\sqrt{\overline{q}_p}\right)}{\sqrt{6\mu}} \qquad (3\text{-}8)$$

In order to express the above equation with respect not to $\overline{q}_p$ but q_p, Parr clarifies the relationship between the equilibrium output at price p as q_p

and the maximum feasible scale of output as $\overline{q}_p$ through the expression of the supply-area inflator. The supply-area inflator I_p is defined as:

$$I_p = \frac{\overline{L}_p}{L_p} \qquad (3\text{-}9)$$

Applying the description of $\overline{L}_p = \overline{q}_p / \mu$ and $L_p = q_p / \mu$ to the above expression, $\overline{q}_p$ can be solved as:

$$\overline{q}_p = q_p I_p \qquad (3\text{-}10)$$

Parr finally shows the relationship between supply-area inflator I_p and spatial configuration as:

$$I_p = \frac{\cos^2\theta_p \tan 30}{\cos\theta_p \sin\theta_p + \frac{\pi(30-\theta_p)}{180}} \qquad (3\text{-}11)$$

3.3. Supply Area of a Single-Plant Producer

Parr (1993b) investigates the position of the single multi-plant producer who maximises economic profit within the region of operation. The following assumptions are considered in his analysis. First, a particular raw material is required for the production within a region. The raw material is used as a unit of factor for producing a unit of output. This implies that the input-output ratio is the same at all levels of production. In addition, the raw materials are uniformly distributed and no rent for the raw material is present from the production. Factor price has constant returns to scale and the producer needs to pay a uniform transportation rate for obtaining the raw material. Regarding the transportation, extra-regional shipping of the output is also paid by the producer. This also has a uniform transportation rate. Finally, the producer behaves as a price taker. The plant output q is denoted with supply-area size per square kilometre $L(q)$ and the output per square kilometre of the supply area as:

$$q = \mu L(q) \qquad (3\text{-}12)$$

As examined in the previous section, Parr analyses three types of supply-area configuration, namely circular, hexagonal and truncated circular configurations. Average assembly cost for supply-area configuration θ as A_θ is defined with the notations of the freight rate per tonne-kilometre τ and the amount of raw material required to produce one unit of output λ, in addition to the above shown symbols:

$$A_\theta = \frac{q\lambda\tau s_\theta(q)}{q} = \lambda\tau s_\theta(q) \qquad 0 \le \theta \le 30 \qquad (3\text{-}13)$$

In the above expression, $s_\theta(q)$ represents the mean distance between plant and all suppliers for processing q. The circular $s_c(q)$, hexagonal $s_h(q)$ and truncated circular $s_t(q)$ cases are expressed as follows.

$$s_c(q) = \frac{2}{3}r \qquad \theta = 0 \qquad (3\text{-}14)$$

$$s_h(q) = \frac{b}{3} + \frac{b\log\sqrt{3}}{2} \qquad \theta = 30 \qquad (3\text{-}15)$$

$$s_t(q) = \left\{\frac{c}{3} + \frac{c^2 - d^2}{3d}\log\left[\frac{c+d}{\sqrt{c^2 - d^2}}\right]\right\}w_\theta + \frac{2c(1-w_\theta)}{3} \qquad 0 < \theta < 30$$

$$(3\text{-}16)$$

where w_θ $(0 < w_\theta < 1)$ represents:

$$w_\theta = \frac{\cos\theta\sin\theta}{\cos\theta\sin\theta + \frac{\pi(30-\theta)}{180}} \qquad (3\text{-}17)$$

From the above results, it becomes clear that the average assembly cost $A_\theta(q)$ has a square-root form with a constant variable k_θ as $k_\theta\sqrt{q}$ and is illustrated in Figure 3-5 (below).

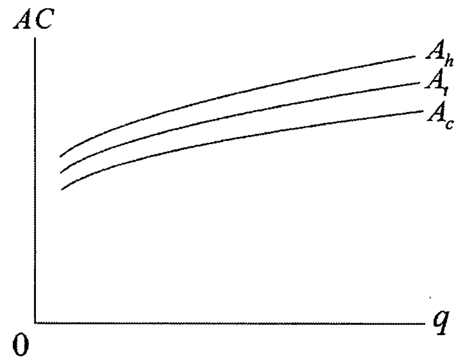


Figure 3-5. Average assembly cost curves (Source: Parr 1993b)

Other costs are factor price, distribution cost of output and other relevant costs. These are all independent of supply-area configuration and are integrated as a processing cost. The processing cost is U-shaped and is added vertically to the average assembly cost. The number of supply area or manufacturing plants having a scale of q as $N_\theta(q)$ is expressed with the extent of the region ρ and the supply-area size of each plant $L(q)$.

$$N_\theta(q) = \frac{\rho}{L(q)I_\theta} \tag{3-18}$$

or

$$N_\theta(q) = \frac{\rho\mu}{qI_\theta} \tag{3-19}$$

The multi-plant producer is assumed to maximise the level of profit Π throughout the region. The analysis initially examines the individual-plant level as in Figure 3-6 (below). In the diagram, the curve $C_\theta C_\theta$ represents the long-run average cost curve for circular, truncated and hexagonal spatial configurations. In addition, average revenue is depicted by the horizontal line a which also represents marginal revenue when $p = a$. Region-wide profit Π_θ is expressed as the following equation, with marginal revenue p, the scale of an individual manufacturing plant q, plant average cost at q as $C_\theta(q)$ and the number of plants at q as $N_\theta(q)$.

$$\Pi_\theta = [p - C_\theta(q)]qN_\theta(q) \tag{3-20}$$

where p, ρ and μ are all constants. Thus, the profit-maximising output at $\hat{q}_\theta$ and average cost $\hat{C}_\theta$ is achieved where $p - C_\theta(q)$ becomes highest. It should be noted that this maximum region-wide profit is interpreted as the maximum profit per unit of output, as the extent of the region and density of input are limited.

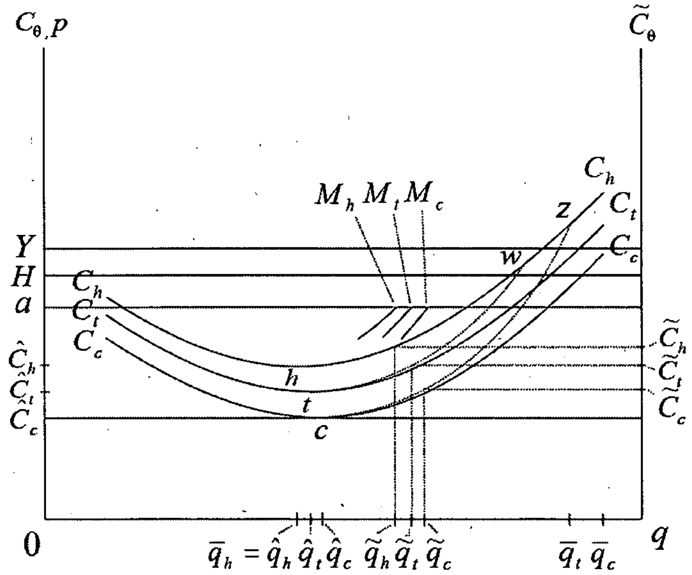


Figure 3-6. The plant average cost curves (Source: Parr 1993b)

Under the condition of horizontal average revenue curve, the profit maximisation is achieved where average cost becomes minimum. As each supply-area configuration has different attributes of average assembly cost, the following relations to other configurations are obtained.

$$\hat{C}_c < \hat{C}_t < \hat{C}_h \qquad (3\text{-}21)$$

$$\hat{q}_c > \hat{q}_t > \hat{q}_h \qquad (3\text{-}22)$$

$$\hat{N}_c < \hat{N}_t < \hat{N}_h \qquad (3\text{-}23)$$

If the extent of the region is large enough to cover the size of the supply area, further analysis can be examined. The profit function can be re-examined as a function in the optimal situation.

$$\Pi_\theta = \left(p - \hat{C}_\theta\right)\hat{q}_\theta \hat{N}_\theta \qquad (3\text{-}24)$$

The above expression shows that the optimal spatial configuration θ^* is a determinant factor for specifying the maximum level of region-wide profit. In addition, it is clear that there is a trade-off interaction between minimisation of average cost and maximisation of the number of plants. As the amount of raw material has a limitation at $\overline{q}_\theta$, the curve $C_\theta C_\theta$ is required to be modified as the plant scale and frequency are determined. The alternative circular long-run cost becomes $C_c h$, the truncated circular case is $C_t w$ and the hexagonal case becomes $C_h t$, as illustrated in the above diagram.

In the diagram, $\tilde{q}_\theta$ and $\tilde{c}_\theta$ represent the feasible production and the relevant cost under each configuration of the monopoly profit-maximising level. The multi-plant total-cost function for configuration θ as T_θ is now derived by the following expression in Figure 3-7 (below).

$$T_\theta = C_\theta(q) q \hat{N}_\theta \qquad 0 \le q \le \overline{q}_\theta \qquad (3\text{-}25)$$

In the diagram, each total cost function is illustrated as the solid-line part of the curve. This ends at the multi-plant output $\hat{Q}_\theta$ where $\hat{Q}_\theta = \hat{q}_\theta \hat{N}_\theta$ and the number of plants $\hat{N}_\theta$ are operating at the scale $0 \le q < \overline{q}_\theta$.

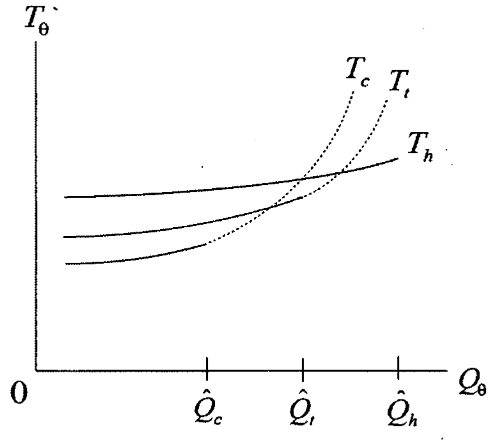


Figure 3-7. Multi-plant total cost curves (source: Parr, 1993b)

The relation to each configuration is summarised as:

$$\hat{Q}_c < \hat{Q}_t < \hat{Q}_h \quad (3\text{-}26)$$

$$\hat{q}_c\hat{N}_c < \hat{q}_t\hat{N}_t < \hat{q}_h\hat{N}_h \quad (3\text{-}27)$$

Figure 3-8 (below) illustrates multi-plant total cost and revenue curves for each spatial configuration. The curve S represents an envelope curve of each end of the solid-line part in the above diagram and shows the minimum multi-plant total cost ratio at each spatial configuration.

$$S = \hat{C}_\theta\hat{q}_\theta\hat{N}_\theta = \hat{C}_\theta\hat{Q}_\theta \quad (3\text{-}28)$$

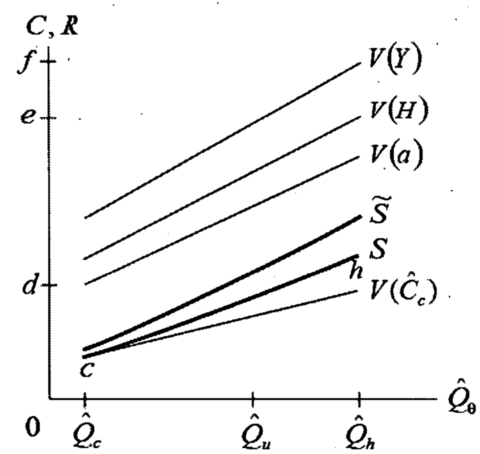


Figure 3-8. Multi-plant total cost and revenue (Source: Parr, 1993b)

In the above diagram, the curve $V(p)$ represents a trans-configurational multi-plant total revenue function.

$$V(p) = p\hat{q}_\theta\hat{N}_\theta = p\hat{Q}_\theta \quad (3\text{-}29)$$

If price level is $p = a$ as shown in Figure 3-6 (shown earlier), the total revenue curve becomes $V(a)$ as illustrated in Figure 3-8 (above). In the above diagram, the optimal multi-plant output Q^* is achieved where the vertical distance between $V(p)$ and S is the largest. This optimal level should also correspond with the optimal spatial configuration θ^* in order to specify the optimal plant scale q^*, supply-area size L^*, plant average cost C^* and the number of plants N^*. In order to specify the optimal multi-plant output Q^*, the given price level $p = a$ is required to investigate with Figure 3-6 (shown earlier). In this diagram, the spatial monopoly profit-maximising scale is expressed as $\tilde{q}_\theta$ and the number of plants $\tilde{N}_\theta$ enables the derivation of the envelope curve $\tilde{S}$ as follows.

$$\tilde{S} = \tilde{C}\tilde{q}_\theta \tilde{N}_\theta = \tilde{C}_\theta \tilde{Q}_\theta \quad (3\text{-}30)$$

It is clear that this envelope curve $\tilde{S}$ does not achieve the region-wide profit maximisation, as the original envelope curve S is lower than $\tilde{S}$ at any levels.

Finally, the influence of exogenously determined price on the spatial structure is examined. If price is $p = \hat{C}_c$ as shown earlier in Figure 3-5, the corresponding multi-plant output is $\hat{Q}_c$ where $S = V(\hat{C}_c)$ in Figure 3-8 (above). Although Parr (1993b) did not state that the maximum separation between S and $V(\hat{C}_c)$ is the output level $\hat{Q}_h$, this maximisation implies that the negative profit is minimised. As a result, the maximised maximum-negative-profit level $\hat{Q}_c$ is the optimal outcome in this circumstance. The relevant supply-area configuration is circular and each plant produces the optimal scale of output $q^* = \hat{q}_c$ under the optimal cost $C^* = \hat{C}_c$. In contrast with other cases, this case achieves monopoly profit maximisation where plant marginal cost equals plant marginal revenue. This spatial case is equivalent to the free-entry competitive model in the previous section. The second case is where price is $\hat{C}_c < p < H$. As examined in the case where $p = a$, the multi-plant profit maximisation is achieved at $\hat{Q}_u$ where $dS = dV(a)$ in Figure 3-8 (above). As a result, the spatial configuration is a truncated circular and each plant produces $q^* = \hat{q}_u$ at the optimal cost level $C^* = \hat{C}_u$. The final case is where price level is $p = h$ in Figure 3-7 (above). In this case, the optimal multi-plant output is $\hat{Q}_h$ and the spatial configuration is hexagonal. The relevant plant scale of output is $q^* = \hat{q}_h$ at the optimal cost level $C^* = \hat{C}_h$. This equilibrium is sustained where $p > H$, as shown in the case $p = Y$. The only difference is the vertical distance between multi-plant total revenue and total cost, unless the condition $dS < dV(Y)$ is changed. In the diagram, the height $f - e$ shows this situation and this can be treated as a rent.

3.4. Industry Cost Curves under Unified Control

Parr and Swales (1996) demonstrate that an industry equilibrium long-run average cost curve for a particular class of economic activity for a given region, has a non-horizontal shape due to the presence of externalities. This curve is used to derive a long-run equilibrium for a regional industry under the condition of unified control. The following assumptions are considered in addition to providing the demand conditions in the analysis. First, a regional industry produces a manufactured good for an extraregional market. The producer uses a raw material for producing a manufactured good and the ratio between them is fixed. The raw material is available at a fixed uniform density and the factor price is fixed. Second, the producer pays an assembly transportation cost which is constant throughout the region. In addition, he also pays a distribution transportation cost to the external region which has a uniform rate per tonne-kilometre. Finally, the production has increased returns to scale for value-added processing.

Parr and Swales initially examine the structure of the cost curve at the plant level. There are two types of curve, which are assembly and non-assembly costs. The former depends on the spatial configuration of the supply area. The latter does not vary with the supply-area configuration, and raw-material acquisition, production of value-added goods and final goods delivery are considered. At the plant level, the long-run average cost $c_\theta(q)$ is with constants β and κ, parameter α_θ and plant scale q.

$$c_\theta(q) = \alpha_\theta \sqrt{q} + \frac{\beta}{\sqrt{q}} + \kappa \tag{3-31}$$

From the above expression, the minimum average cost for any given θ is derived by taking the first derivative.

$$\hat{q}_\theta = \frac{\beta}{\alpha_\theta} \tag{3-32}$$

The value of the minimum average cost $\hat{c}_\theta$ for the supply-area configuration θ is obtained from the above two equations.

$$\hat{c}_\theta(q) = 2\sqrt{\alpha_\theta \beta} + \kappa \tag{3-33}$$

In order to generalise the properties of each variable, the following comparative-static results are provided.

$$\frac{\partial c_\theta}{\partial \theta} = \frac{\partial c_\theta}{\partial \alpha_\theta}\frac{\partial \alpha_\theta}{\partial \theta} = \sqrt{q}\frac{\partial \alpha_\theta}{\partial \theta} > 0 \tag{3-34}$$

$$\frac{\partial \hat{c}_\theta}{\partial \theta} = \frac{\partial \hat{c}_\theta}{\partial \alpha_\theta}\frac{\partial \alpha_\theta}{\partial \theta} = \sqrt{\frac{\beta}{\alpha_\theta}}\frac{\partial \alpha_\theta}{\partial \theta} > 0 \tag{3-35}$$

$$\frac{\partial \hat{q}_\theta}{\partial \theta} = \frac{\partial \hat{q}_\theta}{\partial \alpha_\theta}\frac{\partial \alpha_\theta}{\partial \theta} = -\frac{\beta}{\alpha_\theta^2}\frac{\partial \alpha_\theta}{\partial \theta} < 0 \tag{3-36}$$

The geometric diagram is illustrated in Figure 3-9 (below).

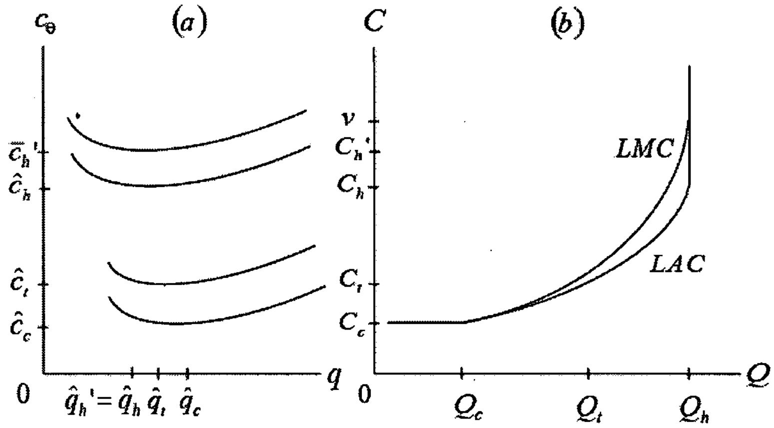


Figure 3-9. Plant and industry cost curves (Source: Parr and Swales, 1996)

The left-hand side of the diagram shows the plant cost and scale which is also provided in the previous section. The right-hand side of the diagram shows the relevant industry cost and output. This is also shown in *Phase* (I) in Figure 3-10 (below).

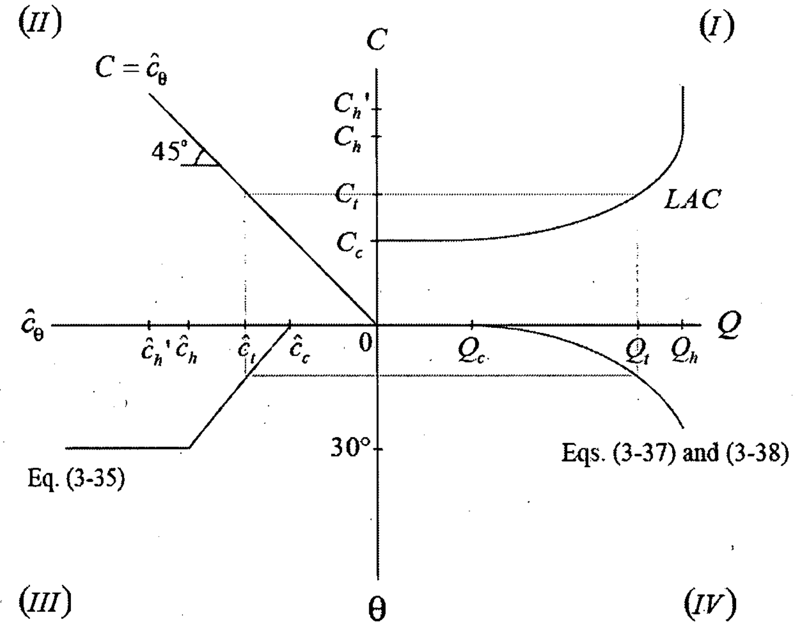


Figure 3-10. Four-dimensional diagram (Source: Parr and Swales, 1996)

The long-run average cost curve for industry in *Phase* (I) in the above diagram is derived from the following two steps. The first is from the relationship between industry output Q and supply-area configuration θ in *Phase* (IV). This includes the extent of land utilisation E_θ, which is the proportion of $RSTV$ and RSU in Figure 3-11 (below).

$$E_\theta = \frac{\tan\theta + \dfrac{\pi(30-\theta)}{180\cos^2\theta}}{\tan 30} \qquad \frac{dE_\theta}{d\theta} > 0 \qquad \textbf{(3-37)}$$

The expression can be restated with the relationship between Q and θ for tightly packed supply areas with the area of the region ρ and manufactured output per square-kilometre of the supply area as μ.

$$Q_\theta = E_\theta \rho \mu \qquad \textbf{(3-38)}$$

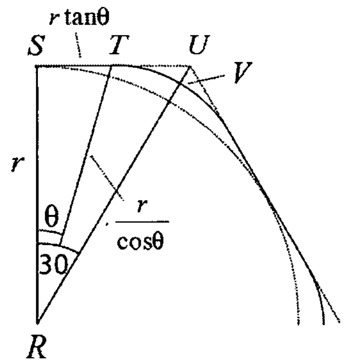


Figure 3-11. Alternative supply-area configurations (Source: Parr and Swales, 1996)

The second factor for deriving a long-run average cost curve for industry is the relationship between supply-area configuration θ and the value of minimum average cost $\hat{c}_\theta$ in Equation (3-33) and illustrated in *Phase* (III) in Figure 3-10 (above). The remaining part of diagram, *Phase* (II), shows the relationship between plant and industry costs which are simply reflected by the angle of 45° at any level of production. The long-run average cost curve in *Phase* (I) is divided into three parts at Q_c and Q_h. The first part is a horizontal line up to industry output Q_c, which has a larger number of plants at $\hat{q}_c$ and $\hat{c}_c$ at each plant level. The second part is the upward-sloping portion between Q_c and Q_h. During this part, there is a negative externality of the technological type. The third part is the vertical section at Q_h. This vertical line implies that there is a rent above cost level C_h to raw-material suppliers or to the owners of their land which is included in κ in Equation (3-32). An example of the rent is the level C_h' in Figure 3-9 and Figure 3-10 (shown earlier). This is referred to as a negative externality of pecuniary type for the industry. In terms of negative externalities, the following condition can be expressed.

$$Q_c < \text{negative technological externality} < Q_h < \text{negative pecuniary externality} \tag{3-39}$$

There are two other spatial structural aspects, which are plant frequency and spacing. These are specified as follows. The frequency of plants F for the circular configuration is expressed as $Q/\hat{q}_c$. For the truncated-circular case, this becomes:

$$F_\theta = \frac{Q_\theta}{\hat{q}_\theta} \qquad (3\text{-}40)$$

where $\partial F_\theta / \partial\theta > 0$, as $\partial Q_\theta / \partial\theta > 0$ and $\partial \hat{q}_\theta / \partial\theta < 0$. The plant spacing G for the circular configuration is:

$$G = \frac{\sqrt{2\rho\hat{q}_c}}{\sqrt{Q}3^{\frac{1}{4}}} \qquad (3\text{-}41)$$

For any cases between the circular and hexagonal cases, this becomes:

$$G_\theta = \frac{\sqrt{2\rho\hat{q}_\theta}}{\sqrt{Q_\theta}3^{\frac{1}{4}}} \qquad (3\text{-}42)$$

where $\partial G_\theta / \partial\theta < 0$. This shows that the plant spacing is involved in the full adjustment to a changed level of industry. In order to clarify this evidence, Parr and Swales introduced the locatinally-constrained long-run average cost as illustrated in Figure 3-12 (below).

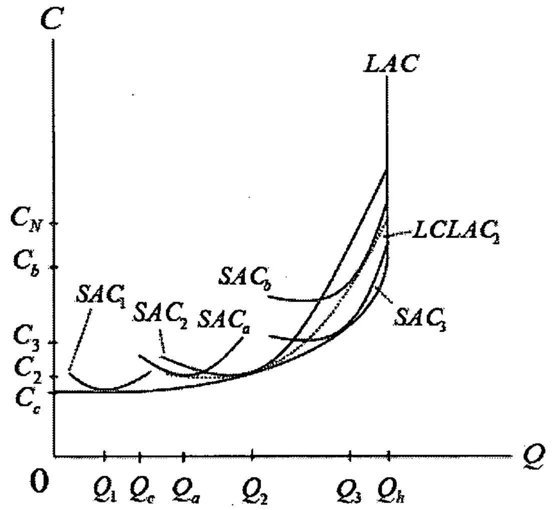


Figure 3-12. The long-run equilibrium adjustment (Source: Parr and Swales, 1996)

In the above diagram, the broken-line curve $LCLAC_2$ represents the locationally-constrained long-run average cost for a particular level of output. Under the condition of the fixed plant spacing, a long-run change in industry output increases average cost levels. This implies that the distance r $(=G_\theta/2)$ in Figure 3-11 (above) is fixed. As a result, the distance $r/\cos\theta$ is a decreasing function of θ. The short-run average cost curves are illustrated as SAC_1, SAC_2 and SAC_3 in the above diagram, and are tangential to the long-run average cost curves at Q_1, Q_2 and Q_3, respectively. The curve SAC_1 solely achieves the minimum short-run average cost at the tangential level to the long-run average cost curve. However, this curve is situated away from the locationally-constrained long-run average cost curve $LCLAC_2$. The curves SAC_a and SAC_b are tangential to the locationally-constrained long-run average cost curve. The full adjustment to LAC and $LCLAC_2$ is achieved at the industrial output Q_3. There are three stages of adjustment from the level Q_2 as follows. First, in the short-run, the curve SAC_2 can only adjust to the industry level Q_3, at the cost level C_N, due to the presence of a fixed factor with respect to plant capital stock and plant spacing. Second, in the long run, the plant capital stock becomes a variable cost, and the curve is changed to SAC_b which is tangential to the curve $LCLAC_2$ and has a lower cost C_b than C_N. Third, over the long run, plant spacing also becomes a variable cost and the curve can shift to SAC_3 which is tangential to the curve $LCLAC_2$ and has a lower cost C_3 than C_b.

Parr and Swales further analyse the long-run industrial equilibrium where the industry is organised in a unified manner. In order to determine the industry equilibrium outcome, the relevant demand conditions for the manufactured good are specified by three different cases. First, price is given as a horizontal line as $p = p_3$, which is shown in Figure 3-13 (below).

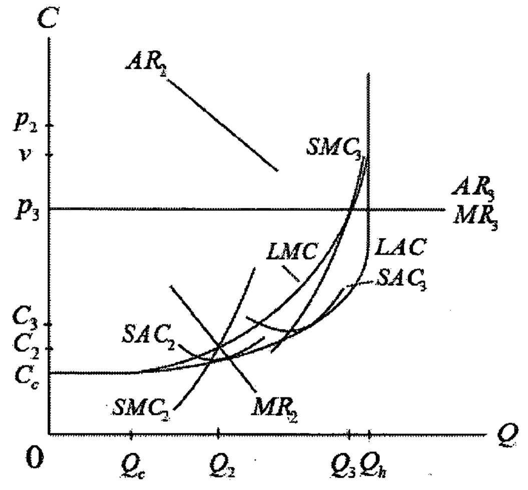


Figure 3-13. Industry equilibrium under certain conditions (Source: Parr and Swales, 1996)

In this case, the optimal industry output is at Q_3, as the marginal revenue MR_3 and marginal cost SMC_3 intersect at this level. The supply-area configuration is a truncated circle, and the profit is expressed as $(p_3 - C_3)Q_3$. This implies that if the price level becomes less than C_c, production will not be operated. Moreover, if the price level is $p = C_c$, production is operated in a zero profit condition and the supply becomes a set of loosely-packed circles. Furthermore, if the price level is $p > v$, the spatial configuration becomes hexagonal and there exists a certain level of rents. Second, price is given by the downward-sloping market demand curve AR_2, which is shown in the above diagram. If the regional industry is assumed to the single source of the supply area, this industry behaves as a monopolist. As a result, the industry output is determined at Q_2 where $SMC_2 = MR_2$ and the price becomes p_2. The supply-area configuration is a truncated circle and profit is expressed as $(p_2 - C_2)Q_2$. The supply-area configuration becomes a set of loosely-packed circles, if the industrial output is below the level Q_c, or hexagonal if the level is Q_h. Third, the regional industry is assumed to be a price maker, as the output is large enough to affect market price. If there are competitors in other regions, this becomes an oligopoly market. By applying the Cournot-Nash equilibrium model with u identical regions and a linear market demand curve $Y = J - jp$, the industry output Q is obtained as follows.

$$Q = \frac{J - jC_c}{u+1} \quad \text{iff} \quad \frac{J}{j} > C_c > \frac{J - 0.9069Q_h(u+1)}{j} \qquad (3\text{-}43)$$

where $Q < Q_c$. The relevant price becomes:

$$p = \frac{J}{j(u+1)} + \frac{uC_c}{u+1} \qquad (3\text{-}44)$$

As $Q < Q_c$, the supply-area configuration is a set of loosely-packed circles.

3.5. Industry Cost Curves under Spatial Competition

Parr and Swales (1999) examine a spatial model of competing firms in a regional industry under conditions of free-entry competition. In this model, there are similar conditions to those assumed in the previous section. Furthermore, the following two conditions are also added. First, competition for the raw material as an input is shipped as an output by each supplier, and is referred to the Nash Equilibrium in an infinitely repeated game. Second, the production of value-added goods at the firm level has a long-run average processing cost which decreases with scale. The analysis initially examines the formation of the industry long-run average cost curve with respect to independent firms. Figure 3-14 (below) shows a hexagonal domain of each uniform-spacing firm. The term r and angle θ $(0 \le \theta \le 30)$ determine the size of the domain and competing firms.

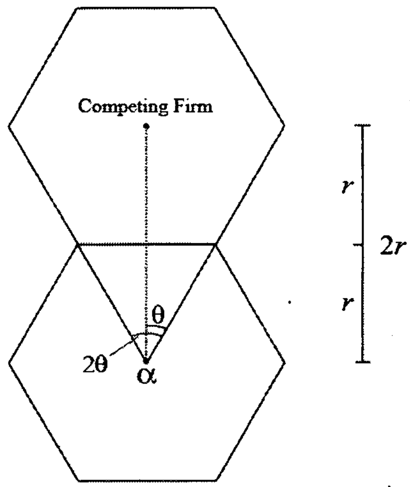


Figure 3-14. The size of the domain and competition

As shown in Figure 3-15 (below), the curve AC has a U -shaped form and this is derived from the combination of the increasing average assembly cost curve AC_A and the decreasing average processing cost curve.

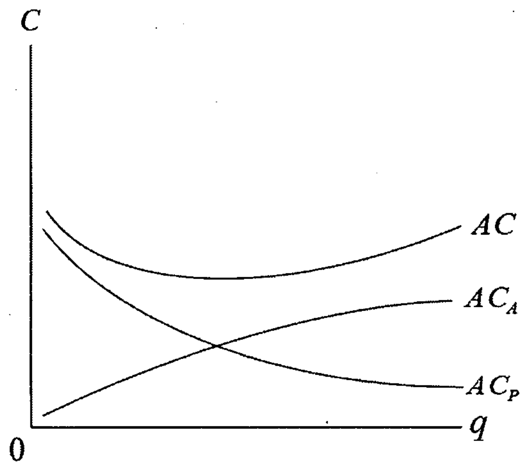


Figure 3-15. Average assembly cost and processing cost

Three spatial configurations, namely circular, truncated circular and hexagonal cases are illustrated in Figure 3-16 (a) (below). As demonstrated by Parr and Swales (1996) in the previous section, the size of domain of each plant is determined by the single decision maker so as to achieve cost minimisation in average cost curves. In order to examine the case under free-entry competition, Parr and Swales (1999) initially investigate technological externalities at the level of the individual firm.

The negative externalities of the technological externality are referred to the combined output of all other firms within the industry in the conventional free-entry competition model. However, the spatial model in this analysis refers to the number of firms in the industry. Above the certain industrial output level, there is a spatial constraint for reducing the size of the domain as the number of firms increases. The increase of the number of firms forces each firm to pay higher costs. As introduced in the previous section, the locationally-constrained long-run average cost curve is drawn in Figure 3-16 (below) in order to observe long-run adjustment under competition.

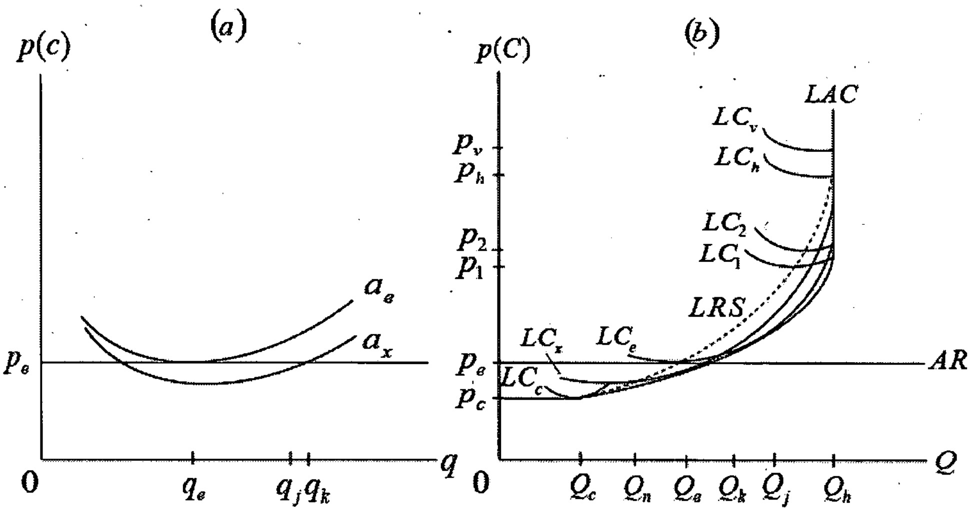


Figure 3-16. Plant and industrial equilibrium outcomes (Source: Parr and Swales, 1999)

In the diagram, the curve LC_x is the horizontal summation of the curve a_x for various firms. This theory applies only to the spatial model for which there is no technological externalities, once firm frequency is fixed. The curve LC_x has an envelope relation to the long-run average cost curve at the industry output level Q_k. At this point, cost minimisation is fully adjusted and the supply-area configuration becomes a truncated circle. A change in this output level has no impact on either the frequency or the spacing of firms. A decline of Q_k reduces the firm scale and supply-area size in addition to forming the more circular shape, while an increase of

Q_k increases them, and the supply-area shape becomes more hexagonal. In the diagram, other LC curves also show the envelope of the long-run average cost curve. However, it should be noted that parts of these curves above LC_1 have no tangency. At any level on the vertical portion of LC, firm scale, size and hexagonal supply-area configuration are all unchanged, as the space is fully filled.

In order to examine the long-run profit-maximising outcome under free-entry competition, the parametric pricing condition of output at the firm level is initially considered. Given the industry demand curve AR is produced as a horizontal line p_e, the locationally-constrained long-run average cost curve must be tangential to the line p_e to achieve profit maximisation in this case. However, if the locationally-constrained long-run average cost curve is below the level of the curve a_e, there exists positive profit and the new entrants appear. This increases the industry curve up to LC_e. At this level, the equilibrium industry output is Q_e which also achieves the minimum level of locationally-constrained long-run average cost. The tangential point of the curve LC_e to the long-run average cost LAC is the industry output level Q_j and the scale of firm is q_j. However, this creates a higher frequency of firms and higher cost of production than Q_e. If the price level exceeds the level p_h, for instance at the level p_v, there exists a rent, $p_v - p_h$ per unit of industry output.

In order to examine the efficiency considerations, Parr and Swales apply a downward-sloping industry demand curve to the analysis, which the regional industry has a significant role to the extraregional product market. As shown in Figure 3-17 (below), welfare maximisation is achieved where average revenue curve AR intersects long-run marginal cost curve LMC. In this circumstance, the community welfare becomes the area $adgu$ which is the summation of the consumers' surplus adz and producers' surplus $zdgu$. In comparison with the aspatial competitive conditions in terms of the technological externalities, the community welfare *adgu* is larger than the aspatial case *afv*. According to Worcester (1969), the formation of the aspatial context of externalities is observed by the upward-sloping curve LAC, which coincides with the industry long-run supply curve in the diagram.

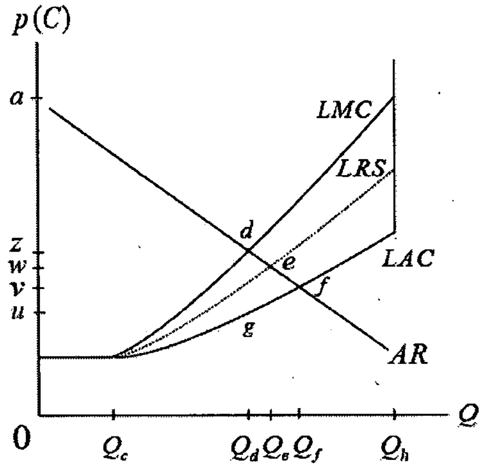


Figure 3-17. Efficiency configuration of the spatial equilibrium (Source: Parr and Swales, 1999)

In this case, the competitive equilibrium is at industry output Q_f and price level v. As the community welfare is solely the part of consumer surplus *afv*, the spatial unified control condition at price $p = LMC = AR$ has a larger impact on community welfare. Furthermore, if the spatial competition is assumed instead of the spatial unified control, community welfare becomes the area *aew* at industry output Q_e and price level w, and this is smaller by area of *wefv* than in the aspatial case.

3.6. Properties of Raw Materials

This section will examine the limitation of raw materials in terms of industry cost curves under unified control and spatial competition. While the limited availability of raw materials has not been investigated in the existing location theory, there is a certain limitation of usage for raw materials with respect to the available amount of deposits of raw materials. In order to avoid the difficulties of dealing with raw materials, the existing literature assumes that the available amount of raw materials is larger than the maximum amount for firm production within a territorial location. As examined earlier, Parr (1993a; 1993b) and Parr and Swales (1996; 1999) assume that the amount of the obtainable raw material is large enough to satisfy the required amount of each firm. If this condition is not assumed, the industry supply curve of raw materials can be drawn as the curve S in Figure 3-18 (below) with respect to the maximum required amount of production $Q_{\max}$.

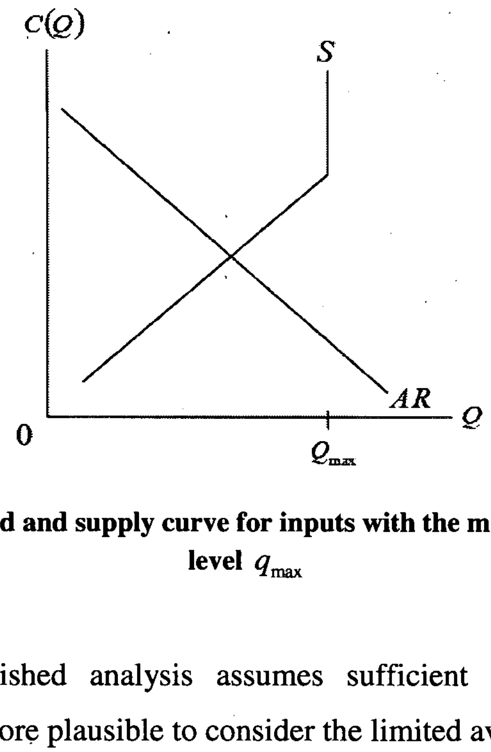


Figure 3-18. Demand and supply curve for inputs with the maximum production level $q_{\max}$

While the established analysis assumes sufficient capacity of input acquisition, it is more plausible to consider the limited availability of inputs. As shown in Figure 3-19 (below), if the feasible level is assumed to achieve up to the quantity of industry output Q_r, the firm requires either a reduction of the supply-area size or imports from other regions to make up the additional amount of input. If the case is the welfare-maximisation situation, the additional amount of input will be $Q_d - Q_r$. If the situation is competitive conditions under the aspatial technological externality type, the amount becomes $Q_e - Q_r$. Finally, the situation under spatial competition will be $Q_f - Q_r$. These results show that the less availability of raw materials, the higher cost burden is imposed to the producer.

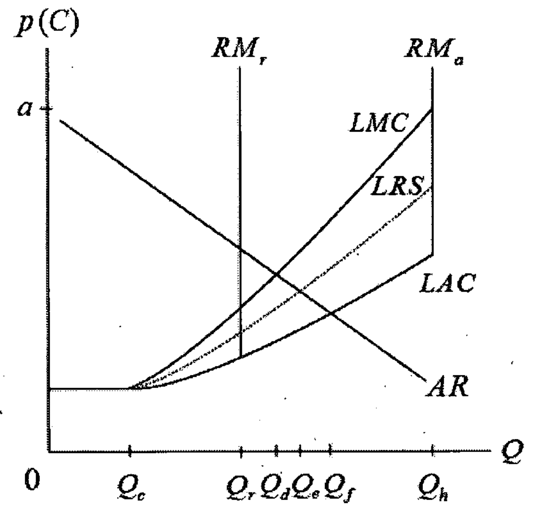


Figure 3-19. Demand conditions and raw material availability (Refer to Parr and Swales, 1999)

In terms of the adjustment of supply-area size, the following examination can be provided. Under the conditions of the given plant long-run average cost curve and demand curve in spatially monopoly model, the optimal production is achieved at the output level q_h as shown in Figure 3-20 (below). If the given demand curve is AR_1, there is too large a supply-area size and the firm decides either to decrease its supply-area size or export to other regions for residual output. By contrast, if the given demand curve is AR_3, there is too small a supply-area size and the firm decides either to increase its supply-area size or import from other regions for the additional required quantity of inputs. In any cases, each firm is faced with the trade-off interaction between sustaining certain profit levels under the optimal level of production, and extra charges for assembly cost by trade with outside distant regions. As the extent of these effects relies on the shape of the demand curve, it is difficult to specify the optimal size of the supply area unless the condition of market areas is taken into account.

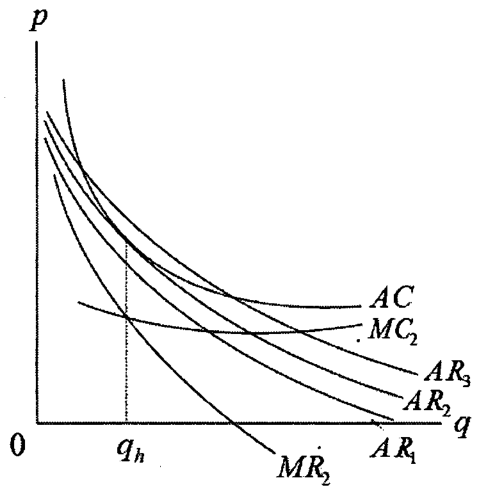


Figure 3-20. Shifts in demand curves

3.7. Further Perspective on Assembly Cost

This section will investigate the structure of assembly cost in terms of transportation attributes between a deposit site of input and the production plant. In market-area analysis, transportation costs are examined within the framework of distribution costs which are the shipping costs between production plant and market area. In supply-area analysis, these costs are examined in the framework of assembly costs which are the shipping costs between suppliers of input and the production plant. The notion of the assembly cost in location theory involves transportation costs for inputs such as raw material and labour. For raw materials, the weight, value and distance between the deposit site and the production plant are the important economic factors. For labour, on the other hand, commuting costs between labours' residences and the production plant is an important factor, as is demonstrated in Kohlhase and Ohta (1989). However, it should be noted that individual decision making regarding the residential location of labour also includes other factors such as the location's proximity to schools, preference of living in a quiet zone, well established infrastructure and so on. These can be referred to the urbanisation type of agglomeration economy detailed in Chapter 5.

This argument will extend the framework of assembly cost. First, the analysis will focus on the framework of raw materials and assembly cost. Average assembly cost AC_A can be expressed in the following form:

$$AC_A = kq^a \qquad 0 < a < 1 \qquad (3\text{-}45)$$

where $a =$ a parameter, $k =$ constant and $q =$ quantity of output. Using the above equation, total assembly cost TC_A becomes:

$$TC_A = kq^{1+a} \qquad 1 < 1 + a < 2 \qquad (3\text{-}46)$$

The marginal assembly cost MC_A will be:

$$MC_A = \frac{\partial TC_A}{\partial q} = (1+a)kq^a \qquad 0 < a < 1 \qquad (3\text{-}47)$$

These expressions can be illustrated in Figure 3-21 (below).

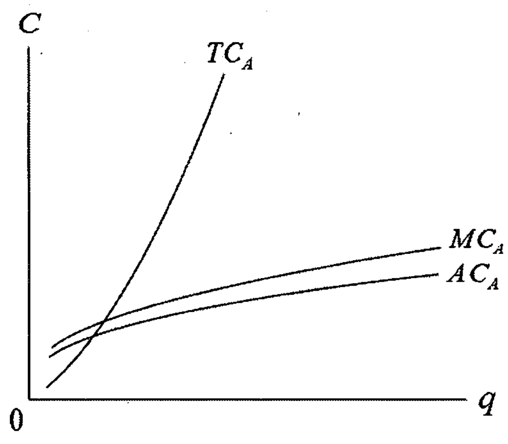


Figure 3-21. Average, marginal and total assembly costs

From these results, it is clear that the total assembly cost has increasing returns to scale and the marginal assembly cost has decreasing returns to scale. This is similar to the average cost except for the form of $(1+a)$. These results demonstrate that the higher the power of the variable a, the steeper the total and marginal cost curves become. As a result, the total assembly cost curve of the hexagonal spatial configuration TC_A^h becomes steeper than that of the truncated circular configuration TC_A^t, which in turn is steeper than that of the circular spatial configuration TC_A^c. This is shown in Figure 3-22 (below).

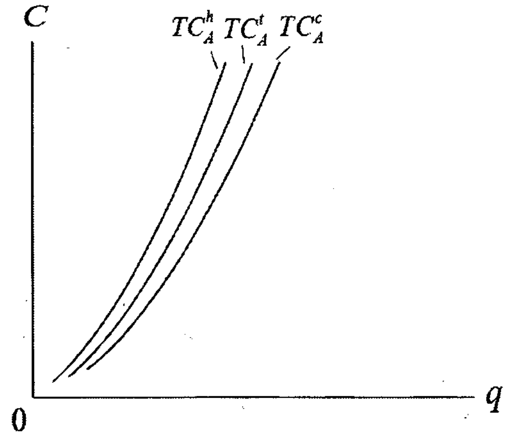


Figure 3-22. Total assembly cost curves of three types of spatial configurations

The simplest case for the observation of the assembly cost is to treat the relevant space as a linear function. If the transportation cost has the property of constant returns to scale, a two-dimensional diagram between cost and distance of input is shown, as in the curve T_rC_L in Figure 3-23 (below). If the curve has increasing returns to scale, the curve becomes T_rC_I. If the freight rate is a multiplier function and has decreasing returns to scale, the curve is shaped as T_rC_D. Also in this diagram, the curves ATC_L, ATC_I and ATC_D are average total costs. These are the sums of average cost AC and each relevant transportation cost curve T_rC_L, T_rC_I and T_rC_D.

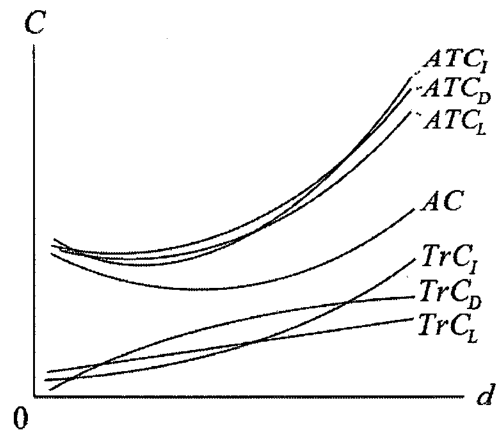


Figure 3-23. Average cost and transportation cost curves

These costs are all categorised as variable costs in economic analysis. There are also other elements of cost - namely fixed costs with respect to transportation from the initial unit of transportation - which should differ from the fixed cost for production. This is called terminal cost F_τ in location analysis. As shown in Figure 3-24 (below), the cost level $F + F_\tau$ is referred to fixed costs and above this level is considered as variable costs. This diagram has two different shipping methods which have the curves T_rC_I and T_rC_D regardless of the price level of inputs. In this case, if the firm has a choice between these two transportation methods, the firm chooses T_rC_I up to the distance d_b and T_rC_D for distances longer than d_b in accordance with cost minimising behaviour.

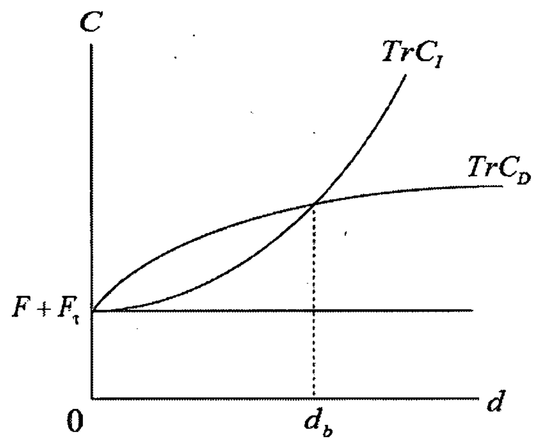


Figure 3-24. Fixed and variable cost for transportation

The following complex cases can also be analysed. First, there is the case in which two or more stage structures of transportation cost exist. In the case of Figure 3-25 (a) (below), there are two stages and the boundary is the point d_b. Figure 3-25 (b) (below) is observed when each stage has a different freight rate. Furthermore, there are three stages which characterise this case: the first stage has an increasing freight rate up to d_1, the second stage has a constant freight rate up to the point d_2, and the third stage has a decreasing freight rate. These discrete attributes at every boundary are caused by changes of transportation methods or a transportation system based on zoning districts.

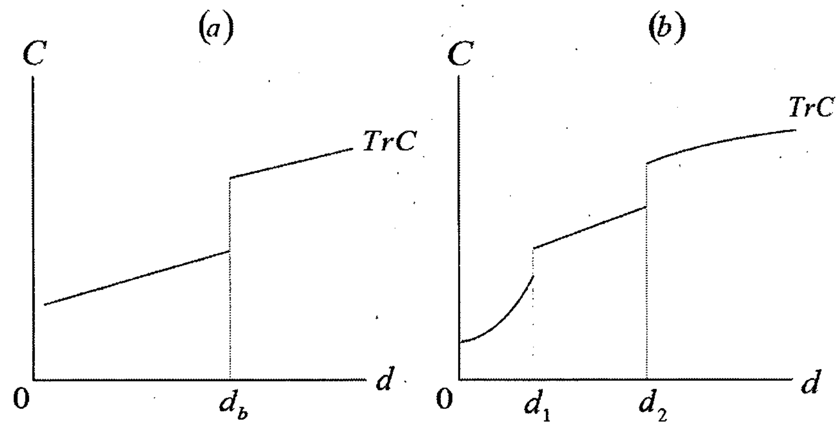


Figure 3-25. Discrete transportation cost curves

As stated earlier, two types of inputs are considered in this analysis, namely raw material and labour. For raw materials, the value and weight have important economic factors apart from the travel distance. As a result, the following four types of pattern can be observed: high value and high weight TrC_1, high value and low weight TrC_2, low value and high weight TrC_3, and low value and low weight patterns TrC_4. Or these four types of curves can also be categorised as high terminal and high line-haul costs, high terminal and low line-haul costs, low terminal and high line-haul costs and low terminal and low line-haul costs. These different structures depend on the proportion of fixed and variable costs of transportation, freight rate differences and the shapes of the transportation cost curves as illustrated in Figure 3-26 (below).

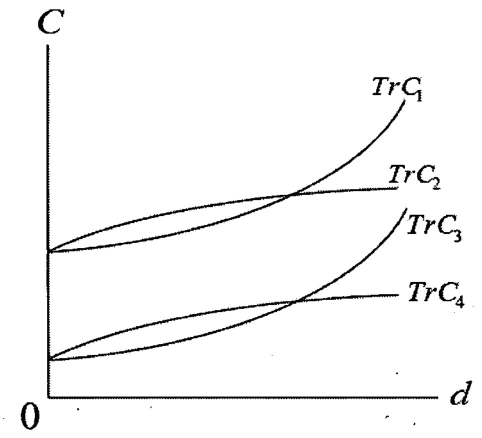


Figure 3-26. Four different types of transportation costs

For labour, on the other hand, the relevant transportation costs are interpreted as a commuting cost. The commuting cost is observable at a constant rate of distance, an increasing rate or a decreasing rate. However, it is more plausible to take into account zoning transportation systems as introduced in many metropolitan areas. These types of transportation cost are illustrated as in Figure 3-27 (below).

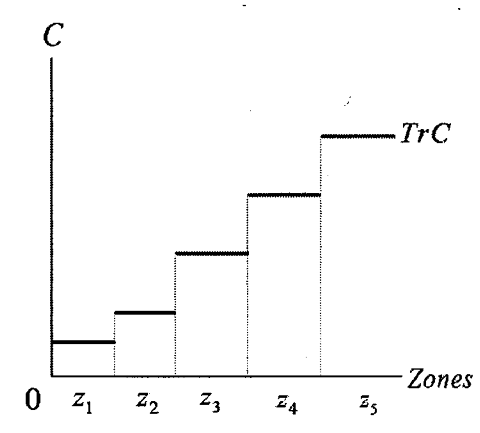


Figure 3-27. Transportation cost with zone-boundary

In location theory, the more realistic observation is to examine the deposit of inputs in terms of an area or plain. This approach can be further examined with respect to the density of inputs, shape and size of supply area (circular, triangular, square, truncated circular or hexagon), unavailable portion of the area, and competition of input with other firms considering price discrimination or bargaining. In this case, the assembly cost curve may be shown as in Figure 3-28 (a) (below). A more complex case is shown in Figure 3-28 (b) (below).

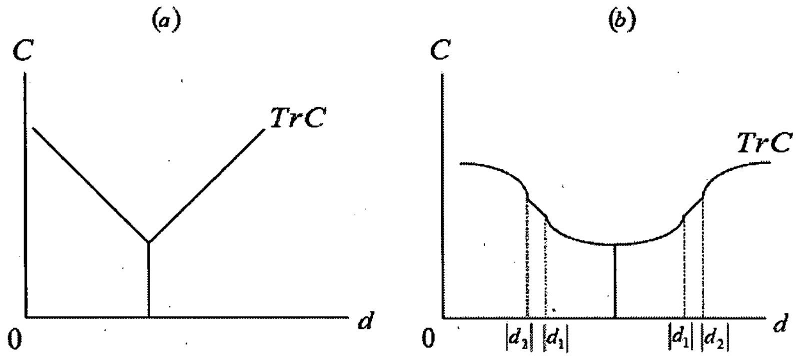


Figure 3-28. Transportation costs in area analysis

Finally, it should be noted that the effect of changes in assembly cost on the production scale is related to the pricing system on the relevant market. The relationship between freight rate and size of supply areas is shown as follows. Under the condition $p = MC$, an increase of freight rate causes a reduction in the size of supply areas through a reduction of the optimal production scale. This is shown in Figure 3-29 (below).

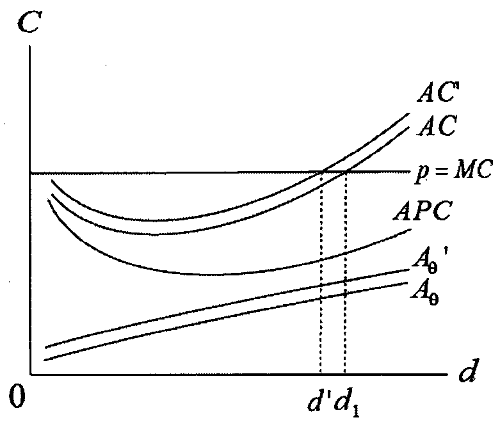


Figure 3-29. Assembly and additional transportation costs in supply-area analysis

$(p = MC)$

The above figure shows the following three points: that an increase of transportation rate moves the average assembly cost curve A_θ to A_θ'; that the shift of the average assembly cost curve moves the average cost curve AC to AC'; and that the upward shift of average cost reduces production scale and distance d_1 to d'. By contrast, under the condition of $p = AC$, an increase of freight rate inevitably causes negative profit. The firm cannot survive in the industry as no adjustment to either the increase or the decrease of production scale can recover profit. This is shown in Figure 3-30 (below). The figure shows that an increase of transportation rate moves the average assembly cost curve A_θ to A_θ' and this shift moves the average cost curve AC to AC'. Due to the price setting condition of $p = AC$, there is no feasible production for firms which satisfies $p \geq AC$ at this cost level.

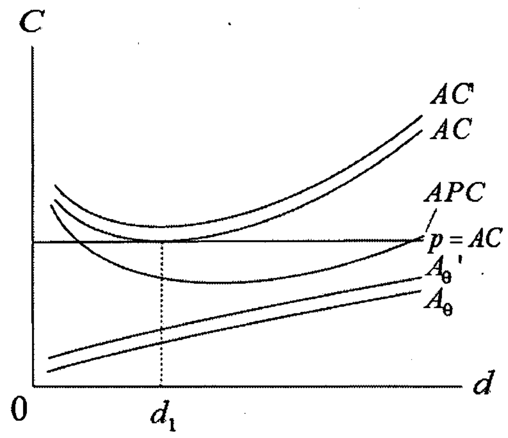


Figure 3-30. Assembly and additional transportation costs in supply-area analysis $(p = AC)$

3.8. The Limitations of Supply-Area Analysis

Supply-area analysis investigates the optimal production scale by considering the relationship between assembly cost and relevant output conditions under the certain assumption of demand. However, the analysis is applicable solely within limited types of economic circumstances. This section will introduce these limitations with respect to a single point of supply, deposit constraint of inputs, contract and negotiation transactions between firms, administrative and geographical conditions and the structure of production function.

3.8.1. Single Points of Supply

There are two broad types of spatial patterns of supply. One is a supply area and the other is a single point of supply. In general, each supply area has a centre as a production plant and inputs are collected within an area of supply. Examples include dairy products, local newspapers and labour sources. They have a model pattern of demand concentration and supply dispersion. However, there is another pattern where supply is not dispersed and inputs are collected from several specific points of supply. This is particularly the case when these inputs are semi-assembled products. These types of concentration of demand are related to advantageous economic factors such as labour concentration. Spatial concentration occurs based on the preferred orientation of firms. First, the raw-material orientation appears if inputs are heavier, bulkier or more perishable than the output. Second, labour orientation can be seen where labour force is crucially important for firms. Finally, market orientation is preferred if products are heavier, bulkier or more perishable than the input and severe *f.o.b.* price competitions exist. In addition, higher levels of land costs should also be taken into account as a factor which can negatively affect the market-orientation. These orientations can be observed in automobile assembly, electronics industries and bottling plants. Each supplier of these industries becomes a set of single points and no area is formed. While the analysis of a single point of supply can solve the optimal output level of each plant, the relevant size and shape of supply areas and the frequency of plants cannot be found, as the condition of continuity is dropped from the assumption.

3.8.2. Availability of Inputs

Supply-area analysis assumes that inputs are ubiquitous across the economic plain. This assumption is one of the necessary conditions of exclusivity of supply areas. Unless inputs are ubiquitous, supply areas which do not have enough share of deposit of input become disadvantageous in terms of transportation cost burdens relating to assembly cost. There are many cases in which inputs are unevenly distributed. In these cases, supply areas can overlap and the exclusivity condition cannot be maintained. If the transportation rate for inputs is set at a high level, the optimal production scale of distant firms becomes smaller and input acquisition availability will be more limited. In this case, firms which are located close to a deposit site of input will have an advantage. As supply-area analysis requires a uniform spatial formation, these types of spatial differentiation analysis cannot be further examined.

3.8.3. Contract and Negotiation Transactions between Firms

It is assumed in supply-area analysis that an independent relationship exists between a purchaser who is located at the production plant, and the supplier. If a purchaser and supplier are not related to each other, their trade is carried out through price mechanisms formed by economic factors which include the input price, optimal production scale, and spatial competition with other neighbouring firms. However, it can be observed that there is contractual trade between distant upstream and downstream linked firms in supply-area analysis. This is especially the case if semi-assembled products are involved in the analysis. These circumstances involve contract and negotiation procedures between both firms. These types of procedures cannot be applied immediately to supply-area analysis due to the static nature of investigation. In addition, it is also necessary to analyse price leadership or price-maker situations, whereas supply-area analysis normally assumes price-taker conditions.

3.8.4. Administrative Boundary and Geographical Supply-Area Conditions

Supply-area analysis assumes that entire spaces are opened to any economic agent and that no restrictions on obtaining inputs exist as long as uniform plain conditions are securely maintained. In general, however, there are international boundaries, trading policies and other administrative economic boundaries. When necessary amounts of inputs exceed the subsistence capacity level within a nation, firms start importing these additional inputs from overseas. As a result, import duties are levied and other extra costs are incurred. In other words, firms have extra transaction costs imposed on them as a result of obtaining inputs from outside their own area. In addition, there may be inaccessible areas within their supply area and these cause further additional cost burdens as firms are required to access wider areas in order to obtain the necessary amount of inputs. Taking into account these economic conditions, there is no uniform spatial cost structure in relation to distance. However, supply-area analysis is only valid under the assumption of uniform cost structure. As a result, the inclusions of administration boundaries and geographical conditions require non-uniform spatial cost structures and this causes difficulties for the analysis of supply areas.

3.8.5. The Structure of Production Function

In supply-area analysis, it is assumed that there is a uniform transportation cost for assembly per unit of distance, and an average input cost which is expressed by constant returns to scale. However, transportation costs may have a decreasing rate of marginal cost in general. As a result, average input cost has to have decreasing returns to scale. However, the existing supply-area analysis assumes production functions with the shape of homogeneous constant returns to scale as expressed by the expression (3-12) as $q = \mu L(q)$. As supply-area analysis disregards external economies, diseconomies and non-constant average assembly cost, the alternative production function should be introduced and expressed with an additional variable k as:

$$q = \frac{1}{k}\mu L(q) \qquad (3\text{-}48)$$

where k $(k \geq 1)$ represents a loss index in the production process. However, this alternative equation cannot be applied in a straightforward manner to supply-area analysis as the relevant cost function is affected by this change and supply-area analysis is valid only if the relevant cost structure is uniform and unchanged.

3.9. The Limitations of Independent Analysis of Supply Areas

Although supply-area analysis has been developed with respect to spatial competition of output, the production function is treated as an external economic factor. As a result, the analysis is unable to examine the production process in a straightforward manner as the parameter is fixed but not a dependent variable. The existing literature on supply areas solves the optimal quantity of output in the production process by applying a simplified linear production function. In order to introduce more plausible conditions, it is necessary to investigate the structure of the production function in terms of an input-output framework. While the differentials between input and output can be explained by internal and external economies, the attempt has not yet been made, as there is the difficulty of including market-area analysis. The next chapter will examine the comparison between both types of area and the following chapters will explore how these two-poled approaches can be simultaneously examined on the same framework.

3.10. Conclusions

As Lösch (1938) and Beckmann (1968) refer to the idea of supply areas in their detailed analysis of market areas, there will be a certain theoretical relationship between the two different types of area analysis. Some of these examinations are attempted in terms of the spatial labour market and the product market. However, market areas and supply areas have not been simultaneously analysed through an input-output framework. In order to attempt further examination in later chapters, the next chapter will investigate the similarity and dissimilarities of both types of area analysis.

Chapter 4. Comparing and Contrasting Market Areas and Supply Areas: Similarities and Dissimilarities

4.1. Introduction

In the previous two chapters, market-area analysis and supply-area analysis were considered independently. There are certain similarities and dissimilarities between these two types of analysis and this chapter will compare and contrast them by the following approaches. First, the similarities of market areas and supply areas will be examined in terms of exclusivity, spatial configurations and relevant economic factors. Second, the dissimilarities of market areas and supply areas will be investigated with reference to the input-output framework, structures of transportation costs and spatial equilibrium procedures. Finally, additional factors for further detailed analysis will be explored in order to combine both types of area analysis in later chapters.

4.2. Similarities of Both Types of Area Analysis

This section will examine the similarities of both types of area analysis with respect to the exclusivity condition, spatial configurations and relevant dependent variables in the analysis.

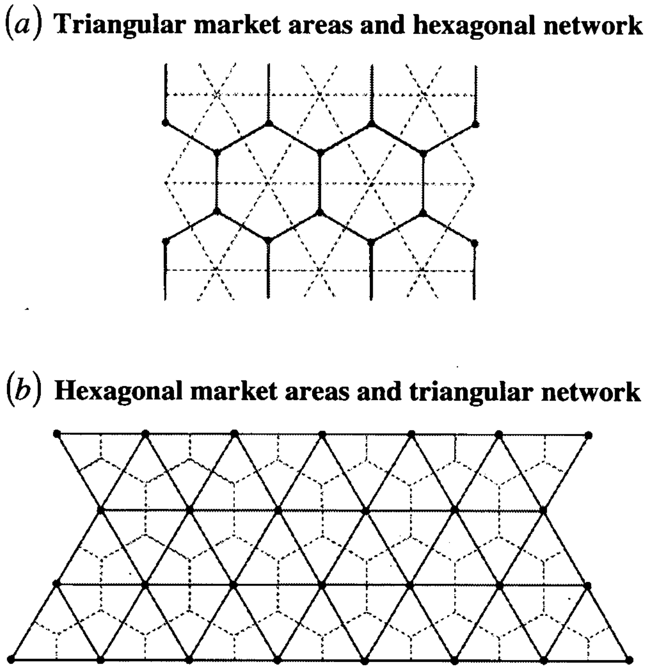
4.2.1. Exclusivity of Both Types of Area

This similarity relates to the exclusivity of areas. Exclusivity implies that a sole firm occupies a space without any overlapping with other competitors. Both types of area analysis generally assume exclusivity of economic space. For market-area analysis, the law of retail gravitation has a form of sharing market areas between two centres. As examined in Chapter 2, this analysis allows sharing of the market according to proportion levels with respect to city size and travel distance. However, market areas are shared by consumers but not by producers. In conclusion, there is a sole distribution point in a market area in market-area analysis, including the law of retail gravitation. Likewise, under the framework of supply-area analysis, each supplier provides input solely to a single assembly point.

4.2.2. Similarity of Spatial Configurations

The properties of spatial configurations have an important role in both types of area analysis and there are similarities between them. As the area configurations are measured on the basis of freight rate and distance, the properties of shape and size are methodologically similar between market areas and supply areas. For example, a circular configuration minimises the mean distance between centres of areas, a hexagonal configuration maximises the number of areas and revenue, with a truncated circular configuration being an intermediate case between the circular and hexagonal spatial configurations. Furthermore, in theory at least, market-area analysis and supply-area analysis have the same transportation network. This implies that market areas and supply areas both share the same transportation route on the plain. Furthermore, both areas are restricted from expanding by the extent of transportation costs. In other words, higher transportation rates act as an incentive for firms to reduce their spatial territories.

Beckmann (1968) shows the relationship between transportation networks and market boundaries, as shown in Figure 4-1 (below). The diagram (a) illustrates regular triangular market areas with hexagonal transportation network and the diagram (a) depicts the more familiar hexagonal market areas with triangular transportation network. According to Beckmann, costs for transportation are calculated by the sum of the horizontal and vertical distances. Although Beckmann applies this only to market areas, it may also be applicable to supply areas if the examination limits the scope of the geographical and mathematical perspectives. The triangular network is geometrically the most efficient allocation with respect to cost minimisation to access other distant locations. On the other hand, the hexagonal configurations are the most efficient in both market-area and supply-area configurations in terms of revenue maximisation and utilisation of space.



•	: Centres
Solid lines	: Transportation network
Broken lines	: Market Boundaries

Figure 4-1. Networks of roads and markets (Source: Beckmann, 1968: 84, modified)

4.2.3. Application of the Same Economic Factors

The economic factors of input and output have similarities and certain linkages between both types of area approach. While market-area analysis and supply-area analysis examine different objectives, several economic factors are common to both types of approach. This can be explained in conventional economic theory as duality theory. Duality theory was first formalised by Shephard (1953) and states that the unknown production function is derived from the given structure of factor cost and cost function. By contrast, the unknown cost function is derived from the given structure of factor cost and production function. An additional interpretation can be supply-area configurations in terms of revenue maximisation and utilisation of space.

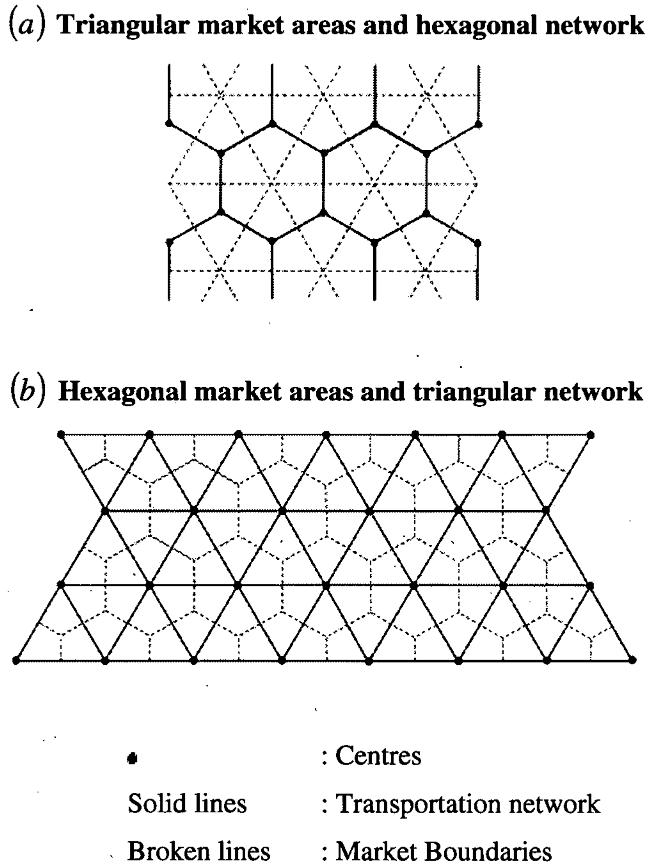


Figure 4-1. Networks of roads and markets (Source: Beckmann, 1968: 84, modified)

4.2.3. Application of the Same Economic Factors

The economic factors of input and output have similarities and certain linkages between both types of area approach. While market-area analysis and supply-area analysis examine different objectives, several economic factors are common to both types of approach. This can be explained in conventional economic theory as duality theory. Duality theory was first formalised by Shephard (1953) and states that the unknown production function is derived from the given structure of factor cost and cost function. By contrast, the unknown cost function is derived from the given structure of factor cost and production function. An additional interpretation can be supplied from these relations using three arbitrary variables: A, B and C. Applying these variables, the duality theory can be restated in terms where C is specified by the combination of A and B. By contrast, B is specified by the combination of A and C. This also implies that A can be specified by the combination of B and C. As shown in Figure 4-2 (below), A and C are directly related to each other through B which itself corresponds to the production function in the original statement of duality theory.

If

$$C \leftarrow \begin{cases} A \\ B \end{cases}$$

$$B \leftarrow \begin{cases} A \\ C \end{cases}$$

then,

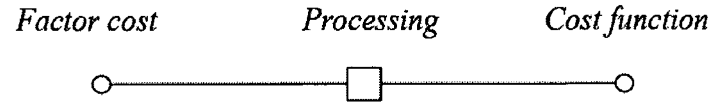
$$A \leftarrow \begin{cases} B \\ C \end{cases}$$

therefore,

$$A \text{ and } B \rightarrow C$$
$$B \text{ and } C \rightarrow A$$

Figure 4-2. The duality theory and an alternative extended form

From the above diagram, the following relationship can be added to the original form of duality theory. *The unknown structure of a factor cost is derived from a given production function and cost function.* As a result, both input and output are specified by the opposite cost structure through the production function. The relationships between factor cost $\sum_{i=1}^{n} C_i(x_i) = w_i x_i + b \quad (i = 1, \ldots, n)$, production function $q = f(x_1, \ldots, x_n)$ and cost function $C = C(q)$ are shown in Figure 4-3 (below).



$$\sum_{i=1}^{n} C_i(x_i) = w_i x_i + b \qquad q = f(x_1, \ldots, x_n) \qquad C = C(q)$$

Figure 4-3. Input, production processing and output linkages

In the above diagram, w_i $(i = 1, \ldots, n)$ represents unit factor price, x_i $(i = 1, \ldots, n)$ represents the amount of input, b = fixed cost and q = quantity of output. This input-output framework shows that the factor cost as input and cost function as output, are connected by the production function $q = f(x_1, \ldots, x_n)$ as processing function.

If market-area analysis and supply-area analysis are considered on an input-output framework, this idea can be applied with some modification to a spatial context. In other words, market areas are specified not only by the demand curve and market organisation, but also by properties of the production function and the structure of the factor cost. Likewise, supply areas are specified not only by the structure of factor cost and competition of inputs, but also by properties of the production function, consumer demand conditions and market organisation. As a result, the following economic variables are similarly applied in both types of area analysis: the structure of factor cost, input-output ratio of technologies, market organisation and demand conditions.

4.3. Dissimilarities of the Market-Area Analysis and Supply-Area Analysis

This section will investigate dissimilarities between both types of area analysis with respect to input-output framework, types of transportation costs and the structure of spatial equilibrium.

4.3.1. Input and Output Analyses

Market-area analysis examines the size, shape and number of settlements of market areas with given structures of factor cost, technologies and demand conditions. Supply-area analysis, by contrast, examines the size, shape and number of suppliers with given assembly costs and demand conditions. Both types of approach are examined on the basis of a shared set of spatial configurations. However, as introduced in Chapter 2, market-area analysis investigates the structure of the market; supply-area analysis, on the other hand, examines the spatial competition of inputs and a part of the demand conditions (see Chapter 3). To elaborate, market-area analysis initially specifies the shipping rate of distribution and the spatial demand curve in order to derive the aggregate terms of the market area. It then considers market organisation in terms of spatial competition. Finally, the equilibrium outcome of the market-area configuration is derived. With supply-area analysis, the first step is to specify the cost structure of the firm, taking into account the assembly shipping cost. Second, the market organisation is considered with respect to the price setting of the market. Finally, the optimal production scale and the equilibrium outcome of the supply-area configuration are derived. In this way, the objectives of spatial competition are clearly different in both types of area analysis.

4.3.2. Definitions of Transportation Costs

In most models, market-area analysis assumes the inclusion of freight-on-board $(f.o.b.)$ distribution cost structure. This suggests that the consumer pays transportation costs between the distribution point and the point of consumption. In this case, transportation costs are added to the unit price of a product as a constant average unit transportation cost t. This is shown in Figure 4-4 (below). These costs directly affect the revenue levels of a firm. In the diagram, u = market-area radius, p = unit price of product, a and b are constants, and AR_1 = demand curve without transportation cost while AR_2 = demand curve with transportation cost.

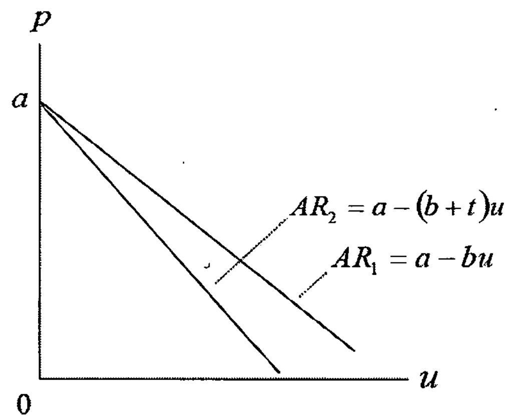


Figure 4-4. Transportation cost t in market-area analysis

By contrast, in most models, supply-area analysis assumes that a producer pays transportation costs between the production plant and the supplier. In this case, the shipping cost per tonne-kilometre is added to the average assembly cost as examined earlier in Chapter 3. This shows that any changes in transportation costs A_θ directly affect the average cost curve AC of the firm, which is the combination of average transportation cost A_θ and average production cost APC as shown in Figure 4-5 (below).

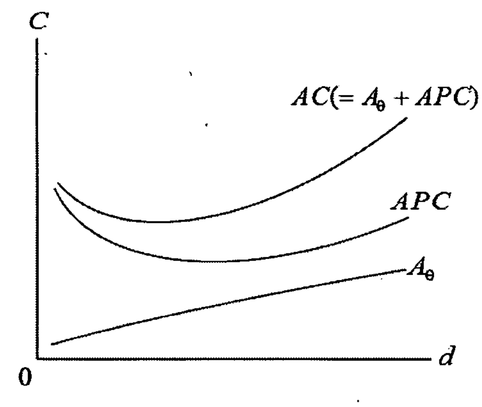


Figure 4-5. Assembly and transportation costs in supply-area analysis

In the above diagram, A_θ = average assembly cost of supply-area configuration θ, APC = average production cost and AC = the vertical sum of these costs as a conventional average cost. As the average assembly cost curve has a property of decreasing returns to scale, the effect of transportation costs per unit declines with distance. This is different from the *f.o.b.* distribution cost of market-area analysis which has a constant rate of transportation cost. In this way, transportation attributes of market areas and supply areas are clearly different and this is one of the most important differences between market-area analysis and supply-area analysis.

4.3.3. Types of Spatial Equilibrium Procedure

In order to clarify the difference between market-area analysis and supply-area analysis with respect to spatial equilibrium procedure, the following two different approaches will be introduced in terms of a priority condition for firm operation. The first case considers a situation in which the supply-area condition has a more important role for firms. In many cases, there is a single market area per product while supply-area analysis has a complex structure due to the presence of more than two inputs or supply areas. Moreover, the consideration of supply areas becomes more important if a producer has the *f.o.b.* pricing system for output and the *c.i.f.* system for inputs. This argument may be compatible with Weber's (1909) location analysis which solves the optimal plant location P with reference to the ratio of transportation costs and weight-bulk of shipments between raw materials RM_1, RM_2 and output MK. This is illustrated in Figure 4-6 (below). The Weber model requires a point analysis as Weber does not apply the concept of areas. Point analysis becomes relevant to market areas when demand is highly concentrated in particular places. For supply areas, point analysis becomes relevant when inputs are available from particular limited suppliers. If the analysis examines single points, and if the primary production determination is supply conditions, Weber analysis can then be acceptable to supply-area analysis.

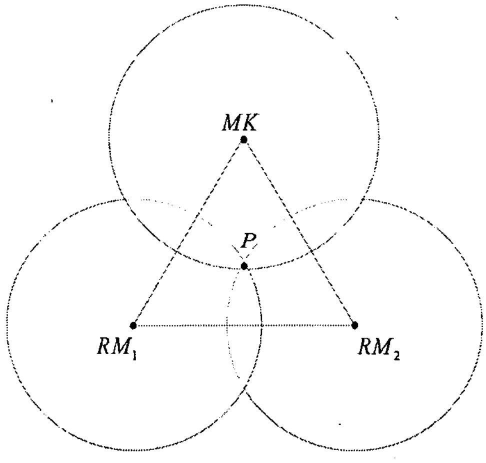


Figure 4-6. Location triangle and the optimal firm location

In many cases, however, production decision is determined by demand conditions on the market. In such cases, market areas will be the primary determination factor for producers, and the spatial equilibrium formation will be different from the above analysis of Weber. The equilibrium procedure follows the Löschian model in this case as demonstrated in Chapter 2. As shown previously, the Löschian model cannot apply the Weber location model due to differences in assumptions between the two models. However, if the market point in the Weber model is expanded and redefined as an area, the optimal plant location may be found in market-area analysis. The difficulty of applying the Weber analysis to market-area analysis is not caused by any inaccuracies in the assumptions which Weber analysis makes; rather, it is caused by the exclusion of spatially constrained internal and external economies in market-area analysis. In this way, one of the dissimilarities between market areas and supply areas can be observed from the standpoint of the analysis of firm location.

4.4. Additional Factors of Market-Area Analysis and Supply-Area Analysis

As indicated in the previous chapters, market-area analysis and supply-area analysis can be simultaneously examined within an integrated framework. However, it is necessary to consider several additional economic factors in both types of area analysis in order to establish this alternative approach.

This section will clarify what sort of economic factors should be included in the alternative analysis. Conventional economic theory partly considers economies with respect to scale, scope and complexity, all of which are spatially unconstrained internal and external economies to the firm. These factors clarify the efficiency and cost-savings of the production process. Another side of economies which has not been taken into account in conventional economic theory is their spatially constrained dimension. This is because the existing approach is aspatial. It should be noted that market-area analysis and supply-area analysis also have not fully taken into account these economies, and that extensive analysis of spatial allocation in market and supply areas has certain theoretical limitations as stated in previous chapters. As a result, inclusions of these spatial economic factors, namely spatially constrained internal and external economies, should be attempted in both types of area analysis. Furthermore, reaction functions of other firms, which have not been sufficiently investigated through market-area analysis and supply-area analysis, should also be considered. These functions, for example, may clarify the relationship between vertical integration in internal economies and activity-complex economies in external economies, observing the locational decision-making process of firms with respect to the integration and disintegration of the organisations.

4.4.1. Partial Inclusion of Spatially Constrained Internal Economies

The spatially constrained internal economies in the framework of market-area analysis and supply-area analysis will now be considered. Both types of area framework assume a simplified form of production function. This poses a problem for an integrated framework approach. The simplification is related to insufficient inclusions of spatially constrained internal economies. The existing market-area analysis and supply-area analysis sufficiently consider economies of horizontal integration or economies of scale as factors which determine the optimal production scale in terms of a cost minimising perspective. However, other elements such as internal economies of lateral and vertical integration are not introduced in the analysis. It is important to take these economies into account as the analysis of operating-cost savings and division of labour may become possible by lateral integration and vertical integration respectively.

4.4.2. Exclusion of Spatially Constrained External Economies

The spatially constrained external economies in the framework of both types of area will now be considered. Regarding the location of the production plant, it may not be situated at the centre of the market area in many cases. For example, certain types of industries prefer to locate with other related firms in areas outside the centre of the market area. This can be seen in many real economic cases, particularly as regards manufacturing. However, such a deviation from the centre cannot be observed in existing market-area analysis and supply-area analysis. The problem is due to an absence of spatially constrained external economies. As a result, it is important to include the following types of economy in both types of area analysis: localisation economies, to examine the negotiation process between relevant neighbouring firms for the purpose of several cost saving opportunities; urbanisation economies, to deal with higher land price and congestion factors and to take advantage of well-established infrastructure; and activity-complex economies to investigate the cost of organising different enterprises in terms of spatial proximity. These economies will be examined further in detail in the following chapter.

4.4.3. Exclusion of Reaction Functions

It is important to consider the notion of reaction functions if a firm relies on upstream and downstream linkages during processing. As a relevant attempt in location analysis, the Weber analysis indicates the negotiation process to the formation of localisation economies. In addition, Hotelling (1929) examines optimal firm location, observing an opponent's economic strategy within the framework of the duopoly price-competition model. Market-area analysis and supply-area analysis should also have certain relationships between producers and the economic behaviour of other neighbouring firms in terms of spatial competition. However, both types of area analysis assume market organisation either to be free-entry competition, a monopoly or monopolistic competition. This focus only takes into account the conditions of demand and cost curves. If economic transactions are included in upstream and downstream linkages, certain reaction functions will be observed which should be included in the analysis if a more detailed integrated framework analysis is to be achieved.

4.4.4. Effects of Each Additional Element

In order to clarify each position of the above introduced economic factors in location analysis, it is important to illustrate the theoretical relationship between these factors in a simple geometric model. The model will consider that there is one input and one output which is produced at an assembly plant. Other spatially constrained variables, such as transaction costs and agglomeration economies, are presumed to be negligible at this stage. A simple case is one in which both a market area and a supply area share a centre at the location of production. It is assumed that an efficient production scale is observed within the firm. These relations are illustrated in Figure 4-7 (below).

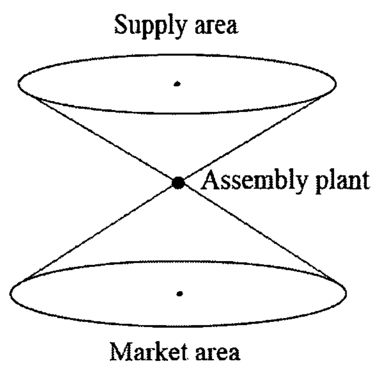


Figure 4-7. A case in which market area and supply area have the same centre

The above diagram shows a case in which the centre of both the supply area and the market area is an assembly plant. If there are transportation-rate advantages and agglomeration economies, however, the optimal firm location may be moved to a different location in the long run. Figure 4-8 (below) shows an alternative case in which there are higher transportation rates on input and a pulling force away from the centre caused by the certain presence of agglomeration economies.

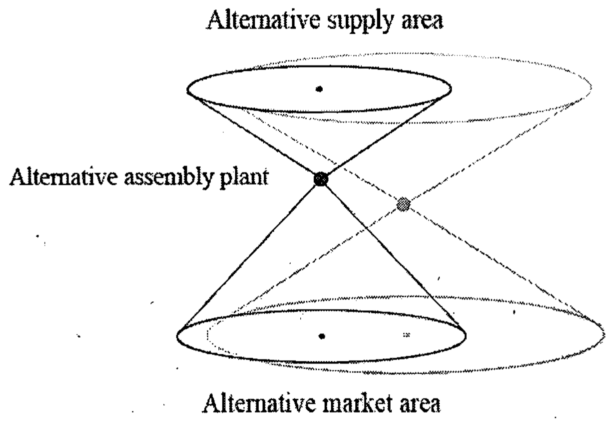


Figure 4-8. The presence of other location factors

The above diagram suggests that a vertical shift from the initial centre can be explained by the transportation-advantage force. In addition, the horizontal shift can be explained by the pulling or pushing force away from, or towards the centre, caused by agglomeration economies. Moreover, a shrinking or enlargement of both market areas and supply areas can also be observed as a result of space-filling competition with other neighbouring competitors. From these considerations of the effects of additional economic factors, it becomes clear that internal economies specify the size of an assembly plant and do not directly affect plant locations. By contrast, external economies specify these locations as regards transportation attributes. Finally, the reaction functions may be examined through the process of the formation of market areas and supply areas with respect to the conditions of other neighbouring competitors. In this way, there are various economic factors which have not been included in the existing framework of market-area analysis and supply-area analysis.

4.5. Conclusion

This chapter examines the similarities and dissimilarities between both types of area analysis. On doing so, it becomes clear that both types of area have certain connections to each other. Although there are several significant similarities between both types of area, these are limited only to the technical basis of theory and do not indicate any similarities from a conceptual standpoint. In other words, market-area analysis deals with output, and supply-area analysis examines inputs for processing. As a result, these similarities do not suggest that market areas and supply areas have symmetrical objectives. While several dissimilarities are found between market and supply areas, these are solely related to theoretical and technical aspects. As a result, these dissimilarities do not suggest that market and supply areas have no connection, or that both types of area cannot be analysed within the same framework. Finally, it becomes clear from this analysis that there are several additional economic factors in the existing framework of market areas and supply areas which should be included in order to investigate the location of firms from the standpoint of area analysis.

Chapter 5. The Introduction of Additional Factors

5.1. Introduction

The previous chapters analyse the spatial structures of market areas and supply areas. In order to examine the location of firms, it is necessary to investigate these independent approaches simultaneously within an integrated framework. For the purpose of this methodological integration, additional spatial economic factors should be taken into account. In this chapter, these spatial factors will be introduced in order to indicate their relation to established market-area analysis and supply-area analysis. In the existing framework of market areas and supply areas, a producer always locates at the centre of an area. This is treated as a fixed condition and further detailed analysis of firm location has not yet been extended to either type of area analysis. As demonstrated in the previous chapter, firms tend to avoid locating their production plant at the centre of a market area in many cases. Likewise, production plants locate away from the centre of supply areas in various cases. As a result, the structural interaction between firm location, market areas and supply areas should be investigated further, applying the established analysis of location of firms. This chapter will introduce the notion of spatially unconstrained and constrained internal and external economies. Although these have not fully been required in the existing framework of market-area analysis and supply-area analysis, they are essential factors for the analysis of firm location. In addition, the relationships between firm operation, location and spatial competition will also be examined. Finally, firm location will be analysed with respect to three industries in order to provide an integrated framework approach.

5.2. Spatially Unconstrained and Constrained Internal and External Economies

The core element of this chapter is spatially constrained internal and external economies -- generally called agglomeration economies. As examined in Parr (2002c), agglomeration economies consist of six elements: scale, scope and complexity dimensions, each of which can have dimensions that are either internal or external to the firm. The dark-shaded area in Figure 5-1 (below) shows spatially constrained internal and external economies where agglomeration economies are observed, while the other parts show spatially unconstrained economies. The un-shaded area represents spatially unconstrained economies and firms are able to obtain internal and external economies without being restricted by their location of production, if their processing requires no distant multiple stages within the firm, and no particular dependence on other firms within the industry or public services.

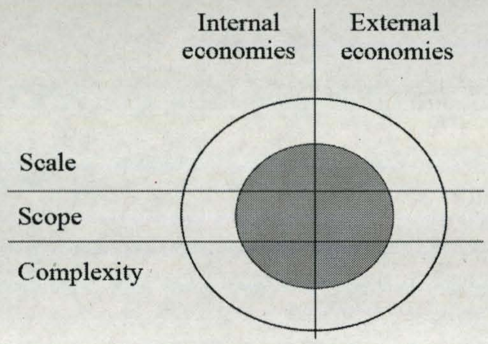


Figure 5-1. Spatially unconstrained and constrained internal and external economies (Parr, 2002c)

If firms require multiple stages within the firm and dependence on other firms within the industry or public utilities, they may have a certain degree of internal and external economies which are brought by particular location conditions. These economies are referred to as spatially constrained internal and external economies or agglomeration economies, as stated above. These economies have a trade-off interaction with transportation costs as investigated by Weber (1909) and this can be referred to as a substitute effect between these two trade-off factors. However, there also exists a complementary effect in terms of transportation costs in addition to the above mentioned substitute effect on agglomeration economies in Weber's framework. This complementary effect is transfer costs between processing stages within the firm in agglomeration economies which are internal to the firm, and between firms in agglomeration economies which are external to the firm and internal to the activity complex. As a result, it should be noted that the argument of the trade-off interaction between agglomeration economies and transportation costs in Weber's approach is solely referred to the substitute type of transportation costs. These costs are non-inter or non-intra firm transportation costs. In terms of this point of view, a dispersed spatial structure is more encouraged. By contrast, the complementary effect requires spatial proximity between two or more firms or industries and it is clear that these two types of transportation costs have an opposite force to each other. The relationship between agglomeration economies and these costs can be illustrated as Figure 5-2 (below).

Agglomeration Economies

// ↖↘

Complementary Transportation Costs ←→ *Subsitute Transportation Costs*

Figure 5-2. Agglomeration economies and two types of transportation costs

If firms rely on economic activity of other firms or industries, it is also important to consider the terms of time saving, transaction cost, and face-to-face negotiation, in addition to the complementary effect of transportation costs. The term time saving works associate with distant transportation as an increasing function. The location proximity reduces these losses with the cost of the complementary type of transportation. Transactions costs and face-to-face negotiation refer to costs for communication and managerial arrangement, which are reduced by spatial proximity with other firms or industries. The following section will examine spatially unconstrained and constrained internal and external economies with respect to firm operation.

5.3. Internal Economies and Firm Operation

There are two dimensions in internal economies, which are spatially unconstrained and constrained. Spatially unconstrained internal economies are able to have benefits without considerations of the location of other processing stages within the firm as illustrated in the conventional aspatial model. By contrast, spatially constrained internal economies are required to refer to the location proximity to other relevant processing stages within the firm. Spatially constrained internal economies are parts of agglomeration economies which are internal to the firm. This section will examine how these economies affect firm operation of processing. In every production process, there is usually some degree of internal economies of horizontal, lateral and vertical integrations. The presence of these economies can indirectly have several effects on firm location, if the economy relies on location proximity. For example, if the total effects of these economies are beneficial for the producer, and spatial proximity is inevitable, dispersed plant locations will be gathered in a specific site and the number of assembly plants within a firm will be reduced by this integration process. By contrast, if these effects are beneficial but do not outweigh other disadvantageous cost factors such as distant higher transportation costs, firm operation may be dispersed and the number of plants within a firm will be increased by disintegration. However, these summaries provide still insufficient details and each economy of agglomeration must be examined in relation to firm operation as follows.

5.3.1. Horizontal Integration and Firm Operation

Horizontal integration indicates that there is a certain cost saving as the production scale increases. There are two types of internal economies in terms of horizontal integration. One is called the economies of scale in terms of the large quantity of production, which is obtained from an increase in the output level along an average production cost curve APC, when further additional production reduces the unit cost of production. This is shown in Figure 5-3 (below). The other type is called the economies of scale in terms of technological change, and is obtained from a change in the long-run average production cost curve $LAPC$. This latter type achieves more cost-saving production as shown by the movement of long-run average production cost curve $LAPC_1$ to $LAPC_2$ in Figure 5-4 (below). The former case is observed in the short-run analysis whereas the latter is observed in the long-run analysis. As demonstrated in the previous chapter, the main difference between short run and long run is whether there is a technological constraint due to the presence of fixed costs.

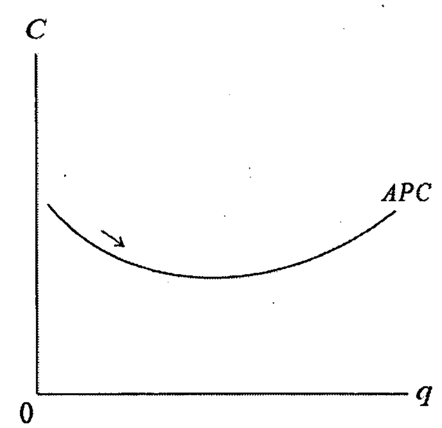


Figure 5-3. Short-run average production cost APC curve

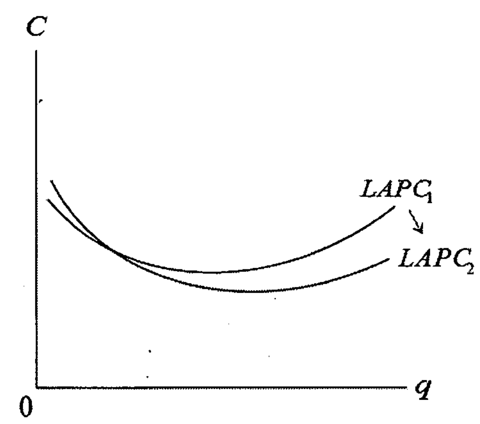


Figure 5-4. Long-run average production cost $LAPC$ curves

Conventional economic analysis suggests that the shift of long-run average production cost curves will follow the movement illustrated in the above diagram. However, it should be noted that this curve might increase as the scale becomes larger in location analysis, due to the presence of the positively sloped average transportation cost curve. In this case, it may not be a straightforward choice to expand the production scale, and existing less efficient production may be maintained in order to avoid cost increases in particular cases. This is shown in Figure 5-5 (below).

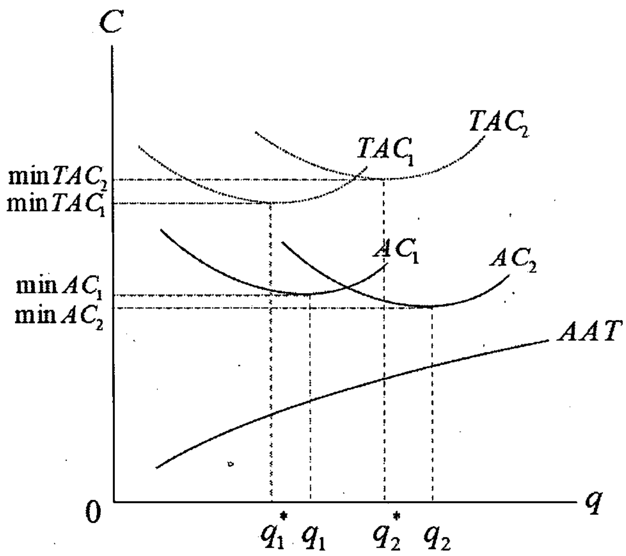


Figure 5-5. Average cost curves and assembly transportation cost

As illustrated in the diagram, the total average cost curve TAC_i $(i = 1,2)$ is the vertically added average assembly transportation cost AAT and the average production cost curve AC_i $(i = 1,2)$. Even though minimum average cost decreases from $\min AC_1$ to $\min AC_2$ due to technological improvements in production, minimum total average cost increases from $\min TAC_1$ to $\min TAC_2$ due to the presence of an increasing average assembly transportation cost AAT. As a result, the output level should not be expanded to q_2^* but sustained at the level q_1^* of the previous technology. Unless technological improvement achieves a dramatic downward shift of the average cost curve in this circumstance, firms will be required to wait for the reduction of the transportation rate. This condition is formally stated as follows.

Subject to:

$$\min AC_1 > \min AC_2$$

$$q_1^* < q_2^*,\ q_1 < q_2 \text{ and } q_i^* < q_i \quad (i = 1,2)$$

$$\frac{\partial AAT}{\partial q} > 0$$

- The combination of the smaller non-advanced technology AC_1 and output level q_1^* is chosen if $\min TAC_1 < \min TAC_2$
- The combination of the larger advanced technology AC_2 and output level q_2^* is chosen if $\min TAC_1 > \min TAC_2$

The horizontal integration is either spatially unconstrained or constrained economies. If it is necessary for firms to locate together in order to take advantage of these types of economies, there will be the spatially constrained type which is referred to as an agglomeration economy. The horizontal integration may also indirectly affect internal economies of vertical integration as will be examined later in this section.

5.3.2. Lateral Integration and Firm Operation

Lateral integration is observed when varieties of production achieve more efficient operation than with the single processing of products. Formally, the following expression can be suggested for total production costs with respect to three types of products within a single firm

$$\begin{aligned} &f(\alpha_1)w_1x_1 + F_1 + f(\alpha_2)w_2x_2 + F_2 + f(\alpha_3)w_3x_3 + F_3 \\ &> f(\alpha_1;\alpha_2;\alpha_3)(w_1x_1 + F_1 + w_2x_2 + F_2 + w_3x_3 + F_3) \end{aligned} \qquad (5\text{-}1)$$

where $f(\alpha_i)$, w_i, x_i and F_i $(i = 1,2,3)$ represent production function, factor price, amount of input and fixed cost for the i^{th} production respectively. The above expression shows the condition that economies of lateral integration experience if the total cost of independent production exceeds the total cost of joint production. More generally:

$$\sum_{i=1}^{n}\left(f(\alpha_i)w_i x_i + F_i\right) > f(\alpha_1;\ldots;\alpha_n)\left(\sum_{i=1}^{n}(w_i x_i + F_i)\right) \qquad (5\text{-}2)$$

This implies that the combined form of production function $f(\alpha_1;\ldots;\alpha_n)$ and individual fixed costs $\sum_{i=1}^{n} F_i$ achieves more efficient production and cost savings than independently established single processing. This state can be brought about either through the concentration of the available technologies in production function, or by the sharing of common facilities in fixed cost. The former case can be categorised as a technological type and the latter as a pecuniary type which will be detailed in the following chapters. This lateral integration can be observed in both spatially unconstrained and constrained circumstances. If the production facility is immobile or inseparable, lateral integration is referred to the part of agglomeration economy which is internal to the firm.

5.3.3. Vertical Integration and Firm Operation

Vertical integration represents the availability of cost saving by integrating several processing stages on the upstream and downstream linkages of a firm. An operational integration can be seen where additional costs by an expansion or reduction of the production scale exceed integrated managerial costs. On the contrary, an operational disintegration can be seen where integral managerial costs exceed separately operated production costs. As shown in Pontes (1992), transaction costs, division of labour and vertical integration have certain relationships with each other. Transaction costs are formally introduced by Coase (1937) in the context of the co-ordination of price mechanisms between contracting firms. Division of labour and vertical integration are analysed in depth by Stigler (1951). Figure 5-6 (below) is a simplified version of the Stigler model. Let us assume that there are two stages, A and B, for processing a product along each average cost curve AC_A and AC_B. If two stages are operated independently, the total average cost becomes AC_{A+B} $(= AC_A + AC_B)$. On the contrary, if the operation is conducted by an integrated form, the total average cost will be AC_I.

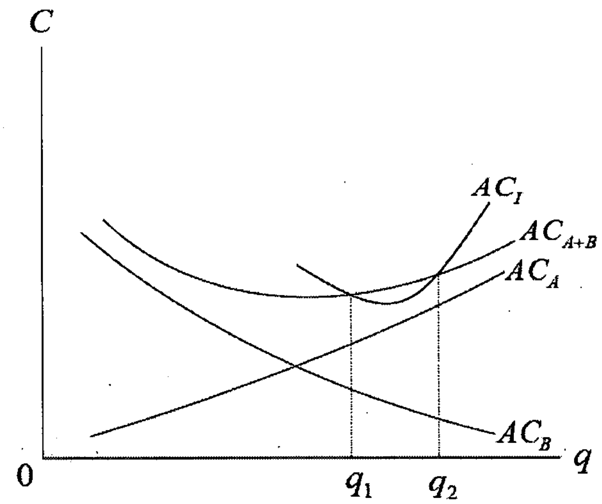


Figure 5-6. Vertical integration and disintegration

In the above diagram, vertical integration can be chosen where the output level is between q_1 and q_2, as the vertically-integrated average cost AC_I is lower than the vertically-disintegrated average cost AC_{A+B} within this range of output. Otherwise, the production will be separately operated by a disintegrated form according to the cost minimisation behaviour of the firm. This analysis can be expanded for more than two stages of processing as demonstrated in Stigler (1951). Vertical integration can be seen in both spatially unconstrained and constrained circumstances. If the integration is achieved with spatial proximity, this may be referred to the part of agglomeration economy which is internal to the firm. As Pontes (1992) states, vertical integration can be enhanced by the availability of spatial proximity, flexible divisions of labour and sufficient information between every stage of production. In spatially constrained terms, vertical integration is encouraged when spatial proximity saves on reheating or liquidity costs between stages on certain kinds of manufacturing process such as iron-steel works and petrochemical plants.

5.3.4. *Effects of Internal Economies on Firm Operation*

All three horizontal, lateral and vertical integrations directly affect the structure of processing costs and production function within a firm in the framework of internal economies. As the combination of processing costs and production function specifies the optimal production scale of the firm, changes in the structure of these integrations provide the extent of change in the production scale. If these economies require spatial proximity, the relationships between firm operation, location and agglomeration economies which are internal to the firm, can be investigated in greater depth.

5.4. External Economies and Firm Operation

This section will concentrate the analysis on spatially constrained economies of scale, scope and complexity, as spatially unconstrained external economies are simply included in economic models by means of reflecting cost structure of firms. Spatially constrained external economies are sub-sets of the various agglomeration economies, which are external to the firm, and over which the firm has no control. This section will examine how these external economies affect firm operation in three dimensions: scale, scope and complexity.

5.4.1. *Localisation Economies and Firm Operation*

In terms of scale, spatially constrained external economies are referred to as localisation economies. Localisation economies are observed when there are possibilities for firms to obtain labour cost savings, joint action for input extraction and specialised services. If these economies are achieved by spatial proximity, it is necessary for these firms to locate at one specific site. In this case, relevant distant firms consider moving to this site, or existing firms try to attract these firms to locate together by negotiation. This procedure is conducted through cooperative negotiation, considering the ex-post advantages from the localisation economies. These economies are one of the elements for specifying the optimal firm location, as will be further examined in a later section of this chapter. Localisation economies have already been indicated by Marshall (1890) as the examination of the localisation of industry, suggesting such physical conditions as climate and soil, and availability of mines or quarries. These physical conditions are both concerned with the importance of specialised skilled labour. Marshall also provides other advantages such as accessibility to a pool of labour, achieving lower assembly, transportation and fuel costs, in addition to such advantages as the sharing of new ideas and information, subsidiary trade, less expensive machinery, specialised services and cooperative joint action regarding the supply of inputs for marketing, and research and development. These ideas have not been included in either the production function in conventional economic theory or in location analysis in the framework of market and supply areas. In location theory, localisation economies are initially formalised by Weber (1909) and this will be introduced in the next section of this chapter. If these economies require no spatial proximity and are still able to obtain certain degree of economies, there exist spatially unconstrained external economies of scale.

5.4.2. Urbanisation Economies and Firm Operation

Urbanisation economies are generally located in metropolitan areas as a result of the various cost saving benefits to be had in such areas. Urbanisation economies can have positive or negative factors for firms. Advantageous factors - which include administrative accessibility, well-organised infrastructure, variety of labour supply, and a highly advanced system of transportation and communication - involve different and unrelated industries in a large urban area. These services are enhanced by the existence of various businesses, municipal and commercial services. However, disadvantageous factors also exist, such as the higher price of land, congestion and pollution. If positive factors exceed these negative factors, it is likely that the plant will situate in the metropolitan area; but the contrary is also true. In this way, urbanisation economies affect the location decision of firms, either towards the metropolitan area or away from the metropolitan area. These net effects are considered with the index of urbanisation economies. For instance, Isard (1956) suggests that hypothetical economies of scale with urban size are composed of the economies of labour, transportation, power and education. In addition, Evans (1972) indicates an index whereby aggregate urbanisation economies can be estimated in terms of scale and costs of floor space, labour, business service and capital. From this point of view, the optimal city size is specified where the total costs are minimised. This will be further studied in Chapter 7.

For urbanisation economies, the following economic characteristics should be noted. First, the location theory, as established, examines the optimal location of the production plant. However, the optimal firm location with respect to conveniences or amenities, such as convenient access to the metropolitan area, well-organised infrastructure or variety of labour force, have not been sufficiently analysed, particularly in the framework of market-area analysis and supply-area analysis. Second, if a production plant is located in a metropolitan area, the firm not only obtains certain opportunities for cost saving, but also faces diseconomies as stated earlier. These factors can be observed in the structure of supply areas if the centre of the supply area is located in an urban area. As will be shown in a later section of this chapter, however, there are a number of cases in which these structural identities between urban areas and supply areas cannot be recognised. Alternatively, the following extensive supply-area analysis can be provided in terms of urbanisation economies. If the production plant is located in a rural area, lower land cost and fewer external diseconomies are achieved in comparison to a metropolitan area. However, various advantages of urbanisation economies cannot be obtained any more. These can be seen where the structure of supply areas and the spatial urban structure are formed in completely different ways. If these economies do not rely on spatial proximity, these are referred to spatially unconstrained external economies of scope. This can be seen where there are particular advantages of well-organised infrastructure, municipal services and convenient transportation at a specific nation or region.

5.4.3. Activity-Complex Economies and Firm Operation

Activity-complex economies rely on trade between different firms in a product chain. These upstream and downstream linkages are encouraged when lower transaction costs and transportation costs are available between succinct stages with other firms. In this case, firms that are relatively flexible to move their plant location, are required to locate at a specific economic site where there is sufficient access to relevant firms or industries. The advantage of spatial proximity in terms of activity-complex economies is having cost savings on energy and on transportation costs for assembly and distribution between production stages. In addition, spatial proximity also encourages better communication systems and increased availability of inputs and outputs between relevant stages. According to Parr (2002c), there are two types of complexity: the first relies on specialised firms at particular production stages, while the second relies on several specialised providers for the supply inputs for the final assembly. This is shown in Figure 5-7 (below).

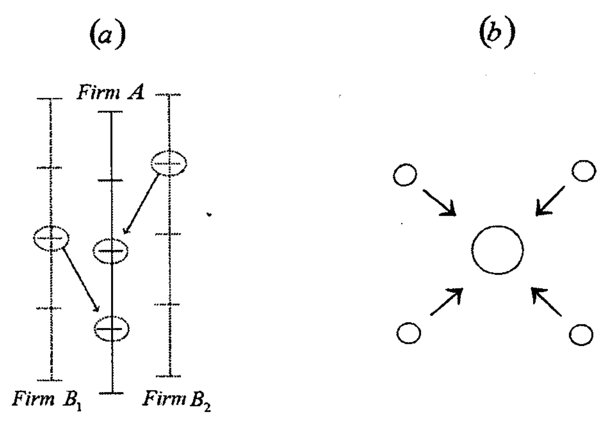


Figure 5-7. Two types of activity-complex economies

The former case in the above diagram (a) represents a case in which *Firm A* is processing a product with five different stages. This firm operates the first, second and final stages with their own facilities. On the other hand, the third and fourth stages are distributed from the third stage of *Firm* B_1 and the second stage of *Firm* B_2, respectively. The latter case in the diagram (b) shows a case in which a producer relies on several independent suppliers within an industry complex. Such a case can be seen at Silicon Valley in California, with the concentration of aero-space production in Los Angels, Seattle and Toulouse, and with concentrations of pottery industries at locations in southern Japan. These economies can also be referred to as spatially unconstrained external economies of complexity, if firms do not require spatial proximity with other relevant firms.

5.4.4. Effects of External Economies on Firm Operation

Localisation, urbanisation and activity-complex economies have an incentive to locate firms at a particular site in order to achieve certain cost savings as spatially constrained external economies. These external economies cannot directly be measured in terms of cost aspects as is the case with internal economies. However, the aggregate effects of external economies, particularly those which are spatially constrained, have an important role in investigating the decision of firms to locate at particular economic sites. As will be demonstrated in the following chapter, the aggregate effect of these economies can be integrated into the structures of the production function and factor cost curve.

5.5. Agglomeration Economies and Firm Location

This section will analyse the impact of agglomeration economies on firm location by the comparison between two locations: with respect to the location triangle approach and an alternative extensive approach.

5.5.1. *Location Incentive of the Firm*

While agglomeration economies are divided into six parts, some of these economies work together or have trade-off relationships between them. In terms of scale, the horizontal integration within a firm closely relates to labour utilisation among industries as localisation economies. With respect to scope, both lateral integration within the firm, and urbanisation economies rely on given potential residual economies in order to share facilities and devices. Regarding complexity, while internal economies of vertical integration are observed within a firm as managerial integration for saving transactions and communication costs, activity-complex economies are observed as cooperation or partnership between different firms. This is relevant to the extent of transaction costs between firms, information availability and the reaction functions of other relevant firms. In this way, these economies of scale, scope and complexity may coexist and work together beyond the categorisation of the six types of agglomeration economies. This can be observed in the following example. Figure 5-8 (below) illustrates the cost curves of a firm for production at two different locations.

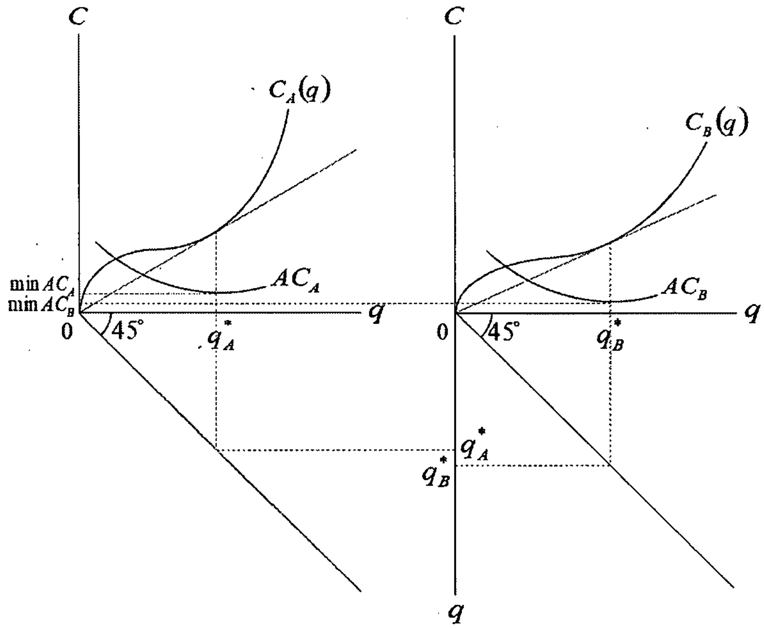


Figure 5-8. Cost curves for a production at location A (left hand) and B (right hand)

In the above diagram, if these productions are operated by the same firm with the same technological condition, the differences, $q_A^* < q_B^*$ and $\min AC_A > \min AC_B$ can be examined by the elements of agglomeration economies: the availability of a better labour pool environment at location B, availability of better public utilities at location B, and availability of better cooperation with neighbouring unrelated firms.

If the cost saving in a particular location is relevant to the obtaining of inputs, either the assembly transportation rate or the factor price must be lower than in the other locations. However, if there is no difference between the input condition for the two locations, these cost differences must be explained by the difference between the cost function and the factor cost curve of the two locations. This is the point at which it is required to introduce the spatial production function. The production function can be added to gain the extra information needed to evaluate the extent of the economies of agglomeration. In order to avoid confusion with the conventional production function, this alternative integrated form will be termed the "spatial production function". The spatial production function will be examined later in the following chapter. The firm chooses either of two locations, A or B in the above diagram. These are chosen on the basis of which is more advantageous in terms of profit maximisation behaviour. Regarding this criterion, the cost structures between locations A and B must be compared in addition to the condition of the assembly transportation cost and the demand conditions. In other words, in order to specify the optimal location of the production plant, it is necessary to investigate economies of agglomeration and transportation costs, in addition to market-area analysis and supply-area analysis. As explained above, every element of agglomeration economies can be observed in these types of analysis and no part of these elements should be neglected for reasons of simplicity.

5.5.2. *Location Triangle and Firm Location*

Agglomeration economies are first introduced to the location analysis by Weber (1909). These economies have two economic factors, namely costs of transportation and costs of labour. Weber examines the availability of localisation economies referring to Launhardt (1885) who, himself, investigated the relationship between travel distance and cost for transportation. Let us assume that there are three firms which are situated at distant locations and belong to the same industry. In addition, each firm is processing a product which is distributed to a point of the market MK, using two types of raw materials RM_1 and RM_2. If these firms locate together, it is possible to have spatially constrained external economies. However, this alternative location may cost more than the original locations for firms in terms of transportation costs for the procurement of raw materials and the distribution of products. As a result, the balance of agglomeration economies and the additional burden of costs for transportation should be taken into account. This is measured by an index in location triangle analysis. This index is called isodapanes after the lines of aggregate minimum transportation costs which radiate in all directions from the centre of the original firm location. In addition, the maximum feasible isodapane is called a critical isodapane. It is assumed that the optimal firm location is situated within a location triangle. The location triangle is illustrated by connecting three apexes of two raw material sites RM_1 and RM_2 and a point of the market MK. The optimal firm location is specified by the condition of cost minimisation for transportation. There are three types of transportation costs: transportation from the raw material sites RM_1 and RM_2 to the production plant, and transportation to the market MK from the plant.

The individual loci of the transportation costs away from three sites are provided radially from the centre. This is initially formalised by Hoover (1937) and named isotimes. The solution for the optimal firm location is exemplified in Figure 5-9 (below). If the firm locates at point A, the total cost of transportation will be 1230 $(=450+480+300)$. Likewise, at other locations, it costs 1250 $(=150+600+500)$ at point B, and 1210 $(=450+360+400)$ at point C. In this way, the minimum total transportation cost will be 1180 $(=300+480+400)$ at point P.

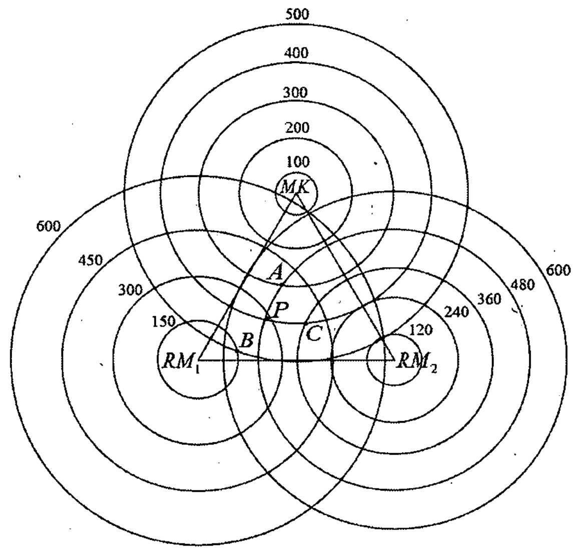


Figure 5-9. Isotims and firm location (Referred to Hoover, 1937: 12)

Figure 5-10 (below) shows plant locations P_1, P_2 and P_3 of each firm as they are respectively allocated by the isotims. Each firm has a critical isodapane which is illustrated as a circle in this diagram. If three critical isodapanes of all firms have an intersection, agglomeration economies can be available at that area. In the case of this diagram, the area E is the alternative common location of these firms.

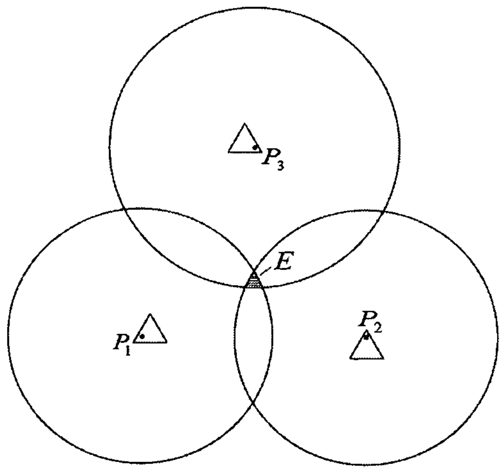


Figure 5-10. Critical isodapanes and the feasibility of agglomeration economies (Source: Weber, 1909: 135, slightly changed)

In the case of Figure 5-11 (below), there is no intersection between three isodapanes, and agglomeration economies will not be achieved.

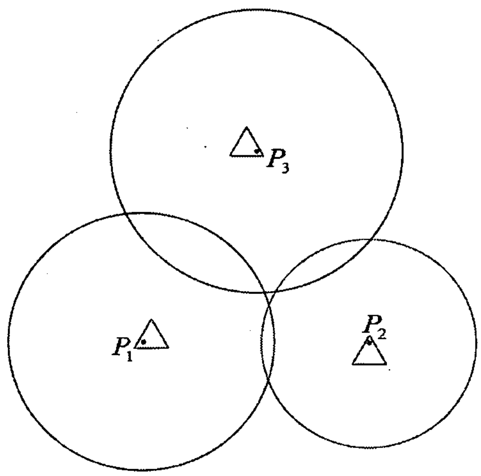


Figure 5-11. Critical isodapanes and no agglomeration economies (Source: Isard, 1956: 177, slightly changed)

According to Isard (1956), there is still a possibility of having agglomeration economies in this circumstance. If two firms offer financial assistance to one firm whose critical isodapane is nearly at a sufficient level to have an intersection between three firms but this has not yet been achieved, this firm can expand the critical isodapane, and eventually all three firms can share the intersection with each other. If the result achieves a less expensive operation than the original separated production, this negotiation should be taken for the purpose of cost minimisation for all three firms. However, this negotiation may not be done if the burden of the financial assistance of two firms cannot be set off by the surplus of the economies of agglomeration between the three firms. This process can be formalised by two methods: by extensions of the bargaining solution with respect to utility maximisation between two individuals, and by the negotiation model by cooperative games between two individuals. Both of these are respectively evolved by Nash (1950; 1953).

5.5.3. Location Triangle and Additional Factors

The primitive Weber model involves the fact that transportation costs $RM_1\ P$, $RM_2\ P$ and $P\ MK$ are equivalent and that there are no agglomeration economies as illustrated in Figure 5-12 (below).

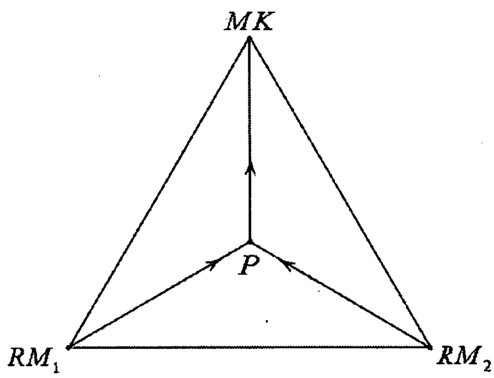


Figure 5-12. A primitive Weber model

An alternative case is that agglomeration economies are available at RM_1 with other firms. The combination of these economies and the saving for transportation cost from RM_1 to P exceeds the extra burden of total transportation costs $RM_2\ P$ and $P\ MK$. These are caused by more distant travel as shown in Figure 5-13 (below).

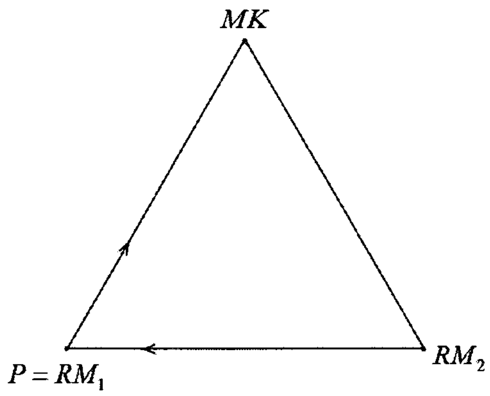


Figure 5-13. A corner solution by the economies of agglomeration

In this case, six different firms may be sharing the raw material RM_1 if the spatial structure is assumed to be regular triangular space, as shown in Figure 5-14 (below). In addition, these firms can share several advantages of localisation economies such as machinery repairing services, a pool of labour and joint research opportunities at point P.

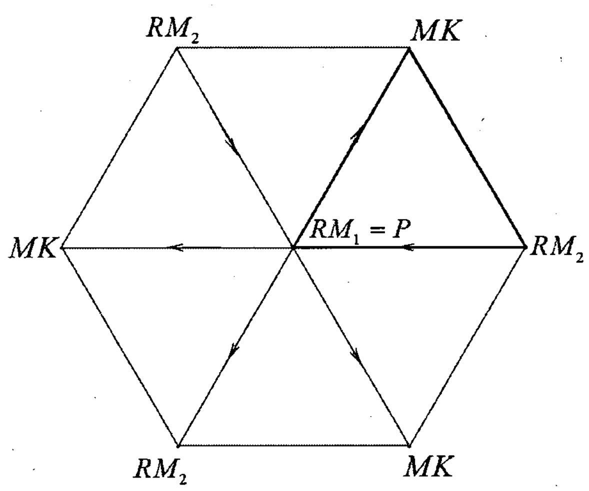


Figure 5-14. Agglomeration economies at RM_1 in the regular triangular space

Another example can also be suggested. Figure 5-15 (below) shows a case in which the transportation cost for distribution P MK is assumed to the $f.o.b.$ as with the Löschian approach. In this case, the firm is not required to consider the cost of transportation between production point P and its market MK, as the distribution cost is imposed on consumers. As a result, firm location is at the middle of the base of the location triangle. This case is valid only if this product is not involved in severe market price competition as will be demonstrated in the following case.

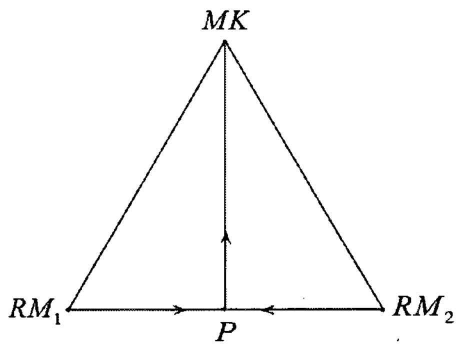


Figure 5-15. The $f.o.b.$ distribution cost and firm location

Figure 5-16 (below) shows a case in which there is severe market competition and the $f.o.b.$ consumer price has to be reduced to the competitive level. This can also be seen when the product relies on municipal services, public utilities and spatial proximity to consumers. This can be referred to as the urbanisation type of agglomeration economy. In addition, this firm does not locate at the market point MK but at a point P. There are two possible reasons for this, as follows. One is due to the presence of remarkably high transportation costs of inputs, and the other is the presence of urbanisation diseconomies at the market point MK, if this point is the centre of the metropolitan area and the production plant requires wide use of land, an un-congested transportation network, clean air and non-polluted water. Further examinations of these economies will occur in the following sections. In terms of market competition, Hwang and Mai (1992) suggest similar evidence for firm location. Although their approach takes into account market competition and demand conditions, the pulling force to the centre of the market by agglomeration economies is neither implicated nor examined.

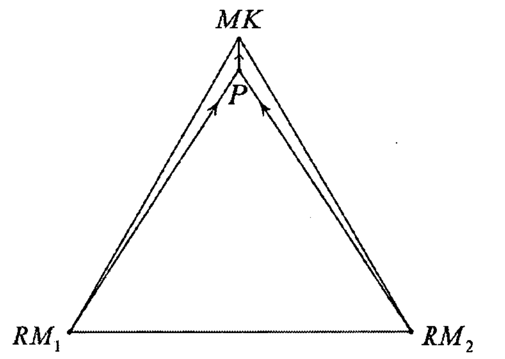


Figure 5-16. Firm operation under severe market competition and urbanisation diseconomies

It should be noted that if the analysis is purely based on Weber (1909), agglomeration economies are achieved when several location triangles stand closer together and the likelihood of their being achieved is based on the level of the critical isodapane of each location triangle, as previously introduced. In other words, his original approach solely referred to localisation types of agglomeration economy for particular manufacturing firms.

5.6. Firm Operation, Location and Spatial Competition

The previous section examines the relationship between each element of agglomeration economies and firm operation in terms of the location of the firm. In order to analyse the location of a firm with agglomeration economies in spatial competitive models, it is required to observe aggregate effects of these economies with transportation cost factors. This section will explore the relationships between agglomeration economies, transportation costs, and market-area and supply-area organisations.

5.6.1. Agglomeration Economies and Transportation Costs

In general, firms consider either production concentration or production dispersion when certain levels of transportation and transaction costs are present. For instance, producers establish branch plants when agglomeration economies decline. In this case, the producer will have dispersed plant locations. In this way, processing costs, fixed cost, and transactions costs for this production, increase if branch plants are established. On the contrary, this dispersion contributes to savings on those transportation costs which used to be borne as a result of sustaining the economies of agglomeration at a particular location. The relationship between agglomeration economies, transportation costs TrC, processing cost PC (which is related to horizontal integration), fixed cost level F (which is related to lateral integration), and transactions costs TRS (which is related to vertical integration) may have the following relationship in the term of absolute value.

$$|TrC \downarrow| > |PC \uparrow + F \uparrow + TRS \uparrow| \qquad (5\text{-}3)$$

The above equation shows that if the absolute value of increases in the total processing cost PC, fixed cost F, and transactions costs TRS, is lower than the absolute value of the saving of transportation cost, the decision may be taken to disperse the plant. On the contrary, branch plants may be reduced, and the production scale of the original single assembly plant expanded, when agglomeration economies increase, and the following expression is provided:

$$|TrC \uparrow| > |PC \downarrow + F \downarrow + TRS \downarrow| \qquad (5\text{-}4)$$

This case shows that the absolute value of the reduction of the total of processing cost PC, fixed cost F and transactions costs TRS, is lower than the absolute value of the additional burden of the transportation cost. The trade-off interaction between agglomeration economies and transportation costs is examined in the framework of Weber analysis. The following section will expand this approach to more complex cases in terms of area analysis.

5.6.2. Market-Area and Supply-Area Organisations and Agglomeration Economies

As examined in the previous chapter, market-area organisation deals with competition of output, and supply-area organisation relates to extracts of inputs. These are of significant importance in determining the optimal firm location and the optimal scale of production. The optimal scale is examined in the existing location theory with respect to given spatial demand conditions and price levels in the framework of market-area analysis. By contrast, the analysis of the relevant supply-area organisations tends to be insufficient, leading to the conclusion that these investigations should be conducted further.

Supply-area analysis is based not only on the elements of the spatial competition of input, but also on the spatial competition of market-area analysis. As a result, these two independent approaches should be simultaneously examined in a single framework. However, attempts at methodological integration experience difficulties with regard to the additional factors required to complete the element. As is suggested in the previous sections, these additional elements comprise a set of spatially unconstrained and constrained internal and external economies. In addition, it becomes apparent at this stage that the analysis of firm location requires consideration of agglomeration economies. As existing market-area analysis and supply-area analysis exclude the notion of agglomeration economies, these approaches have difficulties when used to try to determine the optimal firm location. While it is difficult to measure the extent of agglomeration economies in a straightforward manner, these economies should be introduced to the analysis of spatial competition. The combination of market-area analysis and supply-area analysis may have residual economic factors in an input-output framework and a part of these can be explained by the production function, while others can be explained by economies of agglomeration.

As examined in the previous sections, agglomeration economies are divided into internal and external dimensions. Internal aspects can be simply included in the production cost curve. However, the external aspects cannot be directly added to the production cost curve. According to Meade (1952), external economies are further categorised into two groups: technological and pecuniary external economies. As Scitovsky (1954) demonstrates, both types of external economies can be added to the conventional production function and cost curves. According to these categorising methods, external economies can be included in the cost and production analyses. In this way, the integrated analysis will examine market areas and supply areas simultaneously with the elements of agglomeration economies. The next section will demonstrate a hypothetical model of firm location, applying these methods in terms of the three industries.

5.7. Firm Location in Terms of the Three Types of Industry

This section will exemplify several location patterns of the production plant in terms of the three industries. It is generally stated that market-area analysis considers one producer in one market area, and the production plant is assumed to locate at the centre of the market area. In addition, supply-area analysis assumes one producer in one supply area, and the production plant is assumed to locate at the centre of the supply area. Taking into account these conditions, the integrated framework analysis will initially assume that there are *m* regions and *n* producers affected by the economies of agglomeration. Firm location can be explained by agglomeration economies with respect to the three industries, namely primary, secondary and tertiary industries.

Let us now suppose that there are two types of input RM_1 and RM_2 for processing an output level q. Figure 5-17 (below) illustrates the market area of this output, the metropolitan area, and the supply area as part of the inputs for this production. Point M_0 represents the centre of the market area of this product and also that of the metropolitan area. Point P is located at the centre of the supply area of input RM_1. Point RM_2 is a single supply point of another input RM_2 and the transportation costs for shipping this input RM_2 are assumed to be negligibly small. Finally, point I is situated somewhere between the metropolitan area and the supply point or supply area.

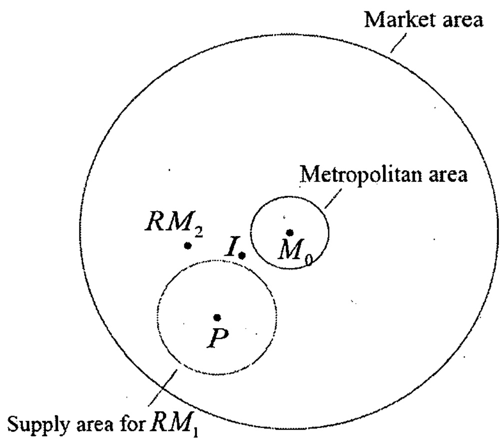


Figure 5-17. Two inputs, one production site, and its market area

In the above diagram, it can be stated that point P is the possible production plant location for primary industries, point I is the location for secondary industries and point M_0 for tertiary industries. The evidence for this claim will be examined in the following subsections.

5.7.1. Primary Industries

Agricultural, forestry and fishery industries are categorised as primary industries, in the sense that they work directly with natural resources. In this case, it is preferential that their production site be located close to the supply area. If several relevant firms who engage with the same industry locate together and achieve sufficient levels of economies, localisation economies can be observed at this location P in Figure 5-17 (above). In this way, it may be more common for localisation economies to occur with supply-area oriented industries. This case is illustrated in Figure 5-18 (below). The supply area would usually be in a completely different location from the metropolitan area unless there were severe market competition or transportation problems for distribution, as previously demonstrated in Figure 5-16 (shown earlier).

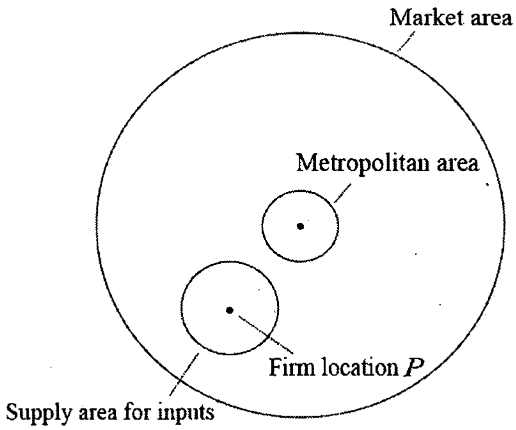


Figure 5-18. Market area and supply area in the case of primary industries

5.7.2. Secondary Industries

This type of industry is characterised by manufacturing or processing. Although the mining industry is generally included in this category, it will be excluded in this argument as this generally does not involve a processing stage. In secondary industries, the production location can be between the primary and tertiary cases -- for instance at the point *I* in Figure 5-17 (shown earlier). In this case, activity-complex economies can be observed if their processing involves multi-stage production, and the transactions costs of the upstream or downstream firms can be kept at low levels. Their location tends to move towards the metropolitan area if market price competition becomes severe and assembly transportation costs are at a relatively low level. By contrast, their location tends to move towards the supply area if costs for locating in a metropolitan area are sufficiently high or the assembly transportation cost is at a remarkably high level. In this way, manufacturing and processing industries which are willing to avoid urbanisation diseconomies, may locate at distant points from the centres of market areas, as shown in Figure 5-19 (below) which illustrates the discount rate of urbanisation diseconomies r_{ud} in two market areas, which have centres MA_1 and MA_2. In this diagram, the point P achieves the minimisation of urbanisation diseconomies.

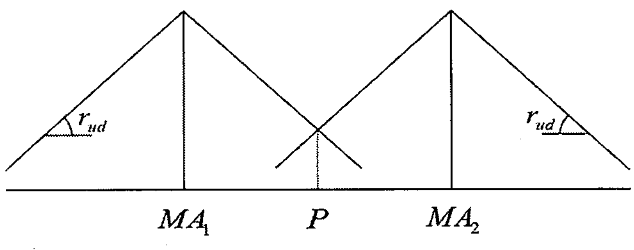


Figure 5-19. Location of production and urbanisation diseconomies

This conflicting force against urbanisation economies is also indicated by Marshall (1890), who cites higher ground-rates at central sites of large towns. He concludes that firms are not required to locate at the centre of the market if the costs of communications and transactions between a distant production site and the centre of the market are at a sufficiently low level. However, it is also necessary to state that this should be applied solely to particular types of industry, namely secondary industries, with the exception of the mining industry. While Marshall assumes a point analysis, this idea can also be applied to area analysis. In area analysis, if the entire space is formed by regular circular market areas, each firm is hypothetically surrounded by three market areas as shown in Figure 5-20 (below). Contrastingly, each market area is surrounded by six producers. As a result, each producer distributes one third of their products over each market area, and each market area is distributed by six firms each taking a one sixth share. These spatial patterns may have an opportunity of localisation and activity-complex types of agglomeration economies, if firms locate at a common site, and they rely on certain benefits from economic activity of other firms by location proximity, which will be examined in Chapter 7. To summarise, firms in secondary industries tend to be located towards their supply areas for the purpose of economising on assembly costs, but also in order to avoid urbanisation diseconomies, even though additions in distribution costs will be incurred, which are typically less than the above mentioned cost savings.

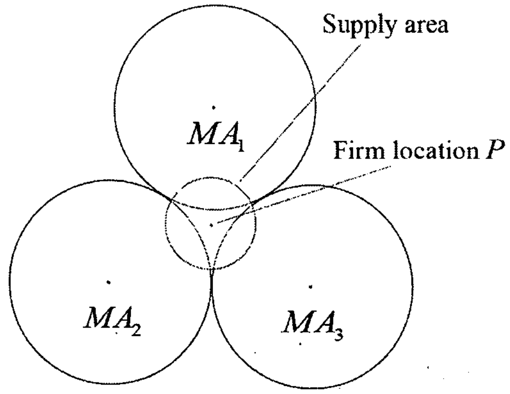


Figure 5-20. Market areas and supply areas in the case of manufacturing industry

5.7.3. Tertiary Industries

Tertiary industries are characterised by commerce, transportation, communication and the service industries. They tend to locate at M_0 in Figure 5-17 (shown earlier), namely in the metropolitan area. Tertiary industries, which take advantage of urbanisation economies, will locate at the centre of the market area as long as the advantage of locating at an expensive metropolitan area is higher than the advantage of locating at a less expensive rural area. The situation is illustrated in Figure 5-21 (below) and it can be stated that urbanisation economies typically occur with market-oriented industries. In general, firms in tertiary industries (particularly those serving households) tend to be located close to the centres of their respective market areas. This is in order to economise aggregate distribution costs, incurred by consumers. It is also the case that inputs (labour, municipal services, public utilities and access to wholesales) are ubiquitous, so that the notion of supply areas loses its significance.

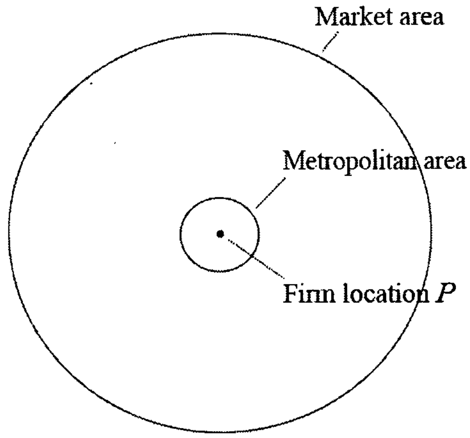


Figure 5-21. Market area in the case of tertiary industries

5.8. Conclusion

This chapter first examines the impact of each element of agglomeration economies on firm operation. In addition, the primitive approach of the location triangle is extended in terms of the trade-off interaction between firm location and agglomeration economies. Furthermore, it analyses the alternative relationship between firm location and agglomeration economies within a single framework of market-area analysis and supply-area analysis. The important evidence in this chapter is that agglomeration economies and spatial area distribution have certain inevitable relationships. These examinations enable us to address theoretical evidence of the interaction between particular firm locations and production scale. However, these attempts have not yet been conducted in existing location analysis. The following chapters will include these additional economic factors in market-area analysis and supply-area analysis to arrive at an integrated framework approach. It should be noted that all the categorisations in the last section are merely the major possibilities and that there are a number of exceptions in the observation of agglomeration economies.

Chapter 6. Spatial Equilibrium Analysis in an Integrated Framework

6.1. Introduction

This chapter will develop an integrated-framework analysis of market areas and supply areas. This will be valid with the inclusion of additional economic factors in the spatial model. As previously discussed, the necessity of this integrated framework has arisen on the understanding that the optimal firm location can only be properly determined if the relevant market areas, supply areas and spatial economic factors are sufficiently included in the analysis. This approach will refer to the input-output framework with the given configuration of the production function. The structural relationship follows the alternative form of duality theory as examined in Chapter 4. At this stage, however, market areas and supply areas will be applied to the limited case of firm location. The result will show that the market area and supply area have an identical centre, which is not a plausible spatial pattern and is in need of improvement in order to examine more general cases. As a result, this applied spatial duality theory (SDT) should also introduce additional economic factors which are relevant to the location problem, as observed in Chapter 5. This will require the re-examination of the structure of factor cost and production function. The modified forms will be called the spatial factor cost and spatial production and cost functions. The integrated framework will be generated in this way and this chapter will examine each dependent variable with comparative-static analysis.

6.2. An Outline of the Integrated Framework

This section will introduce a model framework and technical terms. The objective of the analysis is to examine an integrated framework of market-area analysis and supply-area analysis. There are certain limitations which arise when combining both types of area, as several essential economic factors are excluded from each established theoretical framework, as examined in the previous chapter. The integrated framework will first be demonstrated on a simple aspatial economic condition in order to clarify the structure of alternative duality theory which is examined in Chapter 4. This will then be extended to the alternative spatial duality theory (SDT), which contains spatial economic factors with respect to distance and technologies. In order to include these additional factors in the SDT model, the structure of factor cost, cost and production functions must be modified. These spatially modified factors will then enable us to conduct the SDT model as an integrated framework of market-area analysis and supply-area analysis.

The technical terminology of this analysis as it is generally used is defined as follows:

- $f.o.b.$: Freight on board
- $c.i.f.$: Cost, insurance and freight
- AR : Average revenue
- MR : Marginal revenue
- MC : Marginal cost
- AC : Average cost
- LAC : Long-run average cost
- TC : Total cost
- TR : Total revenue
- F : Fixed cost
- F_τ : Fixed terminal cost for assembly transportation
- F_ε : Fixed additional factor
- τ : Transportation rate in assembly per tonne-kilometre
- t : Transportation rate in distribution per tonne-kilometre
- ε : Index of additional factors (external to the firm)
- p : Output price
- Π : Profit

w : Factor price

x : Input

RM : Raw material

K : Capital

L : Labour

q : Quantity of production

Q : Total output

u : Market-area radius

U : Maximum market-area radius

s : Supply-area radius

q_F : Consumer demand

AR : Average revenue

D : Density of demand

D_X : Density of input

$LAPC$: Long-run average production cost

X_T : Technical efficiency

X_I : Internal economies

X_E : External economies

The model assumes first that transportation for market areas has a $f.o.b.$ pricing system with a constant transportation rate t. Second, transportation for supply areas has a $c.i.f.$ pricing system with a constant transportation rate τ. Third, consumers and inputs are distributed uniformly and continuously on the plain at densities of D and D_X respectively. Finally, every consumer has an identical individual demand curve for products.

6.3. The Theoretical Foundation of Integrated Market-Area Analysis and Supply-Area Analysis

This section will introduce objective dependent variables in an integrated framework analysis of market areas and supply areas. These dependent variables will be observed in further detail with various hypothetical examples in the following chapter.

6.3.1. Sizes and Shapes of Market Areas and Supply Areas

The size and shape of market areas and supply areas will be analysed through four types of case between market areas and supply areas: similar size and similar shape cases, similar size and different shape cases, different size and similar shape cases and different size and different shape cases. As examined in the previous chapters, sizes and shapes are affected by different factors. While the size is determined where the shape of the spatial structure is specified, the shape is formed through the process of spatial competition.

6.3.2. Differentiated Inputs and Products

Differentiated inputs and products between market areas and supply areas will be examined through the following four cases: non-differentiated inputs and non-differentiated products, non-differentiated inputs and differentiated products, differentiated inputs and non-differentiated products, and differentiated inputs and differentiated products. Differentiated inputs can be brought about by accessibility to the deposit site of inputs, special value and quality at a specific deposit site, or other advantages in cost perspectives. Differentiated products are the same as the notion of product differentiation in conventional economic theory.

6.3.3. External Trade Opportunities

Four patterns of external trade opportunities between market areas and supply areas will be considered: non-external trades for both inputs and products, non-external trade for inputs and some external trade for products, some external trade for inputs and non-external trade for products, and some external trades for both inputs and products. External trade for inputs occurs when some of the inputs are imported from other regions. Likewise, external trade for products takes place when some of the outputs are exported to other regions.

6.4. Basic Components of the Integrated Framework Analysis

This section will develop how duality theory is applied to the integrated framework of market-area analysis and supply-area analysis. First, a simple aspatial model will be examined with regard to the framework of the alternative duality theory. Second, the alternative spatial duality theory (SDT) will be applied to the analysis. Third, the relationship between the dual problem and the production function in the context of internal and external economies will be examined. Finally, an integrated framework analysis will be demonstrated.

6.4.1. Aspatial Equilibrium Analysis with Duality Theory

As previously examined in Chapter 4, duality theory solves cost function from a given factor cost through production function. We now examine a derivation of cost function from the given factor cost and production function. Figure 6-1 (below) shows a factor cost curve $C(x) = wx + F$, where w, x and F represent the factor price, amount of input, and fixed cost, respectively.

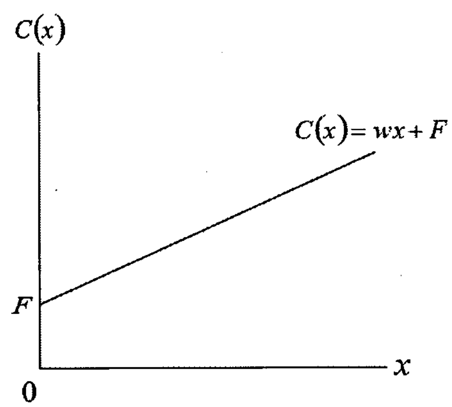


Figure 6-1. Conventional factor cost curve

Figure 6-2 (below) illustrates the conventional production function $q = f(x)$, where q represents the quantity of output.

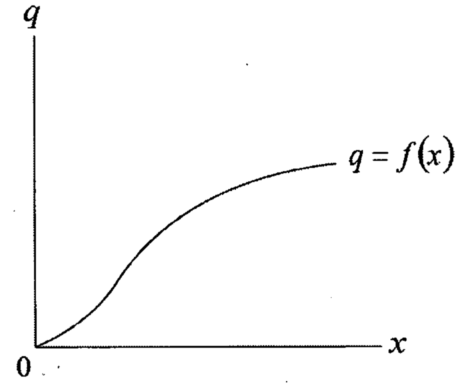


Figure 6-2. Conventional production function

In order to derive a total cost curve as cost function from the combination of the conventional cost curve and the production function, Figure 6-1 (above) and Figure 6-2 (above) should be plotted on the same diagram. Figure 6-3 (below) plots factor cost curve in *Phase* (II) and production function in *Phase* (III). The total cost curve in *Phase* (I) is derived from *Phase* (II) through *Phases* (III) and (IV).

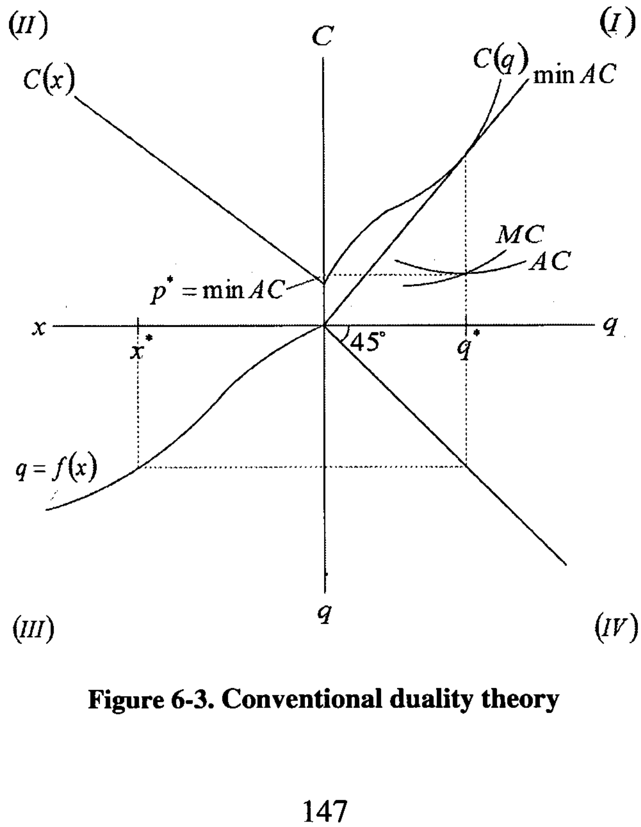


Figure 6-3. Conventional duality theory

In *Phase* (I), average cost AC is derived from total cost curve $C(q)$ and the minimum average cost is achieved where the $\min AC$ line touches cost function $C(q)$. Under the condition of free-entry competition, this output level q^* will be in equilibrium with price level p^*. Duality theory enables the determination of the optimal input level x^* in *Phase* (II) from *Phase* (I) through *Phases* (IV) and (III).

Furthermore, it is possible to analyse a monopoly market, if relevant downward-sloping consumer demand curve AR is plotted in *Phase* (I) as shown in Figure 6-4 (below).

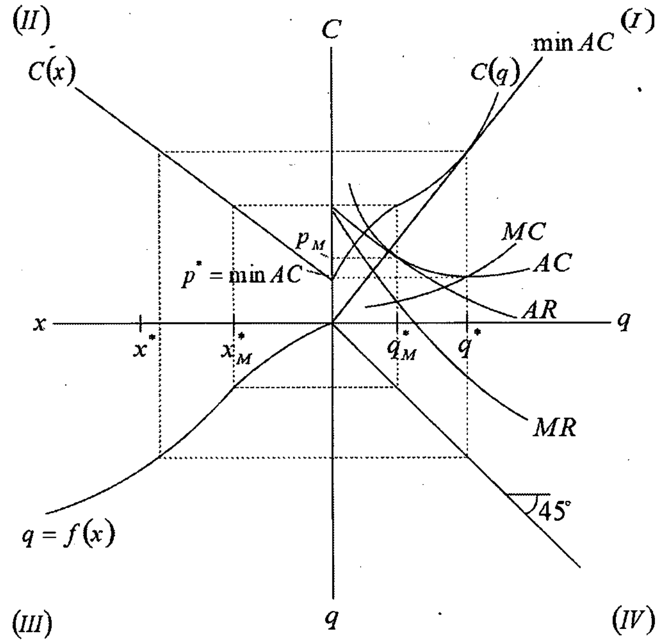


Figure 6-4. Conventional duality theory and monopolistic competition

From the above diagram, equilibrium is available at the output level q_M^*, where marginal revenue MR equals marginal cost MC. Marginal revenue curve MR is derived from the demand curve AR, and marginal cost curve MC is derived from the cost function $C(q)$. In this case, production cost and output price will be p_M. Furthermore, the optimal amount of input

x_M^* can be derived in *Phase* (II) from *Phase* (I) through *Phases* (IV) and (III).

It becomes clear that the equilibrium quantity of output q_M^* and input x_M^* under the condition of imperfect competition, is less than the free-entry competitive case q^* and x^*. In addition, the production cost is lower and the output price is higher due to the lower quantity of output levels. In order to show these relationships in comparative static methods, it is necessary to express the production function in a quadratic form for reasons of simplicity. The examination begins with a simple case. Let us assume that 2 units of input x are required in order to produce an output q, 4 units of x in order to produce $2q$, and 9 units of x in order to produce $3q$, and so on. There are no other relevant costs for this production apart from factor price w. In these circumstances, the production function $q = f(x)$ becomes $q = \sqrt{x}$ as commonly approximated, and the square root of the technical transformation exists for the production process according to the input-output ratio. Now suppose also that the factor cost curve $C(x)$ is expressed as:

$$C(x) = wx + F \qquad (6\text{-}1)$$

In the above equation, w = unit factor price, x = amount of input and F = fixed cost. As $q = \sqrt{x}$, this expression can be re-expressed as $x = q^2$. x can then be substituted into the above equation so that total cost $C(q)$ is derived from the following equation:

$$C(q) = wq^2 + F \qquad (6\text{-}2)$$

Average cost $AC(q)$ is derived from the above equation by dividing it by q:

$$AC(q) = \frac{C(q)}{q} = wq + \frac{F}{q} \qquad (6\text{-}3)$$

Under the condition of perfect competition, the optimal output level q^* is determined at the point where the average cost reaches a minimum. The value q^* will be solved by taking derivatives of $AC(q)$:

$$\frac{\partial AC(q)}{\partial q} = w - \frac{F}{q^2} = 0 \qquad (6\text{-}4)$$

$$wq^2 - F = 0$$

$$q^2 = \frac{F}{w}$$

$$q^* = \sqrt{\frac{F}{w}} \qquad (6\text{-}5)$$

This is the optimal production scale to satisfy the requirements of cost minimisation and the equilibrium level under the condition of perfect competition. In order to find the corresponding optimal amount of input x^*, production function $q = \sqrt{x}$ is substituted into the above equation and x^* has two solutions:

$$x^* = \pm\frac{F}{w} \qquad (6\text{-}6)$$

As $w > 0$ and $F > 0$ by the general assumption in conventional economic analysis, x^* will be a unique solution:

$$x^* = \frac{F}{w} \qquad (6\text{-}7)$$

By contrast, for a situation of monopoly, the optimal input level is not derived in a straightforward manner, and market demand conditions are required to be taken into account. As a result, the average revenue curve AR should be introduced on the first stage:

$$AR = a - bq \qquad a > 0 \text{ and } b > 0 \qquad (6\text{-}8)$$

Here, a is a positive constant value and b is a slope of this curve. Total revenue TR is the multiplied average revenue AR by output level q:

$$TR = AR \cdot q = (a - bq)q \qquad (6\text{-}9)$$

Marginal revenue MR is a partial derivative of total revenue TR with respect to q:

$$MR = \frac{\partial TR}{\partial q} = a - 2bq \qquad \textbf{(6-10)}$$

From Equation (6-2), likewise marginal cost MC is a partial derivative of cost function $C(q)$ with respect to q:

$$MC = \frac{\partial C(q)}{\partial q} = 2wq \qquad \textbf{(6-11)}$$

The impact of a unit of factor price change on marginal cost is:

$$\frac{\partial MC}{\partial w} = 2q \qquad \textbf{(6-12)}$$

Under the condition of monopoly, the optimal output level q_M^* is a point at which marginal revenue MR equals marginal cost MC. Using Equations (6-10) and (6-11),

$$(MR =)\ a - 2bq = 2wq\ \ (= MC)$$

$$q_M^* = \frac{a}{2(w+b)} \qquad \textbf{(6-13)}$$

Substituting production function $q = \sqrt{x}$ into the above equation, the optimal input level x_M^* is specified:

$$x_M^* = \left[\frac{a}{2(w+b)}\right]^2 \qquad \textbf{(6-14)}$$

From Equations (6-13) and (6-14), it becomes clear that both optimal input and output levels are determined by parameters a, b and factor price w under the quadratic form of the production function. These results can be summarised as follows.

$$\frac{\partial q_M^*}{\partial a} > 0,\ \frac{\partial q_M^*}{\partial w} \text{ and } \frac{\partial q_M^*}{\partial b} < 0 \qquad \textbf{(6-15)}$$

$$\frac{\partial x_M^*}{\partial a} > 0,\ \frac{\partial x_M^*}{\partial w} \text{ and } \frac{\partial x_M^*}{\partial b} < 0 \qquad \textbf{(6-16)}$$

The above indicates that fixed cost has no effect on the derivation of equilibrium under the condition of monopoly, while equilibrium is expressed by the ratio of fixed cost and variable factor price under the condition of free-entry competition. Under the condition of monopoly, the corresponding input level is determined by the index of the technological transformation, factor price, the intercept of the vertical axis and the slope of the market demand curve.

6.4.2. Spatial Equilibrium Analysis with Duality Theory

The analysis will now apply the alternative spatial duality theory (SDT). The above investigation examines the relationship between the input and output of a product by applying duality theory. However, spatial aspects have not been included in the analysis and the approach will now refer to Lösch (1954) in order to introduce spatial economic interpretations. The derivation process of the relationship between quantity of output and market-area radius is illustrated from the $f.o.b.$ distribution freight rate and the individual conventional demand curve. This process enables not only the maximum market-area radius $U(p_1)$ under price level p_1 to be found, but also the optimal market-area radius $u^*(p_1)$ under price p_1 to be specified once the individual conventional demand curve is replaced by the aggregate spatial demand curve.

The individual conventional demand curve can be converted into the aggregate conventional demand curve by the horizontal summation of the individual demand curve, if all consumers have the same demand curve. In location analysis, by contrast, the conversion into the aggregate spatial demand curve cannot be achieved in such a straightforward manner, as not all consumers locate at the same site. However, each location of individuals is affected by the distribution cost but not the individual spatial demand curve. As a result, the aggregate spatial demand curve can be derived from the horizontal summation of all individual spatial demand curves as long as all consumers have the same conditions of indifference curve for commodities. In this way, the derivation process of the relationship between quantity of output and market-area radius can now be replaced by the aggregate spatial demand conditions. This alternative approach enables spatial equilibrium analysis under the condition of spatial monopoly to be examined.

As demonstrated in Lösch (1954), a spatial-monopoly profit-maximising production sustains positive profits, encouraging new entrants into the market. This situation is shown in Figure 6-5 (below) as the combination of price p_1 and output Q_1 with profit level $(p_1 - c_1)Q_1$ under the condition of demand curve AR_1.

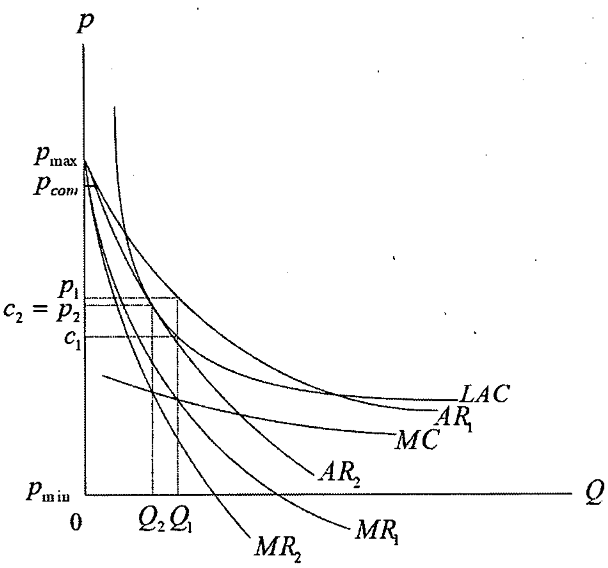


Figure 6-5. Demand and cost curves (Source: Denike and Parr (1970), changed some expressions with additional cost and revenue curves)

The entry process is then referred to spatial equilibrium under free spatial competition. This process then causes a reduction in the incumbent firms' consumer share. Correspondingly, the decreased number of consumers shrinks the extent of the demand curve from AR_1 to AR_2. Under these conditions, the long-run spatial equilibrium condition requires an alternative adjustment in the quantity of output at a point where the aggregate spatial demand curve AR_2 touches the long-run average cost curve LAC. At this alternative output level Q_2, price p_2 equals cost c_2 and marginal cost MC equals alternative marginal revenue MR_2. It has not been made clear whether the marginal cost curve has an upward or a downward slope between the total output levels Q_1 and Q_2. However, this is not important as far as the concern is to clarify the equilibrium levels of production.

During the process of the long-run equilibrium, the long-run average cost curve LAC is shifted to adjust the alternative demand curve generated by spatial competition with other firms. The long-run average cost curve LAC is pulled and pushed in eight different directions: towards the north, northeast, east, southeast, south, southwest, west and northwest, as shown in Figure 6-6 (below). These forces may depend on factors of internal and external economies. Although these interactions have not been investigated in established location theory, the analysis may indicate the adjustability for alternative demand conditions as follows.

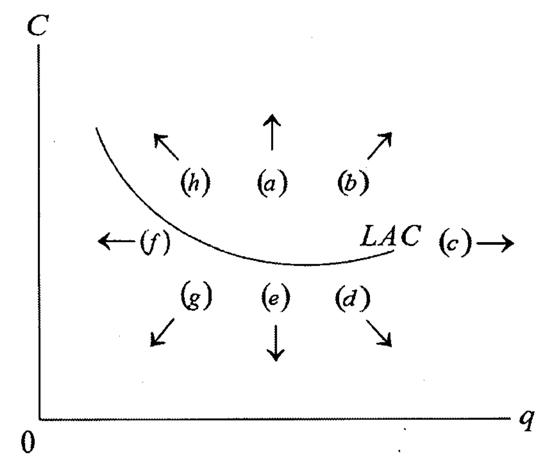


Figure 6-6. Shifts of long-run average cost curve LAC

In the above diagram, the long-run average cost curve LAC can be divided into three indices: technical efficiency X_T, internal economies X_I and external economies X_E. These three indices of economies exert pushing and pulling forces on the long-run average cost curve LAC in a certain direction. The index of technical efficiency X_T represents technological improvements to the production process and assembly transportation. If the aggregate economies are increased by X_T, the movement will be (d) in the above diagram. The index of internal economies X_I shows the impact of internal economies on the structure of long-run average cost curve LAC. If the aggregate internal economies are increased by X_I, the movement of the curve will be (d). By contrast, the increased diseconomies shift this curve to either (a), (b), (g), (f) or (h). The index of external economies X_E, which represents the impact of external economies on the structure of long-run average cost curve LAC, has a similar property. At this stage, it is possible to conduct comparative static analysis with respect to the effect of technological improvement X_T, internal economies X_I and external economies X_E on the market-area radius, $\partial X_T / \partial u$, $\partial X_I / \partial u$ and $\partial X_E / \partial u$ respectively. However, this may have not only the projected result, namely that technological improvement and internal and external economies contribute to an enlargement of the market-area radius, but also the contradictory result showing the opposite effect due to the diagonal movement of the long-run average cost curve LAC to (b) and (g).

We will now solve these spatial problems by comparative-static analysis. Let us assume that an individual firm produces an output q which requires an input x and certain types of technology for processing. The transportation cost for distribution t is expressed through a combination of market area radius u and the $f.o.b.$ transportation rate t. The individual consumer demand q_F is expressed as:

$$q_F = a - b(p + tu) \qquad \textbf{(6-17)}$$

If the market area has a regular shape, the total sales Q are expressed with the maximum radius U and density of demand D as introduced by Mills and Lav (1964) (see Chapter 2). For reasons of simplicity, the analysis in this section applies a simplified circular market-area case. As noted in

Chapter 2 with respect to Denike and Parr (1970), the generalised model in Mills and Lav (1964) has a theoretical problem for particular shapes of market-area formations. At this stage of the analysis, however, the problem may not be expected unless the particular market-area formation is numerically calculated with respect to the dodecagon spatial structure. For a circular market area, the total sales Q are expressed as:

$$Q = D\int_0^{2\pi}\left\{\int_0^U [a - b(p + tu)]udu\right\}d\theta \qquad (6\text{-}18)$$

As a result,

$$Q = D\int_0^{2\pi}\int_0^U [(a - bp - btu)ududθ]$$

$$= D\pi U^2\left(a - bp - \frac{2}{3}btU\right) \qquad (6\text{-}19)$$

As the symbol U expresses the maximum radius of the market area, consumer demand q_F in Equation (2-39) becomes zero at U and price p is specified as follows:

$$a - bp - btU = 0$$

$$p = \frac{a}{b} - tU \qquad (6\text{-}20)$$

In order to find total sales Q, Equation (6-20) is substituted into Equation (6-19):

$$Q = D\pi U^2\left(a - b\left(\frac{a}{b} - tU\right) - \frac{2}{3}btU\right)$$

$$= \frac{1}{3}Dbt\pi U^3 \qquad (6\text{-}21)$$

Total revenue TR is defined by $p \cdot Q$:

$$TR = p \cdot Q = \left(\frac{a}{b} - tU\right)\left(\frac{1}{3}Dbt\pi U^3\right)$$

$$= \frac{1}{3}Dt\pi U^3(a - btU) \qquad (6\text{-}22)$$

Marginal revenue MR is solved by the above equation of the partial derivative with respect to U:

$$MR = \frac{\partial TR}{\partial U} = \frac{1}{3} Dt\pi U^2 (3a - 4btU) \qquad \textbf{(6-23)}$$

Total cost TC is marginal cost MC multiplied by total output Q. In order to find marginal cost, the spatial factor cost must be derived, and that should be a function of output q. The spatial factor cost requires some additional elements in the conventional factor cost (6-1) with respect to the amount of input x and factor price w, namely assembly transportation rate τ and assembly transportation terminal cost F_τ. As a result, the spatial factor cost $C(x)$ becomes:

$$C(x) = (1+\tau)wx + (F + F_\tau) \qquad \textbf{(6-24)}$$

The combination of this equation and the production function will enable the relevant spatial cost function to be obtained. Let us assume that the production function is given as:

$$q = \frac{1}{k}\sqrt{x} \qquad \textbf{(6-25)}$$

where k $(k \geq 1)$ represents an index of technical transformation during the production process. This represents a quantitative transformation of input into output. Solving this production function by input x,

$$x = (kq)^2 \qquad \textbf{(6-26)}$$

The above equation can be substituted into Equation (6-24) so that the cost structure now becomes a function of output q:

$$C(q) = (1+\tau)w(kq)^2 + (F + F_\tau) \qquad \textbf{(6-27)}$$

The optimal market-area radius can be solved by the combination of marginal cost and marginal revenue under the condition of spatial monopoly. In order to examine this combination, the above expression (6-27) is required to transform the quantity of output q into market-area radius u. The relationship between the quantity of output q into market-area radius u can be expressed as:

$$u = \sqrt{\frac{q}{\mu\pi}} \qquad (6\text{-}28)$$

where μ = a constant. Applying this conversion, Equation (6-27) becomes a function of market-area radius u.

$$C(u) = (1+\tau)wk^2\mu^2\pi^2u^4 + (F + F_\tau) \qquad (6\text{-}29)$$

Marginal cost can also be expressed as a function of market-area radius u:

$$MC(u) = \frac{\partial C(u)}{\partial u} = 4u^3(1+\tau)wk^2\mu^2\pi^2 \qquad (6\text{-}30)$$

For reasons of simplicity, the density of demand D is assumed to be $D = 1$ in marginal revenue (6-23). The optimal market-area radius u^* becomes:

$$u^* = \frac{3at}{4\left(bt^2 + 3k^2\mu^2\pi(1+\tau)w\right)} \qquad (6\text{-}31)$$

As all variables are positive in value, it can be specified that $u^* > 0$. Applying the formula (6-28), the optimal quantity of output q^* will be:

$$q^* = \frac{9a^2\mu\pi t^2}{16\left(bt^2 + 3k^2\mu^2\pi(1+\tau)w\right)^2} \qquad (6\text{-}32)$$

which satisfies $q^* > 0$. The optimal amount of input x^* is derived from the combination of the above result and Equation (6-26):

$$x^* = \frac{81a^4\mu^2\pi^2t^4}{256\left(bt^2 + 3k^2\mu^2\pi(1+\tau)w\right)^4} \qquad (6\text{-}33)$$

which satisfies $x^* > 0$. At this stage, it should be suggested that the inclusion of other spatial economic factors will be required in the analysis. These factors can be additionally contained in the structure of factor cost and production function, as will be shown in the following sections. These factors have important roles when firm location is not determined in a straightforward manner. The existing market-area analysis and supply-area analysis implicitly assume that a plant location is situated at the centre of an area. However, there are a number of other cases in which stages of the production process are separated across the plain. In general, disintegrated production operation causes excess cost burdens with respect to the spatially constrained economies to the firm. It is unreasonable for firms to separate their processing stages. However, the separation of branch plants can be suggested when the level of transportation costs of outputs is at a remarkably high level and finishing plants are separately established across the market area, or where there are certain disadvantageous cost factors in the region of the production plant.

The former hypothesis is implausible, as outputs are required to be distributed from each plant not only to their own regions, but also to the assembly plant. This could entail long shipping distances, as will be observed in the following chapter. The latter case involving disadvantageous cost factors seems more straightforward. These factors can be referred to as urbanisation diseconomies if the centre of the area is a metropolitan area and the assembly plant is located this area. In this case, certain urbanisation economies may also be obtained. However, disadvantages such as high land or property rates, congestion and pollution, should also be taken into account in addition to the above stated advantageous factors. The firm may decide to separate a part of processing to other locations if the operational cost at the assembly plant exceeds the additional operational costs and relevant transportation costs of the less efficient separated locations. These types of arguments can be applied to a general case which will be examined in the following chapter as a complex case.

It should be noted that one more theoretical problem arises at this stage. As shown in Equation (6-26), this analysis assumes that there is a technological constraint in the form of a technical transformation between input and output during processing. These constraints can be measured by internal and external economies as examined in the previous chapter. As shown later, the certain part of internal economies to the firm can be observed within the framework of conventional production function. However, some other parts of internal economies and majority parts of external economies cannot be contained in this framework. In order to include whole relevant economic factors, the framework of the conventional production function will be extended in the following part. In addition, some of the internal economies, i.e., economies of scale, have already been included in the framework of the conventional production function. External economies and other parts of internal economies can also be accommodated into the production function if these are related to technological aspects. Otherwise, the remaining economies which do not relate to technological aspects, but relate to pecuniary aspects, can be treated as an additional part of factor cost. These identifications between technological and pecuniary types will be introduced in the latter part of this section.

6.4.3. Duality Theory and Spatial Production Function

In order to revise the structure of production functions in a spatial context, the relationship between production and cost functions should initially be examined. This relationship is systematically analysed by duality theory in Shephard (1953). Duality theory shows the following theoretical interactions: that the input is a function of its relevant production function and that the production function is a function of the cost function. As a result, input x is a function of total cost $C(q)$ through production function $q = f(x)$ and is expressed as follows:

$$C = C(q) \qquad \textbf{(6-34)}$$

$$q = f(x) \qquad \textbf{(6-35)}$$

As a result,

$$C = f(q, x) \qquad \textbf{(6-36)}$$

In the analysis of market areas, cost function becomes a function of the maximum market area radius U, as expressed in Equation (6-37). In addition, as Parr (1993a) demonstrates, input x is a function of the supply-area radius s. These relations are connected as the following functions:

$$C = f(U) \qquad \textbf{(6-37)}$$

$$U = f(u) \quad (6\text{-}38)$$

$$u = f(q) \quad (6\text{-}39)$$

$$q = f(x) \quad (6\text{-}40)$$

$$x = f(s) \quad (6\text{-}41)$$

The above functions can be expressed in an integrated form:

$$C = f(U, u, q, x, s) \quad (6\text{-}42)$$

However, one argument has been left in the conventional field of spatial equilibrium analysis. According to the categorisation in Parr (2002a), the conventional production function solely refers to the internal dimensions of the firm. In addition, spatially unconstrained or constrained cases are not distinguished from each other. Meade (1952) examines the inclusion of external economies in the conventional production function for a case in which there are two indirectly related economic organisations. The example he gives concerns apple-farmers who have their apples fertilised by bees, and bee-keepers who are provided with food for the bees in the apple farm. The results show that the alternative production function contains not only functions of inputs for a single firm, but also functions of inputs and quantities of products relating to other firms. He argues that external economies are not included in the conventional production function, and introduces the following alternative production function between two indirectly related firms

$$q_i = f_i(L_i, K_i, L_j, K_j, q_j) \qquad (i \neq j) \quad (6\text{-}43)$$

where f_i is not necessarily homogeneous to the first degree. This production function shows that a quantity of production is specified not only by the inputs and technical factors of a single firm, but also by the inputs, output levels and technical factors of other firms if there is a certain extent of external economies.

For a single input case, the above expression (6-43) can be expressed more simply:

$$q_i = f_i(x_i, x_j, q_j) \qquad (i \neq j) \qquad \textbf{(6-44)}$$

In our analysis, there are more than two firms. Moreover, the external economies are not brought solely by particular indirectly related firms, as examined in the case of the two specific firms mentioned in Meade. Thus, x_j and q_j cannot be stated in a generalised form. As a result, the external economies will be simply expressed as A in this analysis:

$$q_i = f_i(x_i, A) \qquad \textbf{(6-45)}$$

From this expression, it can be stated that spatially unconstrained and constrained external economies are not included in the condition of the conventional production function. In addition, these economies should be added in between market-area analysis and supply-area analysis as the conventional production function is situated between the framework of input and output of production. The reason of necessity for including these economies is that the core focus of location theory is on the interaction to a firm from the economic activity of other firms and industries. This notion of externality cannot be removed from market-area analysis and supply-area analysis, although these approaches have been excluding them for reasons of simplicity. As a result, the spatially unconstrained and constrained external economies A should be included in the integrated expression (6-42):

$$C = f(U, u, q, A, x, s) \qquad \textbf{(6-46)}$$

Although the above system of function contains all of the relevant spatial economic factors within a single framework, this examination will initially suggest a bisected production function analysis. One is a production function which describes internal economies while the other describes external economies.

The input-output relation in spatial analysis between output q and input x will assume that $q = f(x, A)$ and this can be expressed independently as

$$q = f(x, A) = \left\lfloor f^{\text{int}}(x), f^{ext}(x) \right\rfloor \qquad \textbf{(6-47)}$$

where $q = f^{\text{int}}(x)$ represents the conventional production function and $q = f^{ext}(x)$ shows the external production function. While the position of the conventional production function is not stated in relevant literature, it should be noted that it cannot have a locus beyond the 45° additional line by the general laws of economics, as shown in Figure 6-7 (below).

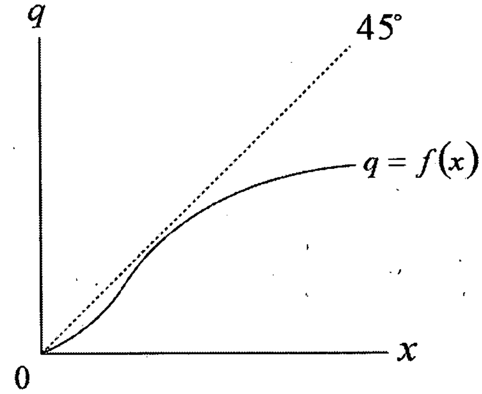


Figure 6-7. The position of the conventional production function

As this analysis examines an individual firm, agglomeration economies may not directly be contained in the relevant cost structure. However, the following interpretation should be considered. Let us assume that this firm produces beer in exclusive market conditions. This firm would not normally have any agglomeration economies. However, it is possible to consider a case in which there are some other industries, such as the wine, whisky or soft drinks industry, with which they share bottles and storehouses within a region. In this case, the argument can be expanded to suggest that the analysis of a single firm can observe the relationship between its own operation and the relevant economies which are obtained from beyond their economic activity. In this way, certain types of localisation economies, urbanisation economies, activity-complex economies, or spatially constrained internal economies to the firm, can be observed in a single-firm investigation.

In order to combine the internal production function and the external production function, let us assume that these production functions,

$q = f^{\text{int}}(x)$ and $q = f^{ext}(x)$ respectively, have the following particular shapes:

$$\text{for } q = f^{\text{int}}(x): \qquad q = x^{0.4} \qquad \textbf{(6-48)}$$

$$\text{for } q = f^{ext}(x): \qquad q = x^{0.45} \qquad \textbf{(6-49)}$$

The alternative production function $q = f(x)$ will be formed as:

$$[q = f(x)] = [q = f^{\text{int}}(x), q = f^{ext}(x)] \qquad \textbf{(6-50)}$$

Equations (6-48) and (6-49) can be combined into the formulation (6-50) as follows:

$$[q = f(x)] = x^{\frac{0.4+0.45}{2}} = x^{0.425} \qquad \textbf{(6-51)}$$

As a result, this can be generalised using coefficients ρ $(0 < \rho < 1)$ for $q = f^{\text{int}}(x)$ and ω $(0 < \omega < 1)$ for $q = f^{ext}(x)$ as:

$$[q = f(x)] = x^{\frac{\rho+\omega}{2}} \qquad \textbf{(6-52)}$$

Spatial equilibrium of market areas and supply areas will now be examined using the formulation of the spatial production function (6-52). Applying the assumption of production function (6-28), the relationship between quantity of output q and input x becomes:

$$q = \frac{1}{k} x^{\frac{\rho+\omega}{2}} \qquad \textbf{(6-53)}$$

Although the relationship between output q and input x is expressed in the above equation, there is a difficulty concerning the generalisation of the spatial production function caused by the mathematics of combining the different power functions of ρ and ω. However, these two different production functions share the same variables, namely quantity of output q and input x. As a result, it is possible to draw these two functions in one figure as shown in Figure 6-8 (below).

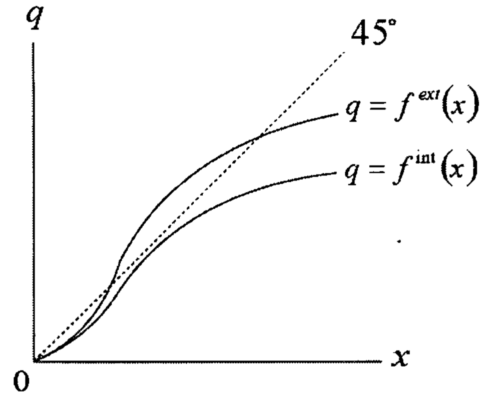


Figure 6-8. Internal economies, external economies and production functions

The curve $q = f^{\text{int}}(x)$ shows the internal production function, and this curve must be located below the 45° additional line as demonstrated with Figure 6-8 (above). By contrast, the external production function $q = f^{ext}(x)$ can be located above the 45° additional line in some areas. These areas represent the positive benefit of the external economies. As these two elements are both situated between market area radius u and supply area radius s, these can be added vertically and the spatial production function $q = f(x)$ will be illustrated as Figure 6-9 (below).

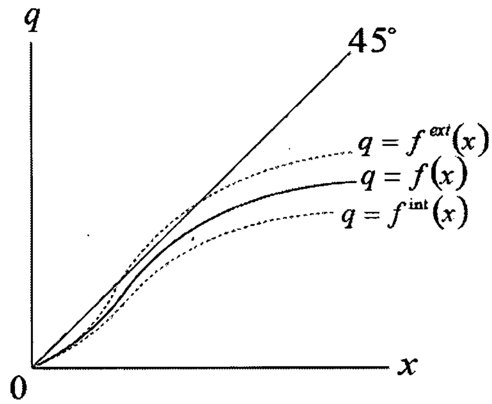


Figure 6-9. Derivation of the spatial production function

Under the full consideration of internal and external economies, the production function $q = f(x)$ is derived by this procedure and can be stated as $q = (1/k)x^{\varphi}$ $(0 < \varphi \leq 1)$. However, it should be noted that this is a technological part of internal and external economies and there is another part referred to as the pecuniary type, according to Meade (1952) and Scitovsky (1954). This latter type of economy can be contained in a part of the spatial factor cost as will be shown in the following part. Although they do not suggest further expansion of this analysis, it can be applied to the Cournot duopoly model which states that one's profit relies on not only one's own quantity of output but also another's quantity of output. If they choose the reasonable strategy for both firms by observing reaction functions of the other firm, a bargaining solution in Nash (1950; 1953) should be taken into account as stated in the previous chapter.

6.4.4. Production Function and Input-Output Framework

In order to transform duality theory into location theory with internal and external economies, the relationship between input x and output q must be reconsidered. For inputs, it is necessary to show how the alternative factor cost curve is formed in spatial analysis. First, as previously examined, the total assembly cost $C(x)$ can be expressed in an extended version of Equation (6-24):

$$C(x) = (1+\tau+\varepsilon)wx + (F + F_\tau + F_\varepsilon) \quad \textbf{(6-54)}$$

This extended equation is developed by adding two additional elements ε and F_ε. These are explanatory variables of internal and external economies which cannot be fitted within the framework of the production function. The element ε represents this additional variable-cost factor and F_ε shows an additional fixed-cost factor. It is assumed that these factors contain transactions cost, communication cost and other relevant explanatory variables of non-technological parts of internal and external economies. The variable factor ε is multiplied by distance s, while fixed factor F_ε does not rely on the amount of inputs and is kept constant. Second, the relationship between input x and supply area radius s can be expressed as:

$$x = \phi\pi s^2 \qquad (6\text{-}55)$$

where ϕ is a constant. The relationship between factor cost and supply-area radius becomes:

$$C(s) = (1+\tau+\varepsilon)w\phi\pi s^2 + (F + F_\tau + F_\varepsilon) = \lambda s^2 + FC \qquad (6\text{-}56)$$

where $\lambda(>0) = (1+\tau+\varepsilon)w\phi\pi$ and $FC(>0) = F + F_\tau + F_\varepsilon$. As a result, the above relationship can be illustrated in Figure 6-10 (below).

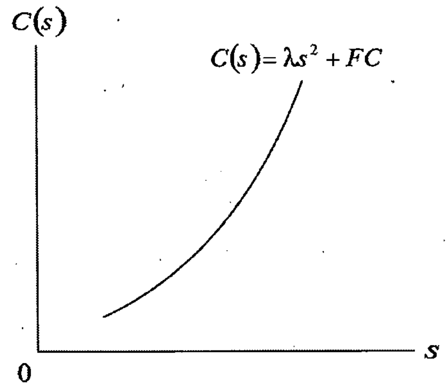


Figure 6-10. Factor cost curve and supply-area radius

For output, it is necessary to demonstrate how to convert quantity of output into market-area radius in spatial duality analysis. Spatial input and production process are connected with respect to quantity of output q. As expressed in Equation (6-28), the relationship between quantity of output q and market-area radius u can be illustrated in Figure 6-11 (below).

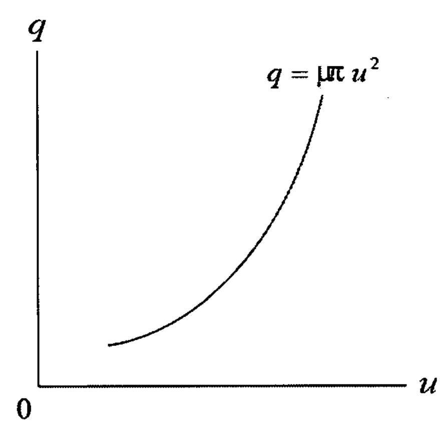


Figure 6-11. Market area radius u and quantity of output q

By combining these interpretations with the spatial production function, an integrated framework of spatial analysis can be demonstrated in Figure 6-12 (below).

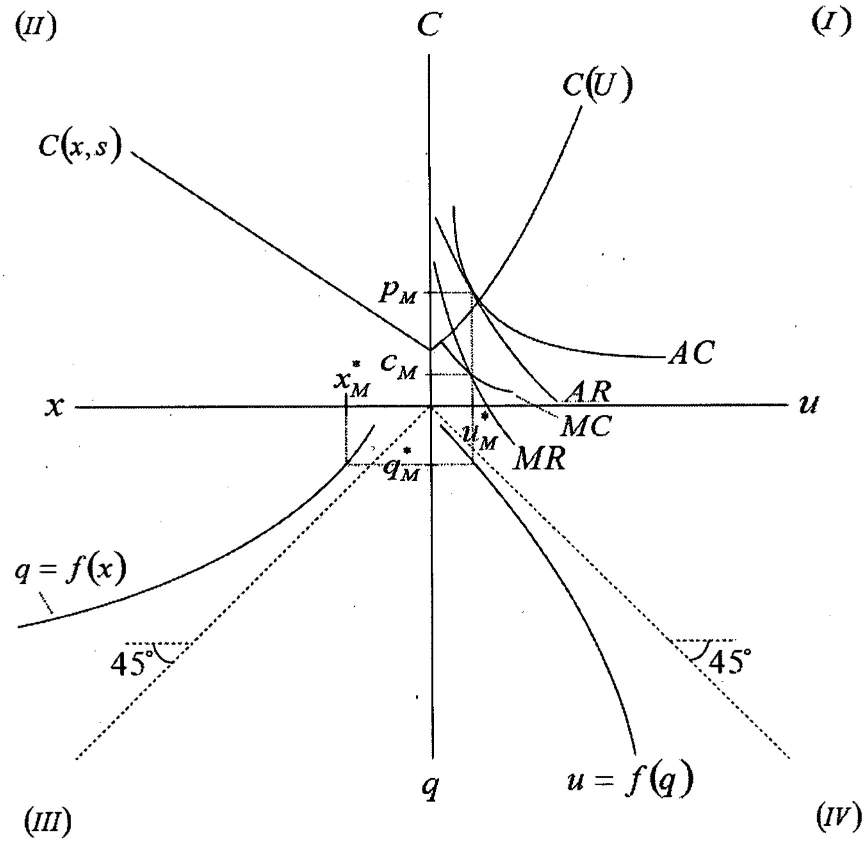


Figure 6-12. Integrated framework spatial analysis

In *Phase* (I) of the above diagram, the spatial cost function $C(U)$ is derived from the spatial input-factor cost curve $C(x,s)$ in *Phase* (II) through *Phases* (III) and (IV). *Phase* (III) represents the spatial production function which is derived in the previous part in Figure 6-9 (shown earlier). *Phase* (IV) illustrates the relationship between market area radius u and quantity of output q as demonstrated in Figure 6-11 (above).

The relevant spatial demand curve can also be added in *Phase* (I) in this diagram. The spatial demand curve *AR* determines the marginal revenue curve MR. As examined in Chapter 2, spatial monopoly equilibrium is achieved at the point at which marginal revenue MR equals marginal cost MC. Marginal cost MC is derived from spatial cost function $C(U)$, and average cost AC is also derived from this spatial cost function. The spatial equilibrium market price p_M under these conditions is where the spatial demand curve AR and the relevant average cost curve AC connect with each other. Moreover, this market-area radius satisfies $MR = MC$ and this is the optimal market-area radius u_M^*. Applying the integrated framework of market-area analysis and supply-area analysis, the optimal amount of input x_M^* is derived through the spatial production function.

In terms of agglomeration economies, it is possible to observe the effect of these economies - with respect to the pecuniary type which appears in *Phase* (II) and the technological type which appears in *Phase* (III) - on the required amount of inputs for profit maximisation and cost minimisation within the firm in market-area analysis and supply-area analysis. It can be projected that the firm requires less inputs and supply areas if either type of agglomeration economy is more readily available. By contrast, the firm requires more inputs and larger supply areas if either type of these economies is less readily available. The relevant market area is observed under the condition of these economic factors and the given spatial demand curve. Regarding Figure 6-9 (shown earlier), the spatial production function is expressed as:

$$q = \frac{1}{k} x^{\varphi} \qquad (0 < \varphi \leq 1) \qquad \textbf{(6-57)}$$

As commonly approximated, let us assume that $\varphi = 0.5$ and substitute the above equation which is solved with respect to x into Equation (6-54):

$$C(q) = (1 + \tau + \varepsilon) w (kq)^2 + (F + F_\tau + F_\varepsilon) \qquad \textbf{(6-58)}$$

Applying the expression (6-28) to the above equation:

$$C(u) = (1 + \tau + \varepsilon) w k^2 \mu^2 \pi^2 u^4 + (F + F_\tau + F_\varepsilon) \qquad \textbf{(6-59)}$$

Marginal cost MC is a partial derivative of the above equation with respect to u:

$$MC(u) = \frac{\partial C(u)}{\partial u} = 4u^3(1+\tau+\varepsilon)wk^2\mu^2\pi^2 \qquad (6\text{-}60)$$

As demonstrated earlier, the results can be shown as follows:

$$u^* = \frac{3at}{4(bt^2 + 3k^2\mu^2\pi(1+\tau+\varepsilon)w)} \qquad (6\text{-}61)$$

$$q^* = \frac{9a^2\mu\pi t^2}{16(bt^2 + 3k^2\mu^2\pi(1+\tau+\varepsilon)w)^2} \qquad (6\text{-}62)$$

$$x^* = \frac{81a^4\mu^2\pi^2t^4}{256(bt^2 + 3k^2\mu^2\pi(1+\tau+\varepsilon)w)^4} \qquad (6\text{-}63)$$

In addition, the optimal supply-area radius s^* can also be derived from the combination of the above expression and Equation (6-55)

$$s^* = 0.5625\left(\frac{a^4k^2\mu^2\pi t^4}{\phi(bt^2 + 3k^2\mu^2\pi(1+\tau+\varepsilon)w)^4}\right)^{\frac{1}{2}} \qquad (6\text{-}64)$$

In comparing the above results with the results in the previous section, which exclude the notion of spatial production function and external economies, it becomes clear that the index of pecuniary external economies ε have certain effects on the determination of the optimal quantity of output q^*, market-area radius u^* and amount of input x^*. The other additional spatial factor F_ε, by contrast, has no impact within the framework of the comparative-static method. This factor ε has a certain impact on the value of total cost as an independent value of the structure of cost function. These results will be tested by the comparative-static analysis in Section 6.6.

6.5. Examinations of the Integrated Framework Model

At this stage of the analysis, it is possible to examine the relationship between market areas and supply areas through changes in particular relevant variables. Five significant cases can be shown from Figure 6-12 (above), two cases in *Phase* (II) and one case each in *Phases* (I), (III) and (IV).

6.5.1. Changes in Spatial Demand Conditions

This case is observed in the change of the spatial demand curve in *Phase* (I) in the diagram. In the short run, the marginal cost curve cannot change its shape and the optimal supply-area radius will increase when the demand curve enlarges. More precisely, the enlargement of the demand curve increases the optimal market-area radius and the optimal quantity of outputs. The increase of the output level expands the amount of input and the relevant supply area. In the long run, by contrast, the cost curves can be moved to adjust the modified demand curve, until the average cost curve touches the demand curve under the condition of spatial free-entry competition, as shown in Denike and Parr (1970).

6.5.2. Changes in Assembly Transportation Rate and Pecuniary External Economies

This case shows a slope change of spatial factor cost curve in *Phase* (II) of the diagram. This increases not only the formation of the spatial factor cost curve itself, but also the structure of spatial cost function in *Phase* (I) through *Phases* (III) and (IV). An increase in the spatial cost function changes the formation of marginal cost. In this way, the increase of assembly transportation rate τ or the index of pecuniary external economies ε changes the shape of the marginal cost curve and reduces the size of the optimal market-area radius. This eventually reduces the size of the supply area in *Phase* (II) of the diagram through *Phases* (IV) and (III).

6.5.3. Changes in Fixed Cost, Terminal Cost and the Explanatory Fixed Factor

In this case, the height of the spatial factor cost curve in *Phase* (II) changes in a parallel movement. An increase of fixed cost F, terminal cost F_τ or explanatory fixed factor F_ε not only changes the height of the spatial factor cost but also increases the level of the spatial cost function through *Phases* (III) and (IV). This increases the height of the marginal cost curve, and the optimal-market radius will be reduced. In addition, these changes eventually reduce the size of the supply area through *Phases* (IV) and (III) of the diagram.

6.5.4. Changes in Spatial Production Function

This is where the slope of spatial production function is increased or decreased by the availability of more advanced production technologies. The former case achieves lower spatial cost function $C(U)$ in *Phase* (I), lower average cost AC and marginal cost MC. As a result, the optimal market-area radius increases, but the size of the supply area is not necessarily increased. This can be achieved by a technological improvement. In the opposite case, a decreased level of technology changes the shape of the spatial production function and increases the spatial cost function $C(U)$ in *Phase* (I) through *Phase* (IV). This causes a reduction of the optimal market-area radius and the optimal quantity of output is reduced. Despite the reduction of the output level, the relevant amount of input and the supply area may increase in this case as the technology level requires more inputs than the previous level.

6.5.5. Changes in Shapes of the Market Area

This is a case in which the spatial configuration of the market area is changed. It affects the shape of the market-area spatial configuration curve in *Phase* (IV) of the diagram. Figure 6-12 (shown earlier) represents a circular case. The regular hexagonal case will be closer to the vertical q axis as the output level increases, and the truncated circular case is situated between these two cases. These shifts affect the structure of the spatial cost functions $C(U)$ in *Phase* (I) and the optimal market-area radius will be changed according to the condition of spatial competition. Furthermore, the optimal size of the supply area is also modified through changes in the optimal quantity of output and the amount of input.

6.6. The Comparative-Static Analysis

This section will demonstrate comparative-static analysis according to the results which are obtained in the previous sections. We can observe the impact of a change in factor price w, distribution transportation rate t, assembly transportation rate τ, index of pecuniary type of economies ε, index of technological transformation k and index of spatial transformation μ on the optimal market-area radius u^*, quantity of output q^*, amount of input x^* and supply-area radius s^*.

6.6.1. The Impact on the Optimal Market-Area Radius

First, the impact of changes in the above stated variables, on the optimal market-area radius u^* is shown as follows.

$$\frac{\partial u^*}{\partial w} = -\frac{9ak^2\mu^2\pi t(1+\tau+\varepsilon)}{4\left(bt^2+3k^2\mu^2\pi(1+\tau+\varepsilon)w\right)^2} < 0 \tag{6-65}$$

$$\frac{\partial u^*}{\partial t} = -\frac{3abt^2}{2\left(bt^2+3k^2\mu^2\pi(1+\tau+\varepsilon)w\right)^2} + \frac{3a}{4\left(bt^2+3k^2\mu^2\pi(1+\tau+\varepsilon)w\right)} \neq 0 \tag{6-66}$$

$$\frac{\partial u^*}{\partial \tau} = -\frac{9ak^2\mu^2\pi tw}{4\left(bt^2+3k^2\mu^2\pi(1+\tau+\varepsilon)w\right)^2} < 0 \tag{6-67}$$

$$\frac{\partial u^*}{\partial \varepsilon} = -\frac{9ak^2\mu^2\pi t w}{4\left(bt^2 + 3k^2\mu^2\pi(1+\tau+\varepsilon)w\right)^2} < 0 \qquad (6\text{-}68)$$

$$\frac{\partial u^*}{\partial k} = -\frac{9ak\mu^2\pi t(1+\tau+\varepsilon)w}{2\left(bt^2 + 3k^2\mu^2\pi(1+\tau+\varepsilon)w\right)^2} < 0 \qquad (6\text{-}69)$$

$$\frac{\partial u^*}{\partial \mu} = -\frac{9ak^2\mu\pi t(1+\tau+\varepsilon)w}{2\left(bt^2 + 3k^2\mu^2\pi(1+\tau+\varepsilon)w\right)^2} < 0 \qquad (6\text{-}70)$$

In the above result, the impact of a change in distribution transportation rate t has an indefinite sign (either $\partial u^* / \partial t > 0$ or $\partial u^* / \partial t < 0$). As examined in Chapter 2, this must be $\partial u^* / \partial t < 0$. Thus, the following additional sufficient condition will be provided:

$$\frac{3abt^2}{2\left(bt^2 + 3k^2\mu^2\pi(1+\tau+\varepsilon)w\right)^2} > \frac{3a}{4\left(bt^2 + 3k^2\mu^2\pi(1+\tau+\varepsilon)w\right)} \qquad (6\text{-}71)$$

6.6.2. The Impact on the Optimal Quantity of Output

Second, the impact of a change in each variable on the optimal quantity of output q^* is shown as follows:

$$\frac{\partial q^*}{\partial w} = -\frac{27a^2k^2\mu^3\pi^2t^2(1+\tau+\varepsilon)}{8\left(bt^2 + 3k^2\mu^2\pi(1+\tau+\varepsilon)w\right)^3} < 0 \qquad (6\text{-}72)$$

$$\frac{\partial q^*}{\partial t} = -\frac{9a^2b\mu\pi t^3}{4\left(bt^2 + 3k^2\mu^2\pi(1+\tau+\varepsilon)w\right)^3} + \frac{9a^2\mu\pi t}{8\left(bt^2 + 3k^2\mu^2\pi(1+\tau+\varepsilon)w\right)^2} \neq 0$$

(6-73)

$$\frac{\partial q^*}{\partial \tau} = -\frac{27a^2k^2\mu^3\pi^2t^2w}{8\left(bt^2 + 3k^2\mu^2\pi(1+\tau+\varepsilon)w\right)^3} < 0 \qquad (6\text{-}74)$$

$$\frac{\partial q^*}{\partial \varepsilon} = -\frac{27a^2k^2\mu^3\pi^2t^2w}{8\left(bt^2 + 3k^2\mu^2\pi(1+\tau+\varepsilon)w\right)^3} < 0 \qquad (6\text{-}75)$$

$$\frac{\partial q^*}{\partial k} = -\frac{27a^2k\mu^3\pi^2t^2(1+\tau+\varepsilon)w}{4\left(bt^2 + 3k^2\mu^2\pi(1+\tau+\varepsilon)w\right)^3} < 0 \qquad (6\text{-}76)$$

$$\frac{\partial q^*}{\partial \mu} = -\frac{27a^2k^2\mu^2\pi^2t^2(1+\tau+\varepsilon)w}{4(bt^2+3k^2\mu^2\pi(1+\tau+\varepsilon)w)^3} + \frac{9a^2\pi t^2}{16(bt^2+3k^2\mu^2\pi(1+\tau+\varepsilon)w)^2} \neq 0 \quad (6\text{-}77)$$

In the above results, the impacts of a change in distribution transportation rate t and a change in the index of spatial transformation μ have indefinite signs. The former case can be suggested to have the following additional sufficient condition, as $\partial q^* / \partial t < 0$, regarding Chapter 2.

$$\frac{9a^2b\mu\pi t^3}{4(bt^2+3k^2\mu^2\pi(1+\tau+\varepsilon)w)^3} > \frac{9a^2\mu\pi t}{8(bt^2+3k^2\mu^2\pi(1+\tau+\varepsilon)w)^2} \quad (6\text{-}78)$$

The latter case can be treated in the same manner as the density of demand in this analysis. As a result, the sign must have $\partial q^* / \partial \mu < 0$. In this way, the additional sufficient condition will be:

$$\frac{27a^2k^2\mu^2\pi^2t^2(1+\tau+\varepsilon)w}{4(bt^2+3k^2\mu^2\pi(1+\tau+\varepsilon)w)^3} > \frac{9a^2\pi t^2}{16(bt^2+3k^2\mu^2\pi(1+\tau+\varepsilon)w)^2} \quad (6\text{-}79)$$

6.6.3. The Impact on the Optimal Amount of Input

Third, the impact of a change in each variable on the optimal amount of input x^* is shown as follows:

$$\frac{\partial x^*}{\partial w} = -\frac{243a^4k^4\mu^4\pi^3t^4w}{64(bt^2+3k^2\mu^2\pi(1+\tau+\varepsilon)w)^5} < 0 \quad (6\text{-}80)$$

$$\frac{\partial x^*}{\partial t} = -\frac{81a^4bk^2\mu^2\pi^2t^5}{32(bt^2+3k^2\mu^2\pi(1+\tau+\varepsilon)w)^5} + \frac{81a^4k^2\mu^2\pi^2t^3}{64(bt^2+3k^2\mu^2\pi(1+\tau+\varepsilon)w)^4} \neq 0 \quad (6\text{-}81)$$

$$\frac{\partial x^*}{\partial \tau} = -\frac{243a^4k^4\mu^4\pi^3t^4w}{64(bt^2+3k^2\mu^2\pi(1+\tau+\varepsilon)w)^5} < 0 \quad (6\text{-}82)$$

$$\frac{\partial x^*}{\partial \varepsilon} = -\frac{243a^4k^4\mu^4\pi^3t^4w}{64(bt^2+3k^2\mu^2\pi(1+\tau+\varepsilon)w)^5} < 0 \quad (6\text{-}83)$$

$$\frac{\partial x^*}{\partial k} = -\frac{243a^4k^3\mu^4\pi^3t^4(1+\tau+\varepsilon)w}{32(bt^2+3k^2\mu^2\pi(1+\tau+\varepsilon)w)^5} + \frac{81a^4k\mu^2\pi^2t^4}{128(bt^2+3k^2\mu^2\pi(1+\tau+\varepsilon)w)^4} \neq 0$$

(6-84)

$$\frac{\partial x^*}{\partial \mu} = -\frac{243a^4k^4\mu^3\pi^3t^4(1+\tau+\varepsilon)w}{32(bt^2+3k^2\mu^2\pi(1+\tau+\varepsilon)w)^5} + \frac{81a^4k^2\mu\pi^2t^4}{128(bt^2+3k^2\mu^2\pi(1+\tau+\varepsilon)w)^4} \neq 0$$

(6-85)

In the above results, the impacts of a change in distribution transportation rate t, a change in the index of technological formation k, and a change in the index of spatial transformation μ, have indefinite signs. As the first case should be $\partial x^* / \partial t < 0$ regarding the previous sections, the additional sufficient condition is given as:

$$\frac{81a^4bk^2\mu^2\pi^2t^5}{32(bt^2+3k^2\mu^2\pi(1+\tau+\varepsilon)w)^5} > \frac{81a^4k^2\mu^2\pi^2t^3}{64(bt^2+3k^2\mu^2\pi(1+\tau+\varepsilon)w)^4} \quad \textbf{(6-86)}$$

The second case should have the form $\partial x^* / \partial k < 0$ regarding the previous sections. Thus, the additional sufficient condition becomes:

$$\frac{243a^4k^4\mu^3\pi^3t^4(1+\tau+\varepsilon)w}{32(bt^2+3k^2\mu^2\pi(1+\tau+\varepsilon)w)^5} > \frac{81a^4k^2\mu\pi^2t^4}{128(bt^2+3k^2\mu^2\pi(1+\tau+\varepsilon)w)^4} \quad \textbf{(6-87)}$$

The third case should be $\partial x^* / \partial \mu < 0$ regarding the previous sections. As a result, the following additional sufficient condition is required:

$$\frac{243a^4k^4\mu^3\pi^3t^4(1+\tau+\varepsilon)w}{32(bt^2+3k^2\mu^2\pi(1+\tau+\varepsilon)w)^5} > \frac{81a^4k^2\mu\pi^2t^4}{128(bt^2+3k^2\mu^2\pi(1+\tau+\varepsilon)w)^4} \quad \textbf{(6-88)}$$

6.6.4. The Impact on the Optimal Supply-Area Radius

Finally, the impacts of a change in each variable on the optimal supply-area radius s^* is shown as follows:

$$\frac{\partial s^*}{\partial w} = -\frac{3.375a^4k^4\mu^4\pi^2t^4(1+\tau+\varepsilon)}{\phi\left(\frac{a^4k^2\mu^2\pi t^4}{\phi(bt^2+3k^2\mu^2\pi(1+\tau+\varepsilon)w)^4}\right)^{0.5}(bt^2+3k^2\mu^2\pi(1+\tau+\varepsilon)w)^5} < 0$$

(6-89)

$$\frac{\partial s^*}{\partial t}=-\frac{0.28125\left(-\frac{8a^4bk^2\mu^2\pi t^5}{\phi\left(bt^2+3k^2\mu^2\pi(1+\tau+\varepsilon)w\right)^5}+\left(\frac{4a^4k^2\mu^2\pi t^3}{\phi\left(bt^2+3k^2\mu^2\pi(1+\tau+\varepsilon)w\right)^4}\right)\right)}{\phi\left(\frac{a^4k^2\mu^2\pi t^4}{\phi\left(bt^2+3k^2\mu^2\pi(1+\tau+\varepsilon)w\right)^4}\right)^{0.5}\left(bt^2+3k^2\mu^2\pi(1+\tau+\varepsilon)w\right)^5}\neq 0$$

(6-90)

$$\frac{\partial s^*}{\partial \tau}=-\frac{3.375a^4k^4\mu^4\pi^2t^4w}{\phi\left(\frac{a^4k^2\mu^2\pi t^4}{\phi\left(bt^2+3k^2\mu^2\pi(1+\tau+\varepsilon)w\right)^4}\right)^{0.5}\left(bt^2+3k^2\mu^2\pi(1+\tau+\varepsilon)w\right)^5}<0$$

(6-91)

$$\frac{\partial s^*}{\partial \varepsilon}=-\frac{3.375a^4k^4\mu^4\pi^2t^4(1+\tau+\varepsilon)}{\phi\left(\frac{a^4k^2\mu^2\pi t^4}{\phi\left(bt^2+3k^2\mu^2\pi(1+\tau+\varepsilon)w\right)^4}\right)^{0.5}\left(bt^2+3k^2\mu^2\pi(1+\tau+\varepsilon)w\right)^5}<0$$

(6-92)

$$\frac{\partial s^*}{\partial k}=-\frac{0.28125\left(-\frac{24a^4k^3\mu^4\pi^2t^4(1+\tau+\varepsilon)w}{\phi\left(bt^2+3k^2\mu^2\pi(1+\tau+\varepsilon)w\right)^5}+\left(\frac{2a^4k\mu^2\pi t^4}{\phi\left(bt^2+3k^2\mu^2\pi(1+\tau+\varepsilon)w\right)^4}\right)\right)}{\left(\frac{a^4k^2\mu^2\pi t^4}{\phi\left(bt^2+3k^2\mu^2\pi(1+\tau+\varepsilon)w\right)^4}\right)^{0.5}}\neq 0$$

(6-93)

$$\frac{\partial s^*}{\partial \mu}=-\frac{0.28125\left(-\frac{24a^4k^4\mu^3\pi^2t^4(1+\tau+\varepsilon)w}{\phi\left(bt^2+3k^2\mu^2\pi(1+\tau+\varepsilon)w\right)^5}+\left(\frac{2a^4k^2\mu\pi t^4}{\phi\left(bt^2+3k^2\mu^2\pi(1+\tau+\varepsilon)w\right)^4}\right)\right)}{\left(\frac{a^4k^2\mu^2\pi t^4}{\phi\left(bt^2+3k^2\mu^2\pi(1+\tau+\varepsilon)w\right)^4}\right)^{0.5}}\neq 0$$

(6-94)

In the above results, the impact of a change in distribution transportation rate t, a change in the index of technological transformation k and a change in spatial transformation μ have indefinite signs. For the first case, this should have $\partial s^*/\partial t<0$ regarding the previous sections. As a result, the additional sufficient condition will be:

$$\frac{8a^4bk^2\mu^2\pi t^5}{\phi\left(bt^2+3k^2\mu^2\pi(1+\tau+\varepsilon)w\right)^5}>\frac{4a^4k^2\mu^2\pi t^3}{\phi\left(bt^2+3k^2\mu^2\pi(1+\tau+\varepsilon)w\right)^4} \qquad \textbf{(6-95)}$$

The second case should have the form $\partial s^*/\partial k<0$ regarding the previous sections. As a result, the sufficient condition becomes:

$$\frac{24a^4k^3\mu^4\pi^2t^4(1+\tau+\varepsilon)w}{\phi(bt^2+3k^2\mu^2\pi(1+\tau+\varepsilon)w)^4} > \frac{2a^4k\mu^2\pi t^4}{\phi(bt^2+3k^2\mu^2\pi(1+\tau+\varepsilon)w)^4} \quad (6\text{-}96)$$

The final case should have the form $\partial s^* / \partial\mu < 0$ regarding the previous sections. As a result, the additional sufficient condition will be:

$$\frac{24a^4k^4\mu^3\pi^3t^4(1+\tau+\varepsilon)w}{\phi(bt^2+3k^2\mu^2\pi(1+\tau+\varepsilon)w)^5} > \frac{2a^4k^2\mu\pi t^4}{\phi(bt^2+3k^2\mu^2\pi(1+\tau+\varepsilon)w)^4} \quad (6\text{-}97)$$

6.6.5. Conclusion of the Comparative-Static Results

This section observes the effect of changes in factor price w, distribution transportation rate t, assembly transportation rate τ, index of the pecuniary type of economy ε, index of technological transformation k, and index of spatial transformation μ on spatial variables of market areas and supply areas, by comparative-static methods. All results have appropriate signs according to the analysis previously examined. While some variables have indefinite signs, these problems are solved by the revision of the previously introduced models. The general findings and discussion of the integrated framework approach will be provided in the following section.

6.7. The Effect of Market Area Changes on the Spatial Structure of Supply Area and Vice Versa

6.7.1. A Change in the Number of Competitors

A change in the number of competitors in a market area affects the shape of the spatial demand curve through a change in the conditions of market competition. An increase in the number of competitors in the market area results in a more restricted capacity of production levels for every individual firm if other economic conditions are assumed to remain constant. As a result, the number of relevant supply areas will be reduced. In this case, economies of large-quantity production and economies of scale are reduced and market price may increase, due to the cost increase, if the relevant average cost curve is downward sloping. For the reverse case, where the number of competitors of input increases, this may cause an increase in the factor price. The increased factor price also increases the spatial cost function. This eventually reduces the size of market areas if other economic conditions are assumed to remain constant.

6.7.2. A Change in Shapes

A change in the shape of market areas affects the formation of the spatial cost function as examined in the change in shapes of market areas in Section 6.5. Regarding a change in the shape of supply areas, this may affect the structure of spatial factor cost. However, the following point is the most important difference between market areas and supply areas: while the shape of market areas concerns the maximisation of revenue, the shape of supply areas concerns more the minimisation of costs than the maximisation of revenue. This can be illustrated by the fact that the circular shape of a supply area forms lower spatial cost function levels than the hexagonal supply area. The truncated-circular case is situated between these two types.

6.7.3. A Change in Differentiated Product or Input Pattern

A change in differentiated output affects the conditions of the spatial demand curve and the examination refers to spatial competition under the condition of product differentiation. In this case, the shape of the marginal revenue curve will be adjusted to the given marginal cost levels. This also modifies the size of the supply area through changes in output and input levels. By contrast, the presence of differentiated inputs solely affects the level of factor price. This will change the spatial factor cost level and the alternative marginal cost will adjust to the given marginal revenue. This determines the optimal market-area radius and the relevant supply-area size is also specified observing the optimal output and input levels.

6.7.4. A Change in External Trade Pattern

A change in the external trade pattern in the market area affects the structure of the *f.o.b.* price system. An increase in external trade opportunities will increase the *f.o.b.* price and this will change the marginal revenue level. The optimal market-area radius becomes smaller and the relevant size of the supply area also reduces through the reduction of optimal output and input levels. By contrast, a change in the external trade pattern in the supply area affects the structure of the spatial factor cost. An increase in external trade opportunities increases not only spatial factor cost, but also the level of the spatial cost function. This reduces the optimal market-area radius and eventually the size of the supply area becomes smaller through the reduction of output and input levels.

6.8. Conclusion

This chapter first composes the input-output framework approach by means of the duality theory. This attempt also demonstrates comparative-static analysis under the conditions of free competition and monopoly. Second, the framework is applied to market-area analysis and supply-area analysis with the introduction of the spatial duality theory. The spatial duality theory includes spatially unconstrained and constrained internal and external economies in the components of the spatial factor cost curve and spatial production function. The impact of these economies on market areas and supply areas is demonstrated by the comparative static analysis. Finally, the relevant economic variables of the integrated framework of market areas and supply areas are examined, with the effect of change in a market area or supply area on the corresponding supply area or market area. As the examination in this chapter is limited in the range of spatial conditions, this integrated framework analysis should be applied to a wide range of economic circumstances. The following chapter will investigate them with eight hypothetical examples in order to complete the integrated framework analysis of market areas and supply areas.

Chapter 7. An Integrated Production-Stage Analysis of Market Areas and Supply Areas

7.1. Introduction

In the previous chapter, internal and external economies are introduced to investigate firm location. However, the previous chapter does not extend the analysis of internal and external economies in terms of spatially unconstrained and constrained factors. In particular, the interaction between firm location and agglomeration economies is not sufficiently investigated. This chapter will extend the integrated framework model with eight representative hypothetical examples to show the effect of spatially constrained economic factors on firm location.

7.2. An Overview of Eight Hypothetical Cases

This section will introduce the eight cases of hypothetical examples which will be examined in the following section in order to apply an integrated framework analysis to various spatial economic patterns. This extended framework will be based on the following case. Let us assume that there is a producer engaged in producing and distributing a product to consumers within a region. In addition, the plant is located at the centre of the region and there is no obstacle across the plain. As a result, the relevant market area and supply area are expanded from the centre of the region if inputs are ubiquitous. In this simplified case, the size of the market area and supply area can be specified by the distribution transportation rate t and the assembly transportation rate τ respectively. As illustrated in Figure 7-1 (below), the size of the market area and supply area are expressed with respect to each radius u and s. The diagram shows a case in which the transportation rate for assembly τ is higher than that of distribution t.

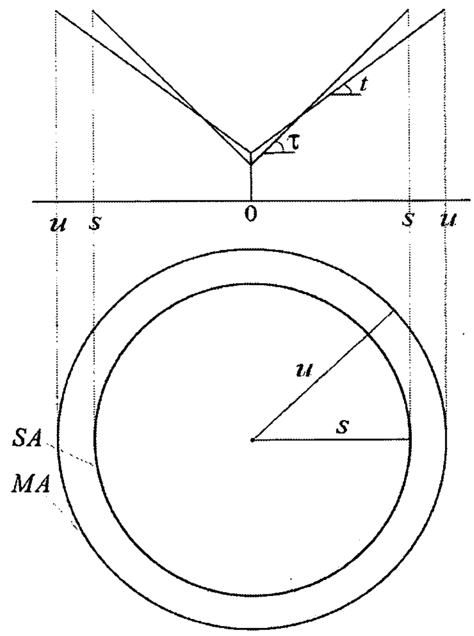


Figure 7-1. Identical centre of market area and supply area

In this way, the analysis will be expanded to various patterns of spatial structures in eight different cases. Although existing market-area analysis and supply-area analysis have not fully taken into account agglomeration economies, these economies have an important role in specifying firm location and the centres of market areas and supply areas. The examination will apply the integrated framework analysis to the following eight cases.

Case I:

This will examine a simple case in which a producer is processing a commodity in his local region and there are no exports or imports. In this case, the maximum market area and supply area have the same size and shape if the following conditions apply: that the size of the region does not exceed the output level which achieves economies of scale of production; and that the region has sufficient capacity of deposit for input. It should be noted that the region is assumed to have the same transportation attributes in both assembly and distribution, in terms of transportation routes and methods as examined in Chapter 4.

Case II:

This will attempt an extension of *Case I.* Whilst the supply area is still limited within a region, the market area expands beyond the region. In this case, the market area has a two-stage distribution cost structure, and the cost curve is divided into two parts at the boundary of each region. At the boundary, there is an additional terminal cost or changing point of transportation rate. As a result, the constant distribution cost structure ends at the boundary of the region. Beyond the boundary, the distribution cost is newly settled and the constant condition for transportation cost is no longer maintained.

Case III:

This case will have eight peripheral finishing plants producing outputs, in addition to an assembly plant at the centre of the economic plain which produces the core element of the product. In general, dispersed division of production processes may cause extra expense due to diseconomies of scale and distant transportation. As a result, this scenario would not seem to have been chosen by a profit-maximising firm with respect to cost saving behaviour. However, such situations can be observed in particular industries by the inclusion of the condition of urbanisation diseconomies.

Case IV:

This will exemplify a complex pattern which is more common in added spatial structures. There are several stages of processing with upstream and downstream linkages between different sections of production. Various elements of agglomeration economies and transportation costs for assembly and distribution will be contained in this case.

Case V:

This case will examine a perfectly overlapping market-area structure. It assumes two independent brands and that their market areas overlap perfectly. The centres of the market areas and supply areas are assumed to be identical, and inputs are shared by both firms.

Case VI:

This case will explore one of the special cases where two brands share the same market area, while owning different centres of distribution. A notable point will be that though they are sharing the same plain, the market areas are not identical between the two firms with respect to the cost minimisation behaviour.

Case VII:

This pattern will show that each market area is exclusively dominated by one brand of product and that both types of area are identical. One possible reason for having these exclusive patterns is that there are gaps between available market-area sizes and the limited production scale of the firm.

Case VIII:

This case will deal with oligopoly competition in *Case VII*. It assumes that there are three independent brands and each market area supplies products exclusively from one of the three brands. The relevant supply area will be the same structure as its market area.

These eight hypothetical cases can be divided into four parts. First, *Cases I* and *II* will be examined in Section 7.3 as a single centre model. Second, *Cases III* and *IV* will be analysed in Section 7.4 as a multiple-centre model. Third, *Cases V* and *VI* will be explored in Section 7.5 as an overlapping model. Finally, *Cases VII* and *VIII* will be investigated in Section 7.6 as an exclusivity-area model.

7.3. A Single Centre Model

This section will examine a single-centre spatial pattern of a market area and supply area. First, a simple pattern where production of goods is operated in a region, and all outputs and inputs are distributed and collected within the region. Second, the analysis will be extended to multiple market-area patterns.

7.3.1. A Coincident Centre of the Market Area and Supply Area within a Region (Case I)

A simple case is one in which a production is operated in a region, and both relevant input and output are exchanged within the region. In other words, there are no imports and exports related to this product. In this case, a simple spatial input-output analysis can be applied as illustrated in Figure 7-2 (below).

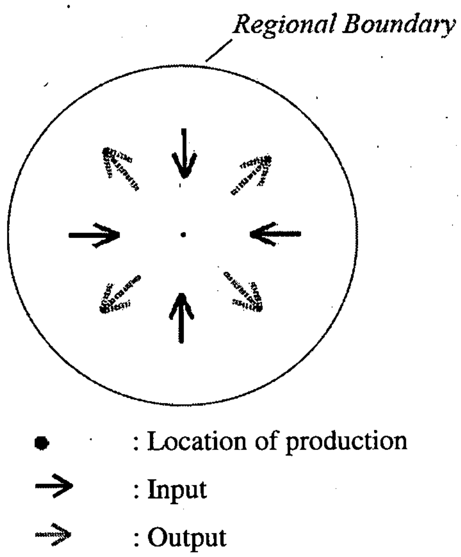


Figure 7-2. An identical centre of market area and supply area

There are three factors which determine the spatial structure of this production. First, the extent of the market area is determined by the combination of spatial competition of consumer demand and distribution transportation cost. Second, the extent of the supply area is specified by the competition of inputs and the assembly shipping cost. Finally, the producer is required to manage the operation within given budget constraints. However, this examination assumes that the firm has a sufficiently large budget to operate the processing. In addition to the above economic circumstances, technological advantages may be present during assembly transportation, distribution transportation and the production process. First, technological improvement on assembly transportation may reduce the producer's assembly cost. Second, the improved technology in distribution may reduce the *f.o.b.* price level. Finally, the improvement in production process may reduce the average processing cost of the producer. This case will examine a scenario in which a single firm produces an output at the centre of a region. In addition, both market areas and supply areas have limited maximum shapes and sizes within the region, as there are additional cost burdens at the regional boundary. As will be examined in *Case II,* a regional boundary has a certain remarkable role in restricting imports from, and exports to, other regions, due to the presence of economic and administrative boundaries.

With respect to the limited economic space in market areas, positive profit attracts additional new entrants to the market, and reduces the market shares of existing producers. If there is an incumbent producer who monopolises a regional market and obtains positive profits, a potential new entrant will join this market. In this case, the market area is either divided into two at the centre, or shares the common area within the region. Both market-area formations obtain an equivalent number of consumers. As a result, there is no difference between the two types of spatial pattern for both producers in terms of the revenue maximisation point of view. For supply areas, the reformation process caused by the entry of an additional producer may be basically the same as for market areas. However, if these two producers use similar types of input and delivery, the supply areas should not be divided into two parts at the centre but should share a common area across the plain. As access to specialised delivery generally requires high levels of terminal cost and transportation rate, more savings can be made in transportation costs by sharing special methods of transportation than would be the case under independent operation. This is applicable to the oligopoly model as well.

7.3.2. A Multi-Regional Coincident Centre (Case II)

This case shows that a production plant obtains input within a region, and distributes output to other surrounding regions in addition to its own region, as illustrated in Figure 7-3 (below). The figure shows a situation in which production takes place at the centre of a region and inputs are collected within the region. However, the output is distributed not only within but also beyond the regional boundary. This is generally the case in location theory, where the market area is larger than its supply area, and products are distributed extensively to various destinations. In this case, the market-area structure may have more than two types of configuration, as there is a multi-regional boundary between this region and the other eight surrounding regions. Beyond the boundary, the output price can be higher, and these extended market areas will be decreased if the price exceeds that of other peripheral competitors. If the supply area also expands beyond the region, input competition may be observed in the same way as market areas. The difference compared to the previous simple case (*Case I*) is that the previous case requires to have the same size and shape of areas while with *Case II,* it is not necessary to maintain this condition.

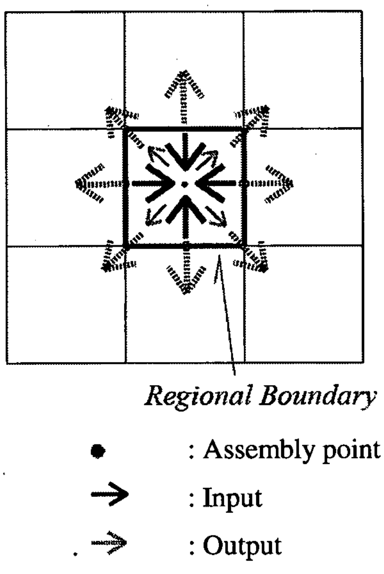


Figure 7-3. Multi-regional case

The $f.o.b.$ transportation cost has a two-stage cost structure with a threshold at the regional boundary B due to the existence of a terminal cost as shown in *Phase* (II) in Figure 7-4 (below). This affects the shape of the demand cone in *Phase* (III) through *Phases* (I) and (IV).

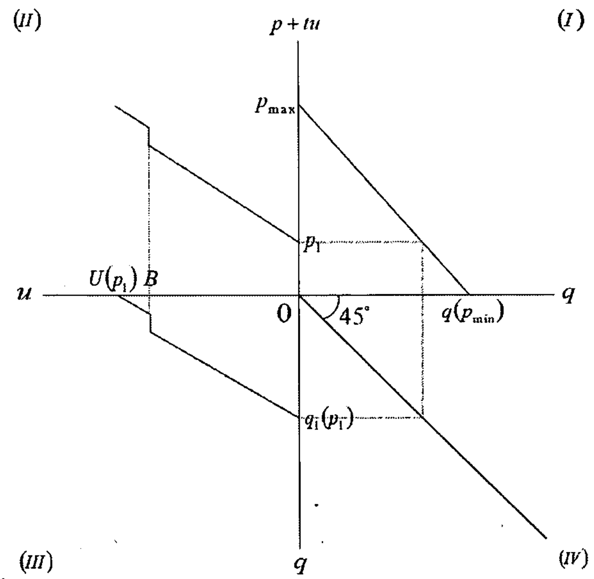


Figure 7-4. The demand cone and regional boundary

For conducting comparative-static analysis, the factor cost curve is required to be a continuous function. However, as shown in the above diagram, the curve has a discrete form. In order to avoid this problem, an alternative factor cost must be derived, which is a linear curve connected between the beginning and the end of this refracted factor cost curve. By way of this procedure, the alternative demand cone also becomes linear in form as illustrated by the broken line in *Phase* (III) in Figure 7-5 (below).

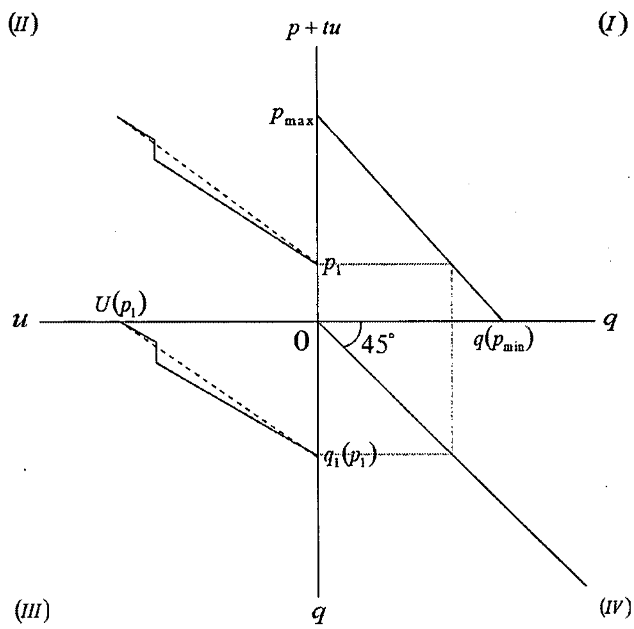


Figure 7-5. Average factor cost and continuous demand cone

As the additional broken lines in *Phases* (II) and (III) are shown by linear forms in the above diagram, it is possible to state that the presence of a regional boundary increases factor cost and reduces the volume of the demand cone in the original region.

In this circumstance, there is a trade-off interaction for the firm between economies of production scale and additional terminal cost beyond the regional boundary. This can be exemplified by the following situation using Figure 7-6 (below).

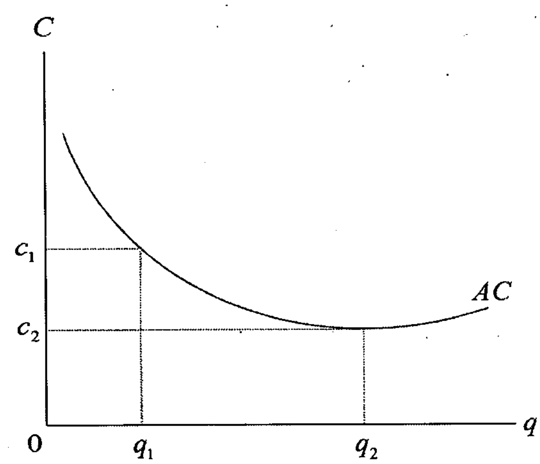


Figure 7-6. Economies of scale and regional boundary

Let us assume that a firm is producing q_1 at unit cost c_1 and distributes its outputs within a region. However, this firm achieves cost reduction up to c_2 if the output is increased to the level q_2. In this case, some outputs exceed the capacity of regional demand so that some are distributed to other regions with extra charges at the regional boundary. Figure 7-7 (below) illustrates the changing point of the distribution costs and terminal cost at the regional boundary u_B.

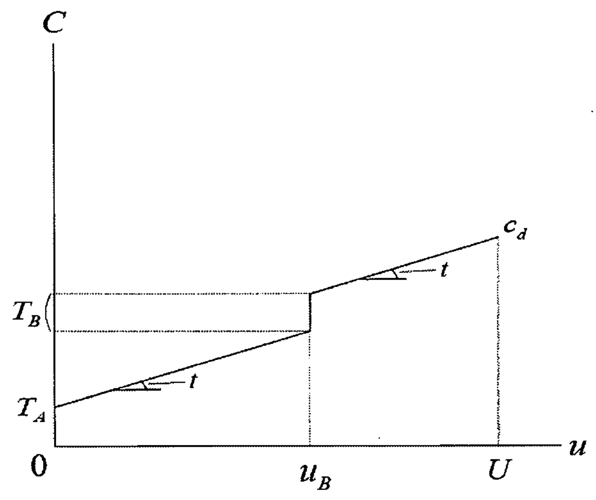


Figure 7-7. Distribution cost and regional boundary

The relevant distribution costs c_d in the above diagram can be expressed as

$$c_d = tu + T_A \qquad 0 \le u < u_B \qquad (7\text{-}1)$$

$$c_d = tu + (T_A + T_B) \qquad u_B \le u < U \qquad (7\text{-}2)$$

where t = distribution transportation rate, u = market-area radius, T_A = original terminal cost and T_B = additional terminal cost at the regional boundary.

As previously demonstrated, the combination of these cost curves is generated by taking the average between the origin and the maximum market-area radius U. Let us assume that the location of the regional boundary u_B is characterised by an index P $(0 \le P \le 1)$ which represents proportions between the distance from the origin to the regional boundary u_B, and the distance from the regional boundary u_B to maximum market-area radius. This shows that the regional boundary locates closer to the origin which is the centre of the region as P approaches zero. The alternative distribution cost c_d can be now provided as:

$$c_d = P(tu + T_A) + (1 - P)[tu + (T_A + T_B)]$$

$$= tu + T_A + (1 - P)T_B \qquad (7\text{-}3)$$

The excess cost c_e which is different from the original distribution cost will be:

$$c_e = [tu + T_A + (1-P)T_B] - [tu + T_A] = (1-P)T_B \qquad (7\text{-}4)$$

This is the extra cost burden of a producer to export products to other regions. If the effects of economies of production scale exceed this cost burden, the firm will maximise its economies of scale and increase the output up to the exporting level. These additional economies are expressed by $c_1q_1 - c_2q_2$ in Figure 7-6 (above). As a result, it can be concluded that the firm will choose export if:

$$c_1q_1 - c_2q_2 > (1-P)T_B \qquad (7\text{-}5)$$

This shows that the condition of economies of scale, proportion between the size of region and maximum market-area radius, and the level of terminal cost at the regional boundary, have important roles for the determination of the adjustment of output level.

7.4. A Multiple-Centre Model

This section will investigate various spatial patterns where a market area and supply area do not share a single centre. First, a simple two-stage processing of a product and its distribution will be examined. Second, a complex pattern, where various processing stages and dispersed market areas are observed, will be demonstrated.

7.4.1. An Assembly Plant and Dispersed Plants (Case III)

The following case shows that there are separations of the production process. Let us assume that there is an assembly plant which produces for eight surrounding finishing plants. These plants produce final outputs to the market in eight separated locations, and to a part of the area of the assembly plant as shown in Figure 7-8 (below).

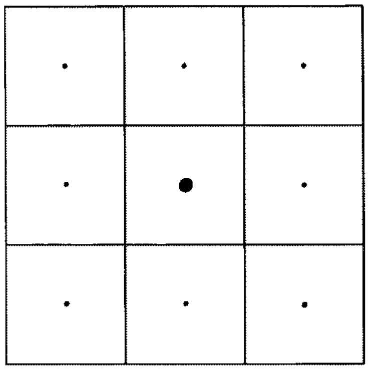


Figure 7-8. An assembly plant and eight finishing plants

This case will assume that an assembly plant processes a semi-finished product at the centre and distributes this product to eight peripheral finishing plants which complete the final-stage of processing. In terms of cost minimising behaviour, producers generally integrate their processing stages in a specific factory and a disintegrated situation can mainly be explained by the following three points: that there are diseconomies of scale at the assembly plant; that transportation costs for inputs are sufficiently lower than outputs; and that the assembly plant is located in the metropolitan area and there are urbanisation diseconomies in addition to the advantages of urbanisation economies. For instance, the high labour costs at the metropolitan area are referred to as one of the urbanisation diseconomies. In addition, production plants at the metropolitan area also suffer congestion, pollution and high property costs. These urbanisation diseconomies attract firms to locate outside of the metropolitan area, even though certain levels of urbanisation economies are unavailable at the non-metropolitan area. In this case, eight finishing plants can be divided into two groups according to the distance between each plant and assembly plant.

The diagonal plants, located at the northeast, northwest, southeast and southwest, solely distribute their outputs within their region. By contrast, the north, east, west and south plants distribute their outputs not only within their own areas, but also to the assembly plant area. Each plant shares 25% of the area in which the assembly plant is located. These divisions are according to the minimum transportation system of the spatial structure in the above diagram. In this way, this producer establishes eight finishing plants in addition to the assembly plant located in the centre of the area. However, it is more plausible to operate only the four finishing plants located at the north, south, east and west from the assembly plant rather than establishing eight plants. This is due to the minimisation of transportation costs with respect to distance and is as shown in Figure 7-9 (below).

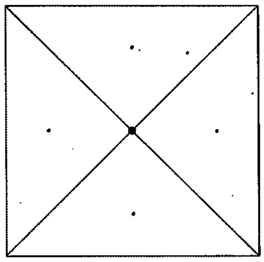


Figure 7-9. An assembly plant and four finishing plants

However, the distribution cost from each finishing plant will not be minimised in this case. Furthermore, this may cause an increase in the market price and, if the firm is competing with other firms, a corresponding reduction in revenue due to reducing consumer demand. From the standpoint of transportation cost savings, it is more reasonable to locate the four finishing plants diagonally as shown in Figure 7-10 (below). In this way, the number of finishing plants and their locations can be determined not only by the extent of the economies of scale, but also by the condition of transportation costs for inputs and outputs.

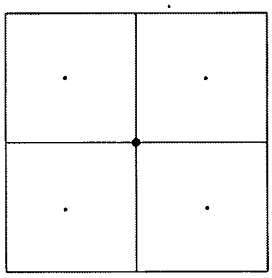


Figure 7-10. An assembly plant and alternative four finishing plants

At this stage, it is possible to compare the operations of the single and separated plants. For the assembly plant, establishing the separated plants reduces managerial cost but simultaneously increases transportation costs and unit cost of production. This alternative result is more detrimental for cost savings and thus outweighs the benefits of the original condition. However, there is a missing point that the results are brought about by a more efficient structure of spatial production function with respect to spatially constrained external economies. These interactions are illustrated in Figure 7-11 (below), where the expressions in subscript 2 relate to the separated plants' operation.

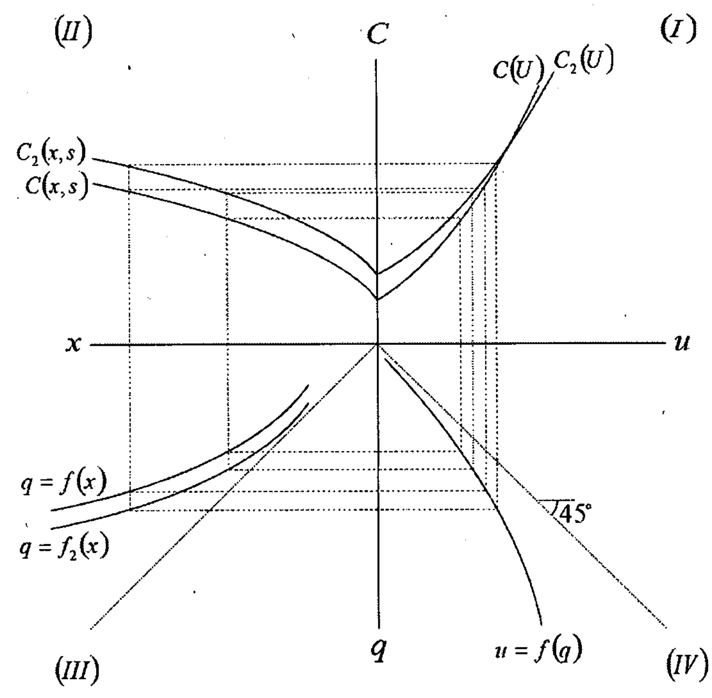
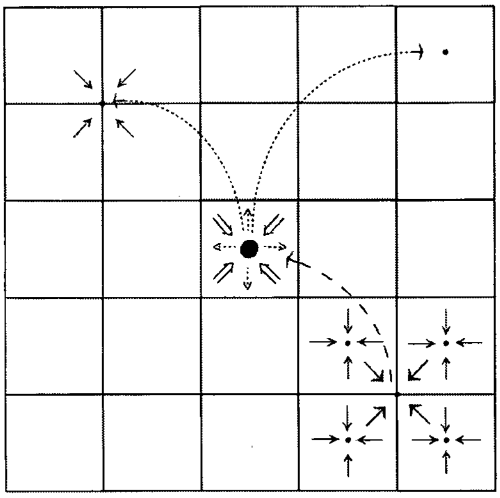


Figure 7-11. Factor cost curves and spatial production functions

In the above diagram, $q = f(x)$ in *Phase* (III) is the spatial production function which contains urbanisation diseconomies, and $q = f_2(x)$ is the spatial production function which avoids these diseconomies. The corresponding cost functions are illustrated in *Phase* (I). The original curve is represented as $C(U)$ and the alternative separated operation is shown as $C_2(U)$. Above the output level where the two curves intersect, the separated plants' operation should be adopted for the purpose of cost savings beyond the production scale where $C(U) > C_2(U)$. Otherwise, the operation should be integrated within the assembly plant to avoid insufficient economies of scale. In *Phase* (II), two factor cost curves are illustrated and the higher curve $C_2(x,s)$ represents the cost curve for the separated plants' operation. This curve has a higher factor cost than the original cost curve $C(x,s)$ due to dispersed plant operation. In addition, the slope is steeper due to the higher average transportation rate τ which includes the extra cost at the regional boundary. These two structures cannot be examined by comparative-static methods as the examination changes more than two variables at the same time. However, it is still possible to demonstrate more suitable location patterns between these two circumstances. As agglomeration economies increase by processing at the finishing plants, $q = f_2(x)$ in *Phase* (III) moves towards the 45° line which minimises economic losses. By contrast, if the production at the finishing plants obtains less agglomeration economies than the assembly plant processing, $q = f_2(x)$ moves away from the 45° line. As a result, this choice of separated plant operation with respect to cost minimisation requires larger a market-area radius.

7.4.2. A Complex Pattern (Case IV)

This case will exemplify a more common spatial economic structure. This examination will assume a non-homogeneous plain, dispersed uneven density of demand, and that there are several complex economic structures. Figure 7-12 (below) shows an example of a complex case of a multi-regional spatial structure.



● : An assembly plant
• : Finishing plants
⇐ : Input 1
<– – : Input 2
<··· : Output

Figure 7-12. A complex case

In this case, there is an assembly processing a product with two different types of inputs. *Input 1* is collected within a region; *Input 2* is assembled in a distant region with inputs provided by peripheral suppliers at the southeast region. Once processing is completed, the output goes to an area within the region, a distant market area and a distant assembly plant as an input for further processing.

From the above diagram, the following economic situations can be considered. *Input 1* is a labour input and is supplied within the region of the assembly plant. *Input 2* is assumed to have two stages before processing at the southeast region. The first stage is carried out by four different processing plants which supply inputs to the producer who makes *Input 2*. The production plant of this second stage is surrounded by these four first-stage plants and may have a sufficient level of agglomeration economies, referred to as the activity-complex type, to achieve lower production costs. In terms of shipping from this production location of *Input 2* to the assembly plant, although the shipping costs of *Input 2* are higher, the sum of lower mill price and higher shipping cost is still lower than obtaining a similar product from any other region.

For output, there are three different distributions. First, there is distribution within the region with the assembly plant at the centre of this market area. Second, the product is distributed to two additional distant locations. One is sold as an input for further processing into four types of product-differentiated goods at the northwest location in the above diagram. The other is sold to consumers located in a small distant village northeast of the metropolitan area. Although these villagers are charged to pay higher shipping costs, the unit shipping cost is much lower than their own management within their region due to insufficient volume of output to achieve economies of scale. In this way, a multi-regional spatial structure achieves the optimal operation for production.

In this type of linkage analysis, it is possible to apply the framework of the production possibilities set to show the technologically feasible combinations of inputs and outputs. For the assembly plant in this model, the production possibilities sets Y will be defined as

$$Y = \{y_1, y_2, y_3, -x_1, -x_2\} \qquad \textbf{(7-6)}$$

where y_i $(i = 1,2,3)$ represents output i and x_j $(j = 1,2)$ represents input j. Let us consider y_1 for output to the plant's own region, y_2 for the distant northeast region and y_3 for the producer in the northwest region. In addition, it is assumed that x_1 represents immobile labour in this region and x_2 represents semi-processed input from the southeast region. In order to analyse spatial equilibrium for this assembly plant, the demand conditions for outputs should be examined in addition to the plant's own cost structures. Transportation rate $t = t_1$ will be applied to output y_1 in a straightforward manner as is the case in established market-area analysis. On the other hand, it will be assumed that output y_2 is a single distant market point and that transportation cost can be expressed as a constant element $t = \bar{t}_2$. Likewise, output y_3 is a single firm who purchases this product as an input, and transportation cost is also defined as $t = \bar{t}_3$. As output y_3 is consumed as an input by a producer, this producer's production possibility sets Y_3 should also be defined as

$$Y_3 = \{z, -y_3, -l_z, -x_z\} \quad (7\text{-}7)$$

where z = the final output of this producer, y_3 = the input from the assembly plant, and l_z and x_z = labour and raw material inputs for processing these outputs. For reasons of simplicity, all inputs are assumed to have constant unit costs, except for labour source l_z which has decreasing returns to scale characteristics with δ $(0 < \delta < 1)$. As illustrated in Figure 7-13 (below), the total average cost TAC_z is expressed with average processing cost AC_z as:

$$TAC_z = c(x_z) + c(y_3) + t_3 + \delta l_z + AC_z \quad (7\text{-}8)$$

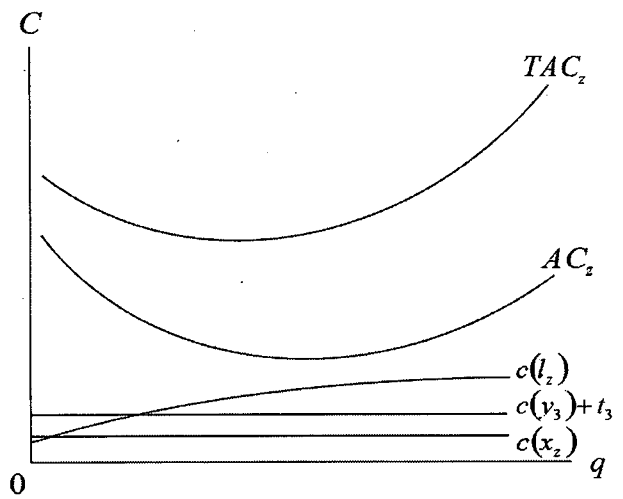


Figure 7-13. Total average cost for production z

Marginal cost MC will be the partial derivative of the above equation multiplied by quantity of output. If the market is in a state of spatial monopoly and the marginal revenue of the relevant demand curve MR is denoted as Equation (6-23) in Chapter 6, the optimal spatial equilibrium output level z will be determined where $MR = MC$.

Regarding the assembly plant, there are two more outputs y_1 and y_2 in addition to y_3. Output y_2 will be examined first. It is assumed that consumers of this output y_2 are located in a distant region and that they need to consume this product for particular reasons regardless of the shipping cost. The product price is set as the producer's marginal cost plus uniform transportation cost t_2. As shown in Figure 7-14 (below), the equilibrium output level becomes q_2^* where this price level p_2 equals the regional demand curve for this product as AR_2.

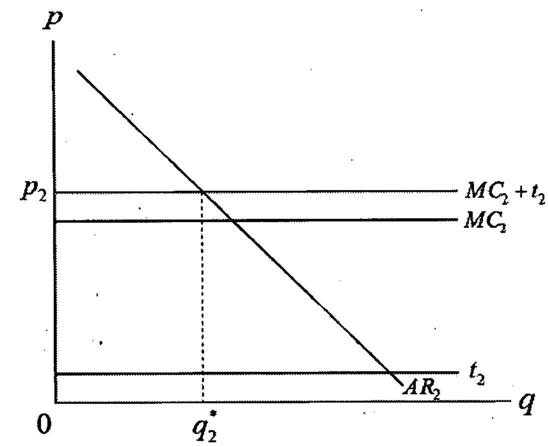


Figure 7-14. Market equilibrium of output y_2

For output y_1, established market-area analysis may be applied in a straightforward manner. It is assumed that there is spatial competition for this market area and that products are not exclusively supplied to this market. In this way, the assembly plant faces three different types of supply condition for outputs y_1, y_2 and y_3. The aggregate demand curve of these outputs is derived by horizontal summation of each demand curve AR_1, AR_2 and AR_3 as shown in Figure 7-15 (below).

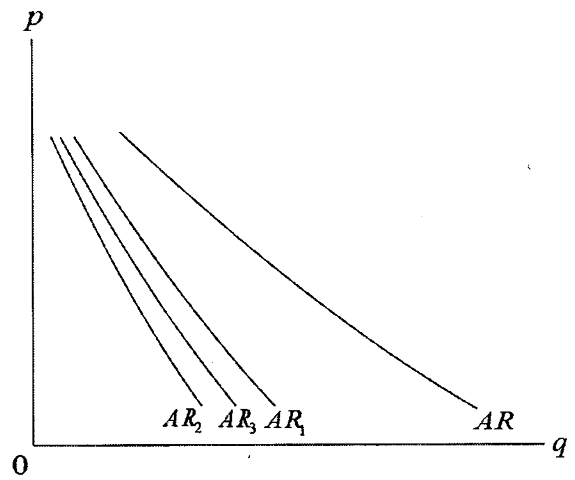


Figure 7-15. The derivation of aggregate demand curve AR

Regarding inputs, this firm has two different types: x_1 and x_2. As input x_1 is an immobile factor from the previously stated assumption, it can be exemplified by the commuting labour of this assembly plant. In this case, supply-area analysis will be able to reverse the structure of market-area analysis with respect to distance and spatial configurations. From the analysis of the supply area of this input, the cost curve of x_1, labour, will be directly determined by the density of the population, commuting transportation costs, the given production function and the condition of outputs. Another element of input, x_2, is a semi-assembled input processed by an upstream firm in the southeast area of Figure 7-12 (see earlier). The cost of input x_2 is generated based on the condition of processing technologies and its factor costs in the southeast area production. These unit input costs can be illustrated as in Figure 7-16 (below), where subscript t_{x1} and t_{x2} represent transportation cost elements for inputs x_1 and x_2. As this plant is located in the metropolitan area, the urbanisation type of agglomeration economy and diseconomy should be considered. This can be evaluated through the urbanisation-economies index curve AE_U. As introduced in Chapter 5, this index curve was studied in Evans (1972) with respect to scale and costs of floor space, labour, business service and capital. The vertical summation of all these curves will be the total average cost curve TAC as shown in the diagram below.

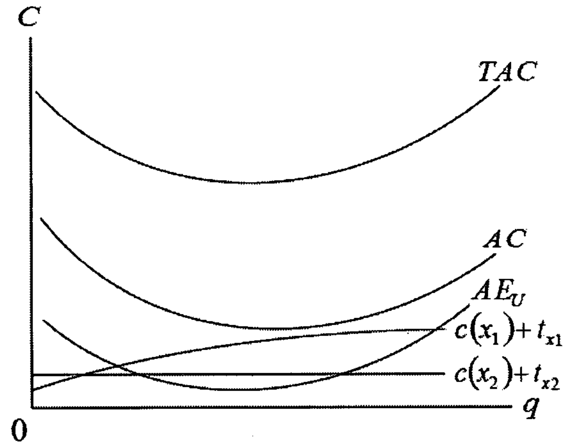


Figure 7-16. The assembly plant input costs and total average cost curve

In order to examine the spatial economic equilibrium of this assembly plant, the aggregate demand curve in Figure 7-15 (shown earlier) and that total average cost curve *TAC* in Figure 7-16 (above) should be mapped on the same diagram as shown in Figure 7-17 (below).

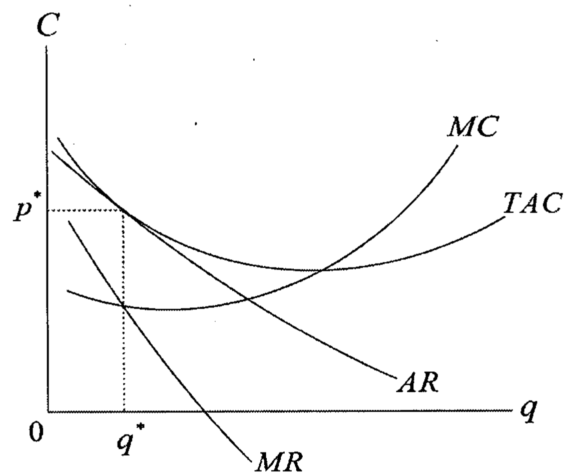


Figure 7-17. Spatial equilibrium mechanism of the assembly plant

Long-run spatial equilibrium is achieved at the production level where the spatial demand curve *AR* makes contact with the total average cost curve *TAC*. At this point, the marginal revenue curve *MR* intersects the marginal cost curve *MC*. If the situation is not optimal, the firm changes production scale or other relevant technological and spatial configurations until the production level is adjusted to the profit-maximising level. Figure

7-18 (below) shows an example of a more general case of the input-output framework.

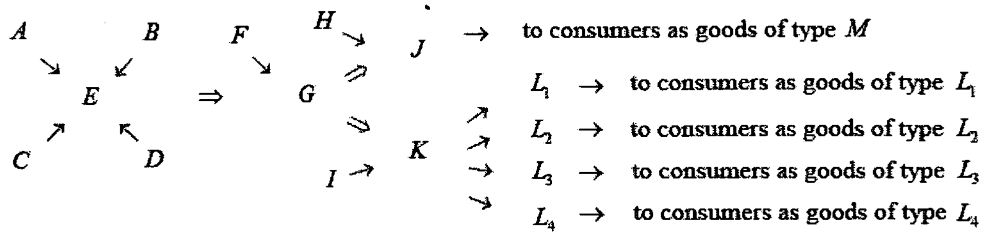


Figure 7-18. An input-output framework

Let us assume that an assembly plant G is processing an output. This production requires two different inputs. One is purchased from a supplier F and the other is supplied from the upstream section E. Section E uses four different types of input from suppliers A, B, C and D. From the standpoint of output, the assembly plant G distributes output to the downstream sections J and K. Section J processes this output with an additional input H and distributes to consumers as goods of type M. On the other hand, another downstream section K processes this output with an additional input I and produces four types of product-differentiated goods distributed to consumers as goods of types L_1, L_2, L_3 and L_4. This example can be seen in car-assembly, pottery works and other assembly-related manufacturers.

7.5. An Overlapping-Area Model

This section will examine an overlapping spatial pattern of market areas. First, perfectly-overlapping duopoly market areas and the relevant structure of supply areas will be analysed. Second, partly overlapping duopoly market area and the relevant supply areas will be demonstrated under the condition of product differentiation.

7.5.1. Perfectly Overlapping Duopoly Market Areas (Case V)

This case shows that there is no exclusivity in the market areas between the two product-differentiated brands α and β. As shown in Figure 7-19 (below), they share the same market area. This case will assume that two brands are distributed by different companies and that they are competing with another brand with respect to output level. It is also assumed that the two brands are similar products and that the two producers have the same technologies and other economic conditions. This spatial pattern shows that two independent firms are sharing common market areas.

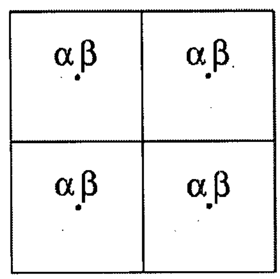


Figure 7-19. Perfectly overlapping duopoly market areas

It seems possible to consider the application of the Hotelling model in this analysis. Hotelling (1929) investigates the determination of price between two competing firms who distribute products at different locations from each other. The equilibrium states that each of the two firms will locate as close as possible to the other firm. However, the Hotelling model cannot be applied to this analysis for the following two reasons. One is that the Hotelling model assumes that the two goods are homogeneous products and therefore not product differentiated; the other is that there is competition over price but not other location factors. As a result, the model of price-level adjustment without product differentiation cannot be compatible with location analysis, as the latter assumes that product differentiation always exists, unless the market price is at a sufficiently high level.

Instead of these approaches, this analysis can be examined in terms of rivalry and cooperative choices for supply areas between two firms. In this analysis, there will be severe competition between the supply areas of two firms as their economic space is very limited and close to each other. These firms choose either rivalry or cooperation. In this case, cooperative behaviour is preferred when their transportation rates are at a sufficiently high level. As examined in Chapter 4, two firms can share transportation methods. This enables both firms to achieve certain cost savings, by applying economies of scale and sharing fixed costs, particularly in the case where specialised shipping is required for both brands of product. In addition, under the assumption of product differentiation, the two firms may have joint production in the upstream production stages as each firm produces similar goods. However, cooperative behaviour may not be observed if there is a severe spatial competition over occupying consumer demand in market areas. As a result, the duopoly model of this spatial pattern also relies on the condition of demand for both brands. This is more plausible in this spatial pattern, as the two brands share the same market areas implying that there is high demand for these brands. In this way, the condition of supply areas depends not only on the transportation network system but also on the structure of market areas.

There is one more thing which should be examined in this spatial pattern with respect to high transportation cost for output. Under the condition of the *f.o.b.* pricing system, certain levels of increase in transportation rate, reduce the volume of demand cone, as examined in Chapter 5. In Figure 7-20 (below), the original demand cone for the market area is illustrated as DC_1.

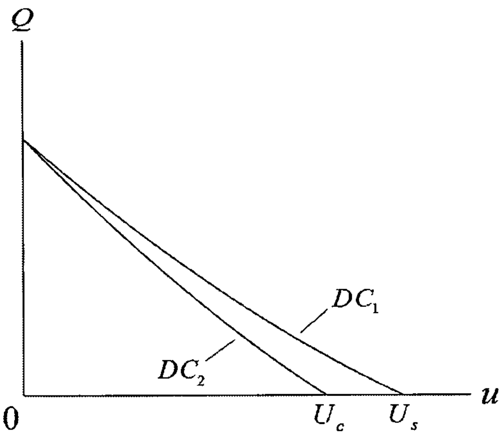


Figure 7-20. Demand cone and transportation rate changes

In the above diagram, a highly increased distribution rate of transportation t shifts the demand cone to DC_2. For reasons of simplicity, let us assume that the end of the market-area radius for DC_2 is U_C which equals the half distance between the centre of the two brands and the end of the market area. Similarly, the end of the market-area radius for DC_1 is U_S which equals half the diagonal distance of the market area of the two brands. As illustrated in Figure 7-21 (below), the original demand cone DC_1 forms square market areas as previously defined in Figure 7-19 (above). The alternative demand cone DC_2 forms circular market areas with radius U_c. Due to the increase in transportation rate t, consumers located in the outer circle will be excluded from the market of these products α and β. In order to avoid the consumer exclusion, the local authority pays subsidies to fill the space, or another Brand γ will enter to the market. If Brand γ appears in the market, the spatial structure becomes an overlapping oligopoly situation. Although the space is completely filled by the three firms in this circumstance, there are still consumer exclusions. Some residents have choices of all brands but others have limited choices of either brands α and β, β and γ, or α and γ. This type of consumer exclusion of oligopoly case will be further explored in later *Case VIII*.

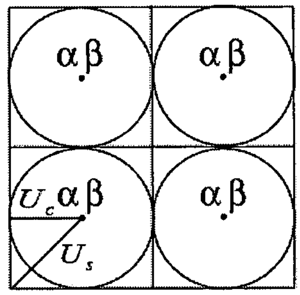


Figure 7-21. Consumer exclusions in perfectly overlapped duopoly market areas

7.5.2. Overlapping Duopoly Market Areas and Product Differentiation (Case VI)

This case shows that there are two brands α and β in the market and that their market areas overlap but do not share the centre of these areas. In this case, consumers choose either brand according to consumer preference between the two brands. As a result, this analysis requires to draw on consumer theory. The examination procedure will be as follows. If a consumer prefers brand β to α, his utility maximisation behaviour can be stated with the expression $\alpha \prec \beta$. Using this statement, a representative consumer's utility maximisation problem can generally be denoted as the following statement.

$$\begin{array}{ll} \text{maximise} & U = U(\alpha, \beta) \\ \text{such that} & M = P(p_\alpha \cdot q_\alpha) + (1-P)(p_\beta \cdot q_\beta) \\ \text{where} & P = 0 \ \text{ if } \alpha \prec \beta \\ & P = 1 \ \text{ if } \alpha \succ \beta \end{array}$$

where U = consumer's utility function, M = consumer's budget constraint and P = parameter. As denoted in the above statement, this consumer cannot maximise his utility by the combination of two brands: he can do so only by choosing either brand α or β if the product is too expansive to purchase two brands, i.e., in the case of a car or a fridge. Figure 7-22 (below) illustrates *Case VI* where Brand α and Brand β have different centres but share the same market areas.

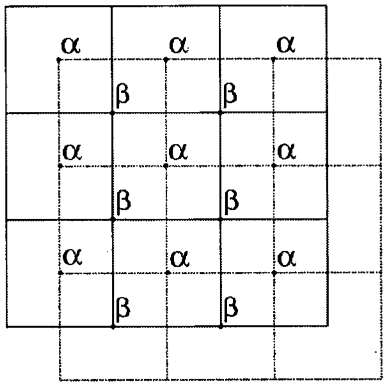


Figure 7-22. Overlapping market areas

In the above diagram, either brand α or β is chosen by consumers according to the balance between consumer preference and transportation costs to the distribution point. This is a trade-off interaction between preferences of goods and the additional transportation cost burden. Figure 7-23 (below) illustrates the $f.o.b.$ price and consumer budget constraint M. In this case, a consumer A located at α will choose Brand β over the nearer Brand α if he prefers Brand β and his payoff π for β as π_β satisfies the following condition.

$$\pi_\beta > M - p \qquad \textbf{(7-9)}$$

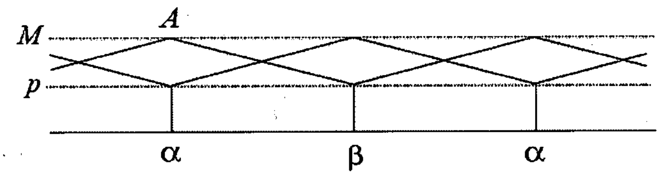


Figure 7-23. Output price and budget constraint

The above case shows that he is located at the distribution of Brand α. This is an extreme case and consumers may be located at any points between the distributions of the two brands. In order to consider a more general condition, let us assume that there is a consumer who prefers Brand β to Brand α but is located closer to the seller of Brand α. In this case, if his preference for Brand β is weaker than the additional transportation cost burden, he will give up obtaining Brand β and compromise to purchase Brand α from the nearer seller. However, if his preference for Brand β is stronger than the additional transportation burden, he will put up with travelling a long distance and paying a higher price to purchase Brand β from a distant seller. From the time-leisure standpoint of view, it can be stated that the temperate-humidity index will increase as the travel distance increases. Let us suppose that Figure 7-24 (below) illustrates the physical constant travel cost curve and two disutility curves $-U_A(\alpha)$ and $-U_A(\beta)$ of this consumer A for obtaining brands α and β.

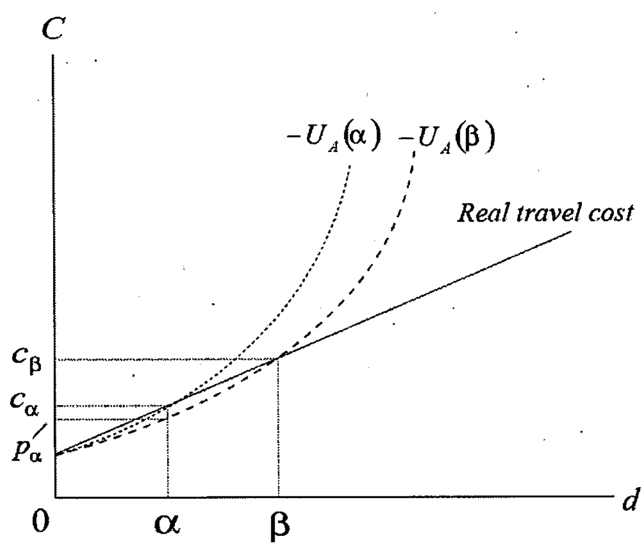


Figure 7-24. Travel cost and consumer preferences

The above diagram shows that consumer A will choose Brand β even though the nearer seller is Brand α. At the location of Brand α, this consumer can purchase Brand α. However, his disutility curve for Brand β is lower than for Brand α at this point. If $-U_A(\beta)$ exceeds the line of actual travel cost at location β, he will not travel to obtain β and instead purchase Brand α as a compromise. This index of compromise can be measured as $c_\alpha - p_\alpha$ in the above diagram. In this case, he will be able to obtain Brand β at the cost of c_β without compromising the value of

$c_\alpha - p_\alpha$. This disutility curve can also be examined with respect to substitution and income effects of the properties of complementary goods.

The formal representation of this spatial consumer utility-maximisation problem can be stated with the travel costs t_α and t_β to the distribution point of the brands α and β:

maximise	$U = U(\alpha, \beta)$
such that	$M = P(p_\alpha + t_\alpha)q_\alpha + (1-P)(p_\beta + t_\beta)q_\beta$
where	$P = 0$ if $\alpha \prec \beta$
	$P = 1$ if $\alpha \succ \beta$

As previously examined, the above case also shows that the consumer cannot choose both brands α and β but can choose either α or β. In addition, consumers will access another market area if their preferred brand is not available within the market area. This can be illustrated by the ideal range. If a consumer locates at the centre of Brand α, he will be able to obtain Brand α without any shipping cost. However, if his preference is denoted as $P = 0$ in the above condition, he will travel to the distribution point of Brand β. In this case, his ideal range can be illustrated as the subscribe circle of the Brand α market area as shown in Figure 7-25 (below).

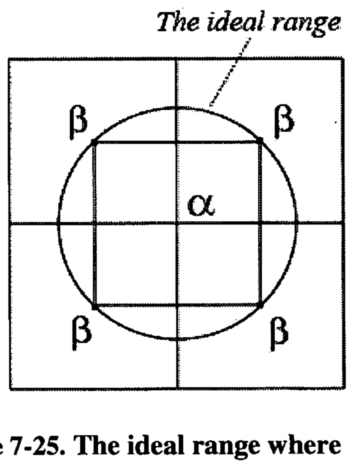


Figure 7-25. The ideal range where $P = 0$

The size of the ideal range depends on the relative levels of price p_i and distribution transportation rate t_i $(i = \alpha, \beta)$ between the two brands.

Regarding the producers, there are four centres of Brand β at the market-area boundary of Brand α. Similarly, there are four centres of Brand α at the boundary of Brand β. The relevant supply areas can be illustrated in Figure 7-26 (below).

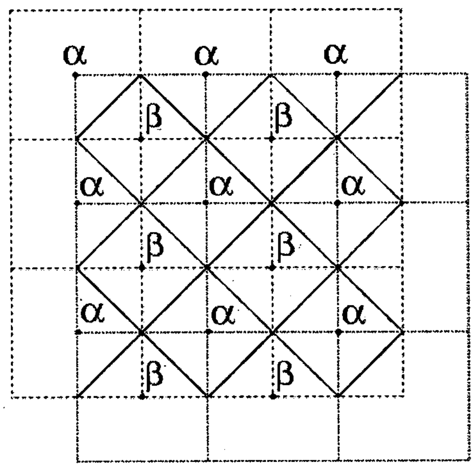


Figure 7-26. Supply areas of overlapping market areas

In this case, the supply areas will not necessarily be shared between two firms if the condition of limited supply does not exist.

7.6. An Exclusivity-Area Model

This section will analyse an exclusivity pattern of market areas. First, a spatial duopoly model under the condition of exclusivity market areas and the relevant supply area will be examined. Second, a spatial oligopoly case will be investigated. This model will also explore the alternative supply-area formation of the joint location of three independent firms with the notion of agglomeration economies.

7.6.1. Duopoly and Exclusivity of Market Areas (Case VII)

This spatial pattern shows that there are two types of similar brands α and β on the economic plain and that each brand is exclusively distributed to each market area. Thus, there is no overlapping area, as illustrated in Figure 7-27 (below).

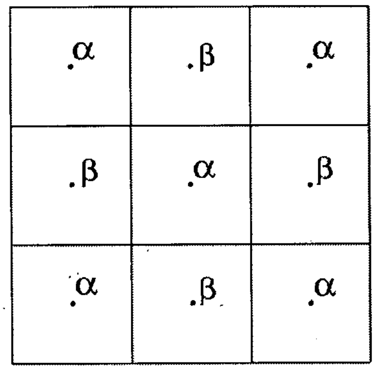


Figure 7-27. Exclusive duopoly spatial structure of market areas

This particular duopoly model can be observed in the following economic circumstances. As shown in Figure 7-28 (below), there is an extremely high level of distribution transportation rate. In addition, this level is too high to distribute goods over more than half of the market area. In this case, the two brands must have a spatially dispersed and exclusive market area structure. The second case is where there is an extremely high level of assembly transportation rate for processing each brand, and the size of each market is small.

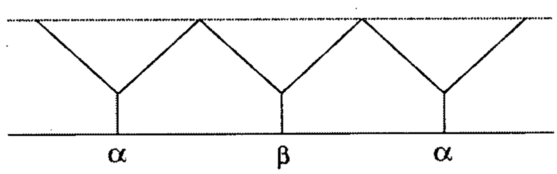


Figure 7-28. Exclusive duopoly market areas

This approach can also follow the economic law of market areas. When the price levels and transportation rates of both products are equivalent, $p_\alpha = p_\beta$ and $t_\alpha = t_\beta$, the boundaries will be shown as in the above diagram. Alternatively, it is not necessary to satisfy $p_\alpha = p_\beta$ and $t_\alpha = t_\beta$ if

$p_\alpha + t_\alpha = p_\beta + t_\beta$. It should be noted that this spatial allocation also corresponds to the supply areas if other variables are indifferent to market areas. However, if $p_\alpha \neq p_\beta$ and $t_\alpha \neq t_\beta$, the supply-area size can differ from the size of market area, even though the above alternative necessary condition $p_\alpha + t_\alpha = p_\beta + t_\beta$ is satisfied, since supply-area size relies also on factor price and assembly transportation cost.

This example can also be found in the following three cases. First, when the size of market areas is extremely large, the feasible distance of delivery is limited by this size, and the relevant competitors cannot overlap their market areas. Second, when the optimal production scale is very small, individual firms cannot satisfy the entire demand of the market areas and their feasible size of market area is limited below the overlapping level. Finally, when Cournot's (1838) duopoly equilibrium is applied, in which unprofitable price adjustment is replaced by quantity adjustment, excess demand will appear and a single firm cannot occupy the entire market. In these cases, the economic plain can be shared between two firms without overlap. In this case, consumers have to accept an elastic supply curve condition beyond certain levels of market price as shown in Figure 7-29 (below).

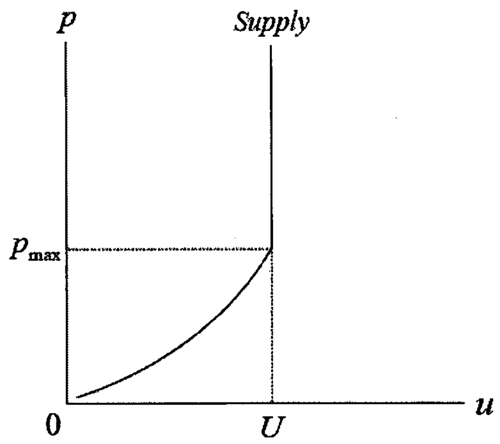


Figure 7-29. An inelastic supply curve

This *Case VII* can demonstrate a spatial equilibrium under the conditions of duopoly and exclusive structure of market areas. The equilibrium model is illustrated in Figure 7-30 (below).

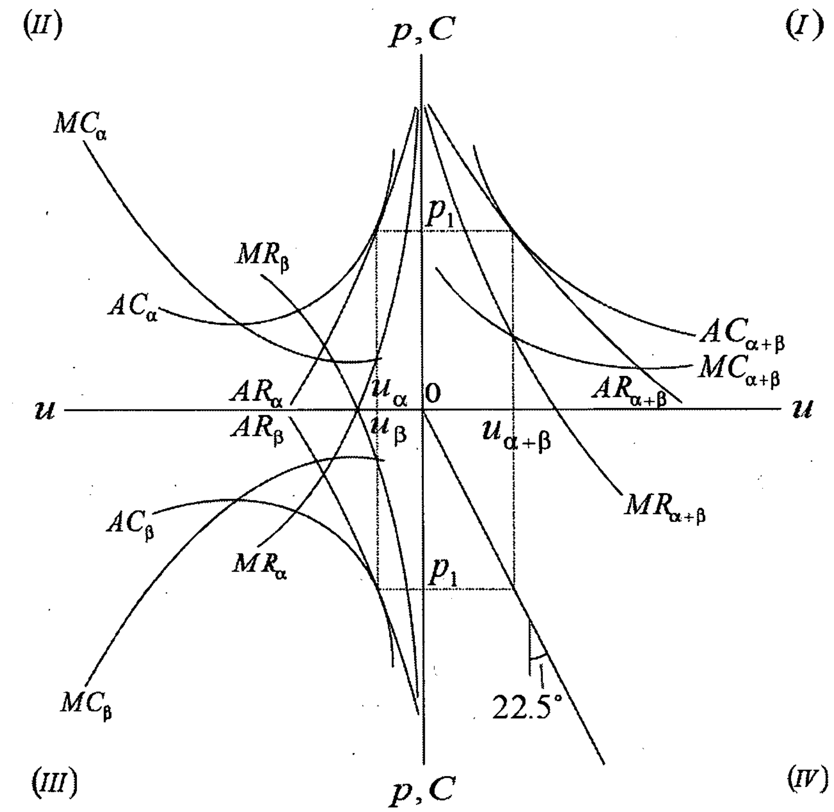


Figure 7-30. Spatial duopoly and exclusive structure of market areas

In the above diagram, *Phase* (I) shows the spatial equilibrium of two market areas, *Phase* (II) depicts the spatial equilibrium in the market area provided exclusively with Brand α, *Phase* (III) shows spatial equilibrium in a market area provided exclusively with Brand β, and *Phase* (IV) represents the 22.5° reflection line, which is the half of the 45° reflection line, which connects *Phases* (II) and (III) to *Phase* (I). If the market areas of the two brands do not have a symmetric price condition, this 22.5° reflection line will become more or less steep in order to adjust the aggregate level in *Phase* (I). Thus, the slope of this line represents the price ratio of market areas between the two brands α and β on the plain. However, the condition of equal market-area size level $u_\alpha = u_\beta$ must be satisfied as shown in the above diagram. This situation; market-area size level but different output level between two brands, can be observed where the transportation rate for either distribution t or assembly τ of one brand is higher than the other. This is one of the ways that the products are differentiated in terms of location analysis. In this case, consumers located at the site of the higher-price brand face consumer exclusion as they cannot choose the less expensive brand due to accessibility in terms of budget constraint. Likewise, consumers located at the site of the other brand in the market area also experience consumer exclusion, as they cannot choose the higher-price brand even if they are willing to obtain this product.

There is one more instance of consumer exclusion in this case of spatial pattern. As shown in Figure 7-31 (below), consumers located at the outer circles of each market area cannot obtain any products if the transportation rate and price are at a sufficiently high level.

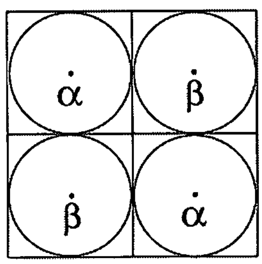


Figure 7-31. Market areas and consumer exclusions in duopoly case

In order to avoid this problem, the local authority may provide subsidies and entire areas will have a space-filling economic pattern. Otherwise, another Brand γ may enter the market to fill the entire space and form an oligopoly monopoly. In this new-entrant case, consumers in each market area of brands α and β will be reduced certain volume of output. If these potential losses exceed cost minimising circular strategy, two existing brands α and β will occupy a square space-filling spatial structure without relying on public subsidies. These losses can be explained by the cost of changing to a smaller scale of production facilities and by the decreased amount of revenue from reduced sales of outputs. This is shown in Figure 7-32 (below) by changes in cost and revenue curves. In other words, the optimal output level is reduced from q_1 to q_2. In addition, the corresponding cost and price levels c_1 and p_1 are increased to c_2 and p_2, respectively.

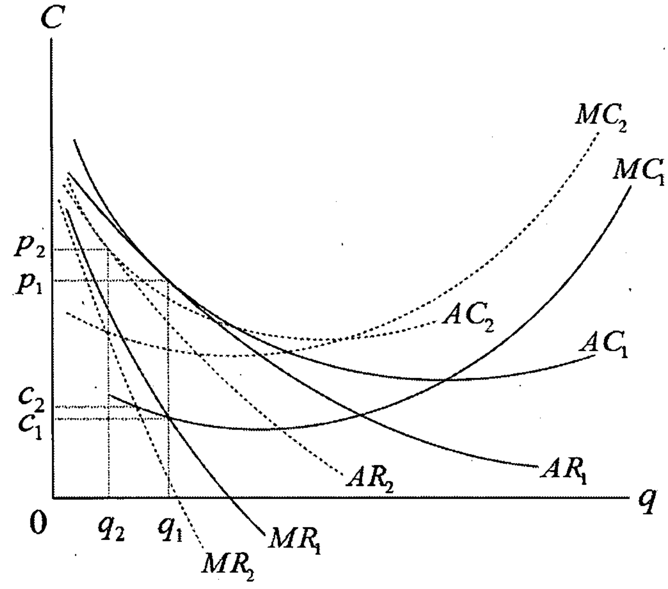


Figure 7-32. Production scale changes and alternative spatial structure

7.6.2. Oligopoly and Exclusivity of Market Areas (Case VIII)

This case will introduce three completely product differentiated brands, α, β and γ, distributed by three independent companies. The previous cases examine duopoly models where two different brands α and β fill the economic space in either an overlapping or exclusive form, and potential new entrants to the market can also be observed. In the cases that follow, the market areas will be of a regular hexagonal oligopoly form once a new entrant joins the market and all firms achieve space-filling equilibria. The situation is either mutually exclusive market areas or perfectly overlapping. The former pattern is shown in Figure 7-33 (below). These types of spatial pattern are examined as the $n-$competitor case of the duopoly model.

There is one more case of a space-filling structure which is an intermediate case between the above two types. This case occurs when the following three conditions apply. First, the entire market area is too large for every brand to be sold. Second, there are insufficient numbers of consumers in each market area for every brand. Finally, the assembly plant of each brand must be dispersed across the economic plain. This final condition is due to the fact that the relevant volume of deposits of inputs is limited per square-kilometre and the assembly transportation rate is at a high level.

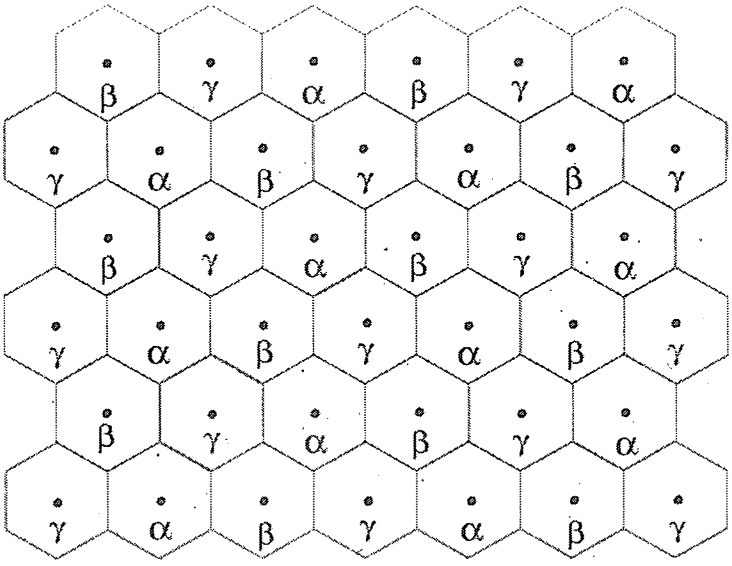


Figure 7-33. Market-area and supply-area territories between three different firms

The above diagram forms a symmetric hexagonal market-area and supply-area structure. However, they are completely different from the other existing hexagonal spatial analysis. This particular case in the diagram is observed only if a further three conditions are assumed. First, there are three independent companies and each company has the same conditions for operating their economic activity. Second, no market areas overlap in order for the exclusivity condition to be strictly kept in the assumption. Finally, the three different products are complementary in order for there to be no incentive for displaying preference for one of the brands.

There can be consumer exclusion in some areas, for example outside the circles shown in Figure 7-34 (below). In this oligopoly case, these areas are smaller than those of the duopoly case since the oligopoly case forms regular hexagons while the duopoly forms squares. It can be interpreted from this case that if the local authority considers subsidising the industry to support consumer demand, the oligopoly case will have a lower cost structure than that of the duopoly square case.

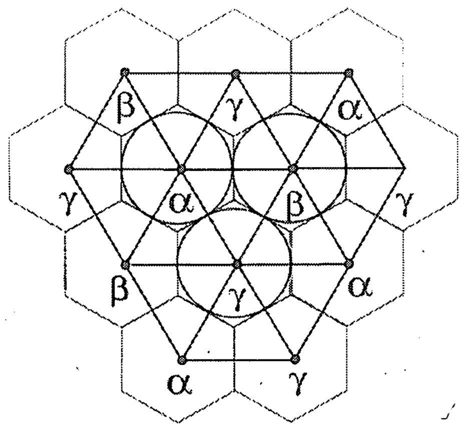


Figure 7-34. Market areas and consumer exclusions in oligopoly case

The formation of duopoly or oligopoly spatial structures with respect to consumer exclusions and price adjustment can be summarised by the following four types of attributes. The duopoly situation is maintained when the existing two firms reduce their price levels down to consumer's maximum reservation price level in order to avoid a new entrant to the market. Another case is when these existing firms receive subsidies from local authorities for the equivalent amount of price reductions. These effects are shown as the changes to the dashed price line in Figure 7-35 (below).

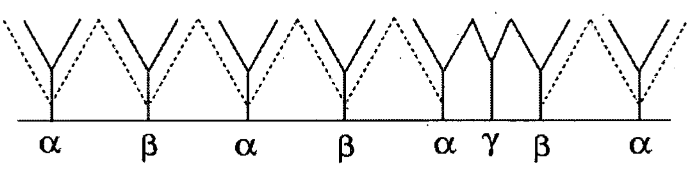


Figure 7-35. Price reduction and entrant barrier

By contrast, the duopoly situation is not maintained and the market becomes an oligopoly when a new entrant γ locates between the two brands α and β, or a new entrant γ sets a *c.i.f.* price setting which is equivalent to the level of the maximum consumer reservation price. The former case is illustrated in the above diagram as γ between two existing brands α and β. The latter case is achieved if the saving cost for the establishment of a new distribution point between two brands α and β exceeds this $c.i.f.$ pricing level for Brand γ. This price level is illustrated in Figure 7-36 (below).

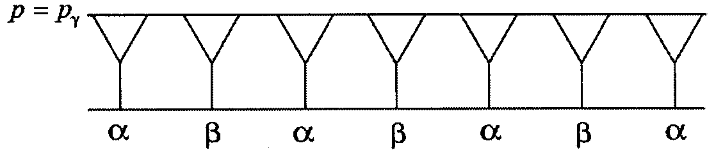


Figure 7-36. New entrant with the uniform $c.i.f.$ price p_γ

The last case enables all consumers to have two choices from two brands and contribute to prevent consumers from consumer exclusions. In terms of consumer exclusions, the overlapping market-area pattern between three brands is preferred to the exclusive market-area circumstance. However, in the case of partly overlapping market areas with three brands, there may still be consumer exclusion of one or two brands. As shown in Figure 7-37 (below), in part of market area α, all three brands are available to some consumers. However, either β or γ are not available to other consumers. In addition, neither brand β nor γ is available in some areas. These exclusions are caused by the combination of the partly overlapped spatial structure of the market areas and the high rate of the $f.o.b.$ transportation rate of outputs.

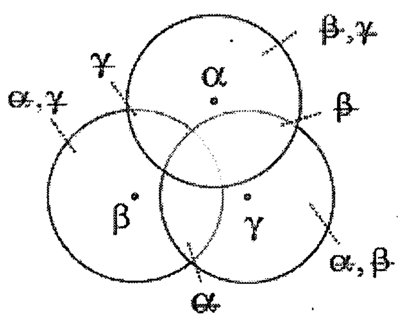


Figure 7-37. Consumer exclusion in overlapping oligopoly case

The above argument can be more precisely examined in Figure 7-38 (below).

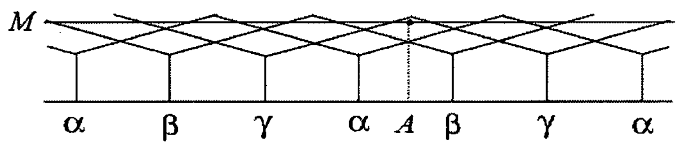


Figure 7-38. Consumer exclusion for Brand γ in overlapping oligopoly case

In the above diagram, the line M represents the budget constraint of a representative consumer. For instance, a consumer A who locates between the distribution points of brands α and β is able to choose from these two brands. However, he cannot purchase Brand γ at this location as the $f.o.b.$ price of Brand γ at location A exceeds his budget constraint level M. As previously examined, this problem may be solved by a subsidiary payment from the local authority should they wish to guarantee its availability.

Figure 7-33 (shown earlier) illustrates a spatial pattern where three brands α, β and γ exclusively occupy every market area. In this case, all processing may be engaged independently. If the centre of each market area is a metropolitan area, the situation could be changed as examined in Chapter 5. If the product does not require to have location proximity to the metropolitan area, firms tend to avoid locating at the centre due to the presence of urbanisation diseconomies. In this case, firms will locate closer to the spatial boundary and other producers. If all three producers come closer to each other due to mutual attempts to avoid production at each metropolitan area, and they are producing product-differentiated but similar types of goods, the three firms can situate at a common location and have certain types of agglomeration economies. Under the condition of a uniform spatial pattern, the optimal firm location can be illustrated in Figure 7-39 (below).

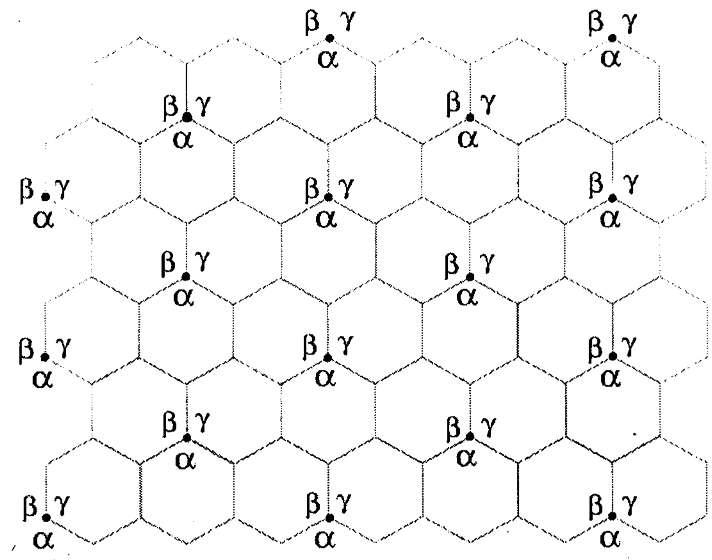


Figure 7-39. The optimal firm location in terms of the integrated framework

The relevant supply-area configuration can be shown as Figure 7-40 (below), if other conditions are kept constant.

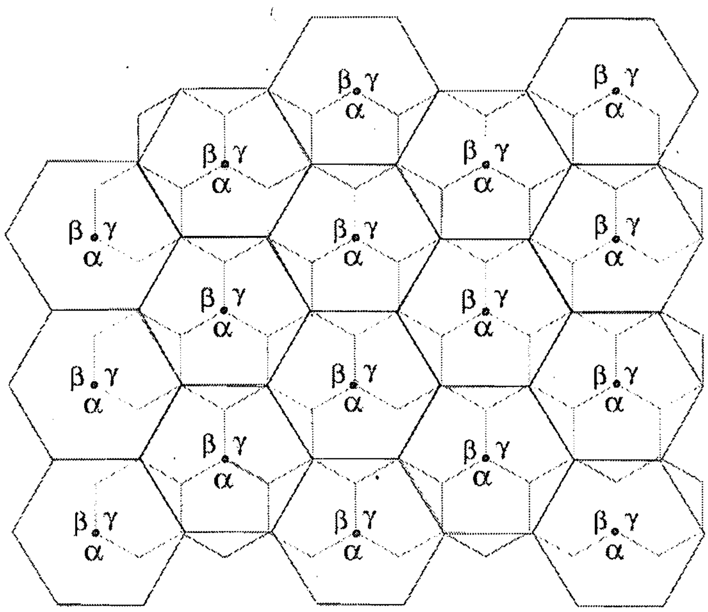


Figure 7-40. The alternative market-area and supply-area configurations

The alternative market-area configuration also becomes the same shape. In this way, including the concept of agglomeration economies may change the structure of market-area and supply-area configuration. Not only does this bring cost savings for producers; it also solves the problem of consumer exclusion for particular products. Thus, this is one of the Pareto improvement solutions which are brought about by the consideration of agglomeration economies. However, it should be noted that there may still be possibilities to have consumer exclusions when the distribution transportation rate increases and market areas become circular configurations. In addition, if the transportation network has an important role for this activity, the production should be operated on the triangular transportation network as demonstrated in Chapter 4. In this way, agglomeration economies and transportation costs cannot be excluded from the analysis of firm location with respect to the integrated framework of market-area analysis and supply-area analysis.

7.7. Conclusion

This chapter initially outlines a theoretical framework involving hypothetical examples, and then introduces a typology of eight cases of spatial structure. These cases are separately examined in the integrated framework of the analysis of market areas and supply areas. The former parts clarify the interaction between both types of area, demonstrated by simplified examples. The analysis attempts to explore more complex cases and therefore the consideration of various location factors must be included in the model framework. The latter parts further extend the analysis to the several irregular spatial formations of market areas and the corresponding structures of supply areas. From the former approaches, it becomes clear that plant location analysis should be conducted in terms of market areas, supply areas, and spatially unconstrained and constrained internal and external economies. From the latter approaches, it becomes obvious that plant location is not required to be investigated, as the individual firm is considered to be operating under optimal-production conditions. However, if a specific exceptional spatial structure is observed, particular locational patterns and production conditions will require to be investigated, taking into account the spatially unconstrained and constrained internal and external economic factors. In addition, it can be stated that economic policies for solving the spatial consumer exclusion problem can be formed by giving full consideration to the effect of market-area and supply-area configurations on spatially unconstrained and constrained internal and external economic factors concerning the location of production firms.

Chapter 8. Concluding Comments (Summary of Findings, Comparative-Static Results and Avenues for Further Research)

The concern of this research has been to clarify the structural relationship between market areas and supply areas, in terms of an input-output framework, as well as additional factors. As the established structure of both types of area framework needs to be introduced, market-area analysis and supply-area analysis are individually examined, following an introductory chapter. The core elements of market areas are analysed in Chapter 2. Following the study of the existing market-area analysis, models of spatial competition under free entry and monopoly are examined, and it is suggested that the existing framework is limited by certain simplifying assumptions made for the purpose of theoretical investigation. Supply-area analysis is examined in Chapter 3, where reference is made to the existing literature and the related assumptions therein, regarding limitation of supply, assembly processing cost and transportation rate. The limitations of supply-area analysis are then examined in the light of these assumptions in order to indicate further possibilities of extension. These limitations mainly regard the measurement of technological elements for production and other relevant forces of economies and diseconomies.

In order to avoid confusion between the two types of area, their similarities and dissimilarities are analysed in Chapter 4, where the findings reveal similarities in spatial terms and dissimilarities in economic terms. Furthermore, attention is drawn to the presence of hitherto neglected factors, namely spatially unconstrained and constrained internal and external economies, in order to investigate both types of area analysis in the same framework. The relationship between these additional economic factors and firm location is further investigated in Chapter 5. This examination clarifies the structural interactions between market areas, supply areas, firm location and additional economic factors by means of the combination of the input-output framework and the duality theory between production and cost functions. In Chapter 6, the means of applying the original duality theory in conventional economic analysis to location analysis is demonstrated as an extensive form of the input-output framework. This then enables an integrated-framework analysis of market areas and supply areas to be applied, in order to determine the optimal radius of market areas and supply areas, quantity of output, amount of inputs and availability of internal economies and external economies.

The spatial equilibrium approach under conditions of spatial monopoly is applied as follows. The factor cost is generated from the supply-area framework, and is directly related to the inputs and the additional factors. This factor cost defines a cost function which represents the total cost of output. In order to derive the cost function, the spatial production function requires to be substituted into factor cost. The spatial production function is the combination of the conventional production function and additional factors. The derived cost function is mapped on the relevant spatial demand conditions and the optimal market-area radius and quantity of output are obtained. Applying duality theory, the optimal amounts of inputs and supply-area radius are also derived from the optimal market-area conditions.

This integrated framework has not yet been attempted in existing location analysis, and may contribute to clarifying the interaction of spatially unconstrained and constrained internal and external economies between market areas, supply areas and firm location. In order to explain further its application to various spatial economic situations, the analysis considers eight representative hypothetical cases in Chapter 7. The hypothetical cases begin with the most simplified homogeneous pattern, *Case I*, where market area and supply area have the same centre and there are no exports and imports. These conditions are relaxed in *Case II* with respect to the presence of external trade opportunities. Here, the spatial structure has discrete cost curves due to external trading transactions at the economic boundary. Such formations can cause problems for comparative-static analysis. However, the potential problem is solved by taking the mean value of the cost curves as an approximation. As a similar method has already been demonstrated in conventional economic theory, in the case of a two-part tariff non-linear pricing system under the condition of price discrimination, the theoretical consistency is valid in this analysis. The condition is further relaxed in *Case III* with a dispersed plants structure in addition to an assembly plant at the centre of the economic space. This hypothesis enables reference to the presence of certain urbanisation economies and transportation costs. These additional factors are further investigated by means of a complex case in *Case IV* . The economic pattern is assumed to have upstream and downstream linkages, involving such factors as transaction and transportation costs, availability of inputs, dispersed patterns of output demands, and internal and external economies. In this multiple economic situation, the integrated framework analysis remains valid with the application of a production possibilities set, composed of generalised formal representations based on the input-output framework approach. From the standpoint of firms, these examinations may clarify the interaction between the optimal plant location and its relevant spatial configuration.

Regarding spatial competition, four further economic patterns are observed under the conditions of duopoly and oligopoly spatial competition. The simplest case is examined in *Case V* , where two firms are sharing the same market areas. This explores either rival or cooperative firm strategies. The analysis can be further expanded to a two-period game where an incumbent firm takes either negative profits at period one and positive profits after period two, or positive profits at period one and zero profits after period two. In addition, the Bertrand price competition model with product differentiation can also be applied in this framework. The case of partly-overlapping market areas is analysed in *Case VI* , introducing consumer-brand preferences and generating disutility curves in the spatial context. This unique solution enables investigation of location problems under the condition of consumer choice of particular products. Moreover, the duopoly and exclusivity of market areas are explored in *Case VII,* referring to the economic law of market areas and alternative space-filling spatial structures where consumer exclusions are present. The case investigates the effect of change in the available amount of demand, due to the appearance of space-filling new entrants, on the profit-maximising adjustment of incumbent firms. This hypothetical case is concluded by the presentation of spatial duopoly exclusive market-area equilibrium under the condition of symmetric market-area size. Although this model enables analysis of the asymmetric price condition by the movement of the 22.5° reflection line, the model can be further extended to asymmetric market-area size and price conditions.

The oligopoly and exclusivity of market areas are finally observed in *Case VIII.* This case shows that the economic space is hexagonally formed but can be formed circular when either a very high transportation burden is required in the case of distribution, or when diseconomies of scale are present in production. In this case, the economic space is not fully filled and vacant areas can be seen as potential economic space for new entrants. The equilibrium process is suggested in this analysis from the standpoint of both firms: on the one hand, incumbent firms reduce price levels of output in order to occupy the entire economic space; on the other, the new entrants try to have only one single plant for cost savings on fixed production facilities and set a *c.i.f.* pricing system. If the *c.i.f.* price of the new entrants is more expensive than the incumbent *f.o.b.* price, but less expensive in some areas, the new entrants can occupy these market areas according to spatial price competition. This new entrant's behaviour improves the level of consumer's surplus if consumer exclusion is present, by exclusive distribution of particular brands to limited market areas.

The findings can be summarised as follows. The analysis attempts to clarify the relationship and connection between market areas and supply areas. This research finds that there are difficulties in combining the two types of area caused by neglecting certain economic factors. One question raised is that the existing framework of market areas and supply areas has not yet required to take into account these neglected economic factors.

During this research, it becomes clear that existing analysis always assumes the production plant to be located at the centre of the market area. If inputs are ubiquitous and all other economic conditions are the same across the plain, this centre is also the centre of the supply area. As a result, the market area and supply area have the same centre. However, many situations can be found where the production plant is not situated at the centre of its market area. According to the Weberian approach this can be explained by the presence of spatially constrained economies and diseconomies. As a result, the theoretical framework of Weber needs to be integrated with spatial analysis. However, this approach considers not areas but single points of economic activity. As the relevant economic factors of the Weber problem can be incorporated into the variables of production and cost functions, it is possible to include the existing framework of market areas and supply areas. This research introduces these additional economic factors to existing market-area analysis and supply-area analysis. It also contributes to combining the independent frameworks of market areas and supply areas by applying the duality theory in conventional economic analysis to demonstrate that they are theoretically related to each other within the input-output framework.

The comparative-static results can be concluded as follows. In order to verify the theoretical accuracy of the hypothesis, comparative static analysis is applied in terms of spatial sizes, freight costs, production costs, densities of demand and input, and indices of internal and external economies. The results are basically consistent with the approaches of conventional aspatial economic conditions where the market and supply areas, and the spatially unconstrained and constrained internal and external economies, have certain relationships through the production function and the framework of the duality theory. Although the densities of demand and inputs are assumed to be constant for reasons of simplicity, it is possible to observe these spatial factors as dependent variables in order to examine more general spatial structures. This analysis also provides evidence showing the extent of the importance of the additional locational factors, with respect to the spatial constraints and spatial enhancement forces of economies. However, it should be noted that some hypothetical cases require dynamic analysis between upstream and downstream linkages or between earlier and later stages of processing. In addition, certain competition models of entries and exits of firms also need dynamic investigation by the framework of game theory. Such extensions are beyond the scope of this research; however, they can provide a basis for further in-depth investigation into location theory.

Regarding further avenues of research, although representative simplified models and hypotheses are generalised in this analysis, other complex exemplified hypotheses can also be generalised by the application of advanced microeconomic theory. First, product differentiated spatial duopoly cases can be examined on the framework of the Cournot-Nash and Bertrand-Nash models. In addition, for the oligopoly case, multi-stage Stackelberg quantity leadership game can be applied to find the spatial equilibrium condition. Related to the game approach, the decision-making between upstream and downstream linkages can be analysed by observing negotiation process and dominant strategies. Spatial industrial integration and dispersion, or operational integration and disintegration can also be examined on the framework of the transactions and contracts of firms. In order to observe the motion of individuals, firms, and local authorities in spatial context, these notions of the equilibrium concept should be applied to the analysis of this integrated framework approach. Finally, as demonstrated in the final part of this analysis, economies of scale and entry barriers of fixed costs can be further expanded with respect to the address model, which is the primitive spatial framework in conventional economic theory.

References

Beckmann, M. J. (1968) *Location Theory,* New York: Random House

Chamberlin, E. (1933) *The Theory of Monopolistic Competition: A Re-orientation of the Theory of Value,* London: Oxford University Press (Eighth Ed., 1966)

Coase, R. H. (1937) "The Nature of the Firm", *Economica,* IV/4:386-405

Cournot, A. (1838) *Recherches sur les Principes Mathématiques de la Théorie des Richesses,* [English Edition: *Researches into the Mathematical Principles of the Theory of Wealth,* New York: Macmillan, 1897]

Denike, K. B. and J. B. Parr (1970) "Theoretical Problems in Central Place Analysis", *Economic Geography,* 46/4: 568-586

Evans, A. W. (1972) "The Pure Theory of City Size in an Industrial Economy", *Urban Studies,* 9/1: 49-77

Fetter, F. A. (1924) "The Economic Law of Market Areas," *Quarterly Journal of Economics,* 38/3: 520-529

Hoover, E. M. (1937) *Location Theory and the Shoe and Leather Industries*, Cambridge, Massachusetts: Harvard University Press

Hoover, E. M. (1971) *An Introduction to Regional Economics,* New York: Alfred A. Knopf

Hotelling, H. (1929) "Stability in Competition", *Economic Journal,* 39: 41-57

Hwang, H. and C. C. Mai (1992) "Production-Location Decision and Free-Entry Oligopoly", *Journal of Urban Economics,* 31/2: 252-271

Hyson C. D. and W. P. Hyson (1950) "The economic law of market areas", *Quarterly Journal of Economics,* 64/2: 320-327

Isard, W. (1956) *Location and Space-Economy,* Cambridge, Massachusetts, and London, England: The M.I.T. Press

Khalili, A., V. K. Mathur and D. Bodenhorn (1974) "Location and the Theory of Production: A Generalization", *Journal of Economic Theory,* 9/4: 467-475

Kohlhase, J. E. and H. Ohta (1989) "General Equilibrium of Spatial Product and Labor Markets," *Journal of Regional Science,* 29/4: 537-553

Launhardt, W. (1885) *Mathematische Begründung von Volkswirthschaftslehre,* Leipzig: Verlag von Wilhelm Engelmann

Lösch, A. (1938) "The Nature of Economic Regions", *Southern Economic Journal,* 5/1:71-78

Lösch, A. (1954) *The Economics of Location,* New Haven, Yale University Press

Marshall, A. (1890) *Principles of Economics,* London: Macmillan and Co., Limited

Meade, J. E. (1952) "External Economies and Diseconomies in a Competitive Situation", *Economic Journal,* 62/245: 54-67

Mills, E. S. and M. R. Lav (1964) "A Model of Market Areas with Free Entry", *Journal of Political Economy,* 72/3: 278-288

Moses, L. N. (1958) "Location and the Theory of Production", *Quarterly Journal of Economics,* 72/2: 259-272

Nash, J. F. (1950) "The Bargaining Problem", *Econometrica,* 18/2:155-162

Nash, J. F. (1953) "Two-Person Cooperative Games", *Econometrica,* 21/1:128-140

Parr, J. B. (1993a) "Competition, Supply Areas and Industrial Location: An Equilibrium Perspective", *Annals of Regional Science,* 27/3: 191-210

Parr, J. B. (1993b) "Supply Areas and Optimal Spatial Structure", *Journal of Regional Science* 33/2: 167-186

Parr, J. B. (1995a) "Alternative Approaches to Market-area Structure in the Urban System", *Urban Studies,* 32/8: 1317-1329

Parr, J. B. (1995b) "The Economic Law of Market Areas: A Further Examination", *Journal of Regional Science,* 35/4: 599-615

Parr, J. B. (1997) "The Law of Retail Gravitation: Insights from Another Law", *Environment and Planning,* 29/8: 1477-1492

Parr, J. B. (2002a) "Agglomeration Economies: Ambiguities and Confusions", *Environment and Planning,* 34/4: 717-731

Parr, J. B. (2002b) "The Location of Economic Activity: Central Place Theory and the Wider Urban System", in P. McCann ed., *Industrial Location Economics,* Cheltenham: Edward Elgar

Parr, J. B. (2002c) "Missing Elements in the Analysis of Agglomeration Economies", *International Regional Science Review,* 25/2: 151-168

Parr, J. B. and J. K. Swales (1996) "Industry Cost Curves and the Spatial Organisation of Product", *Scottish Journal of Political Economy,* 43/3: 301-316

Parr, J. B. and J. K. Swales (1999) "Aspects of Competition in Spatial Context", *Journal of Regional Science,* 39/3: 233-243

Pontes, J. P. (1992) "Division of Labour and Agglomeration Economies", *Estudos de Economia,* XII/2: 123-132

Predöhl, A. (1925) "Das Standortsproblem in der Wirtschaftstheorie", *Weltwirtschaftliches Archiv,* XXI: 294-321

Reilly, W. J. (1929) *Methods for the Study of Retail Relationships,* University of Texas Bulletin No. 2944, Austin: University of Texas

Reilly, W. J. (1953) *The Law of Retail Gravitation,* New York: Pilsbury Publishers

Schneider, V. E. (1935) "Bemerkungen zu eniner Theorie der Raumwirtschaft", *Econometrica,* 3/1: 79-105

Scitovsky, T. (1954) "Two Concepts of External Economies", *Journal of Political Economy,* 62/2: 143-151

Shephard, R. W. (1953) *Cost and Production Functions,* Princeton: Princeton University Press

Smith, A. (1776) *The Wealth of Nations,* New York: Reprinted by Prometheus Books in 1991

Stackelberg, H. von. (1948) *The Theory of the Market Economy,* Edinburgh, 1952. Translated by Alan T. Peacock of *Grundlagen der theoretischen Volkswirtschaftslehre*

Stigler, G. J. (1951) "The Division of Labor is Limited by the Extent of the Market", *The Journal of Political Economy,* 59/3: 185-193

Weber, A. (1909) *Über den Standort der Industrien, Tübingen.* Translated and edited by C. J. Friedrich as Alfred Weber's Theory of the Location of Industries: Chicago, 1928

Worcester, D. A. (1969) "Pecuniary and Technological Externality, Factor Rents, and Social Costs", *American Economic Review,* 59/5: 873-885