Regional Studies, Vol. 37.6&7, pp. 649–663, August/October 2003

Carfax Publishing
Taylor & Francis Group

The Rise, Fall and Rise Again of Industrial Location Theory

PHILIP MCCANN⋆ and STEPHEN SHEPPARD†
⋆*Department of Economics, University of Reading, Whiteknights, PO Box 219, Reading RG6 6AW, UK.*
Email: p.mccann@reading.ac.uk
†*Department of Economics, Fernald House, Williams College, Williamstown, MA 01267, USA.*
Email: Stephen.C.Sheppard@williams.edu

(Received January 2003; in revised form March 2003)

MCCANN P. and SHEPPARD S. (2003) The rise, fall and rise again of industrial location theory, *Reg. Studies* **37**, 649–663. In this paper we will argue that new academic fashions, new international institutional arrangements, new communications technology and new developments in data availability, have all renewed the need for a redevelopment of analytical industrial location theory. Our paper will argue that the microeconomic foundations of industrial location theory must now be reconsidered. In particular, the methodological basis of traditional industrial location models needs to be reconciled with recent models of clustering, the new economic geography literature, and also more aggregate systemic levels of analysis. We will argue that in order to do this it is necessary: first, to specify the transactions-cost assumptions underlying these various approaches; second, to adopt broader definitions of spatial transactions costs; and third, to incorporate environmental characteristics within an orthodox location-production type of framework. The insights gained by this integrated micro-level approach will also provide new directions for more aggregate systemic approaches, as well as clarifying the both the strengths and limitations of the currently fashionable models of clustering.

Location Clustering theory Methodology Transactions Costs

MCCANN P. et SHEPPARD S. (2003) La montée, le déclin et la remontée de la théorie de la localisation industrielle, *Reg. Studies* **37**, 649–663. Cet article cherche à soutenir que les dernières modes académiques, les récentes dispositions institutionnelles internationales, la nouvelle technologie de la communcation, et les progrès dans le domaine de la disponibilité des données, ont tous renoué avec la nécessité de repenser la théorie analytique de la localisation industrielle. Cet article soutient qu'il faut remettre en question les fondements microéconomiques de la théorie de localisation industrielle. En particulier, on doit concilier la base méthodologique des modèles traditionnels de la localisation industrielle avec des modèles d'agglomération récents, la récente documentation sur la géographie économique, et aussi des niveaux d'analyse globaux plus généralisés. Pour y réussir, on soutient qu'il faut, primo, préciser les suppositions quant aux frais de transaction qui sous-tendent ces diverses façons; secundo, adopter des définitions plus générales des frais de transaction géographiques; et tertio, inclure des caractéristiques environnementales dans un cadre orthodoxe du type localisation-production. Les aperçus obtenus grâce à cette microanalyse intégrée fourniront de nouvelles possibilités pour le développement de plus de façons généralisées globales, ainsi que clarifier et les atouts et les inconvénients des modèles d'agglomération qui sont actuellement à la mode.

Localisation Agglomération Théorie Méthodologie Transactions Frais

MCCANN P. und SHEPPARD S. (2003) Aufstieg, Fall und Wiederaufstieg der Industriestandorttheorie, *Reg. Studies* **37**, 649–663. In diesem Aufsatz wird die Behauptung aufgestellt, daß neue akademische Moden, neue imternationale institutionelle Ordnungen, neue Kommunikationstechnologie und neue Entwicklungen in der Verfügbarkeit von Daten zusammen zur Notwendigkeit der Überarbeitung der analytischen Industriestandorttheorie beigetragen haben. Dieser Aufsatz stellt die Behauptung auf, daß die mikro-ökonomischen Grundlagen der Industriestandorttheorie jetzt neu überdacht werden müssen. Vorallem die methodologische Grundlage überkommener Industriestandortmodelle muß mit neueren Clustermodellen, mit moderner wirtschaftsgepgraphischer Literatur und auch mit stärker aggregierten systemischen Ebenen der Analyse in Einklang gebracht werder. Um dies zu erreichen, sei es zunächst nötig, die Vorstellungen, die den verschiedenenartigen Ansätzen zugrunde liegenden Durchführungskosten spezifisch aufzuführen; zweitens umfassendere Definitionen räumlicher Durchführungskosten einzuführen, und drittens charakteristische Umweltmerkmale in ein orthodoxes System vom Typ Standortproduktion einzubeziehen. Die Einsichten, die durch diesen integrierten Ansatz auf Mikroebene gewonnen wurden, werden auch neue Richtlinien für stärker aggregierte systemische Ansätze liefern, und sowohl Stärken als auch Grenzen der gegenwärtig modischen Clustermodelle erhellen.

Standort Clustern Theorie Methodologie Durchführungen Kosten

0034-3404 print/1360-0591 online/03/06&70649-15 DOI: 10.1080/0034340032000108741
http://www.regional-studies-assoc.ac.uk

INTRODUCTION

During the first six decades of the last century many significant breakthroughs were made in the analysis of industrial location behaviour. Yet, within the geographical school, traditional microeconomic location theory focusing on the interrelationships between production function models and the costs of transporting goods became largely out of fashion during the latter two or three decades of the twentieth century. Within the geographical literature this was in part because more aggregate levels of analysis, which emphasized the institutional and systemic changes within the production sectors, tended to dominate the literature. On the other hand, within most areas of economics, prior to the 1990s, spatial issues as a whole tended to be regarded at best, as minor issues. Where these issues were addressed, they tended to be discussed, ironically, in the non-spatial world of international trade theory. Yet, in the post war Bretton-Woods era of fixed exchange rates, high tariff barriers and limited factor mobility, this is to some extent understandable.

The recent renewal of interest in industrial location and trade, however, has spanned a range of disciplines including economics, geography, management science, marketing and international business studies. At the same time, these topics have also become central to discussions within not only the private sector but also within the governmental sector, at the local, national and even at the international levels. The reasons for this renewal of interest are varied, and include the rise of new international institutional arrangements surrounding economic integration and free trade areas, the growth in new information and communications technology, the growth and new developments in data availability, as well as simple changes in academic fashions and trends. While this renewed interest in location is to be applauded by spatial analysts, terms such as 'industrial clustering' or 'agglomeration' are now used more or less interchangeably by many commentators, with little real understanding of the origins, differences or actual meanings of these terms. Similarly, concepts such as 'new industrial areas', 'innovative milieux' or 'competitive clusters' are applied to different situations without any real consideration of the different empirical methods required in order to test whether such claims are justified. At the same time, there is often an inherent bias in the analytical perspective, in that the focus of such discussions is primarily on the advantages of industrial clustering, whereas the forces encouraging industrial decentralization are largely ignored.

From an analytical perspective this bias poses a real problem. The reason for this is that firm clustering and dispersion is a natural outcome of a random process in geographical space (ELLISON and GLAESER, 1997; GABAIX, 1999a, 1999b); even in a spaceless world, the distribution of firms is not even but random. Therefore, how are we to know whether observed firm location behaviour is a result of microeconomic responses to particular cost and revenue signals or simply a random outcome which is unrelated to the local economic environment? From both an analytical and public policy perspective, the most important issue is this: to be able to distinguish between the conditions under which industrial spatial concentration will take place and the conditions under which it will not. While the stylized models in favour of industrial concentration, such as agglomeration, clustering, new industrial areas, etc. can be very instructive, these models themselves require very careful interpretation in order to confirm that what we are observing is indeed what we are arguing. Without such interpretation, we cannot know the conditions under which the countervailing phenomenon of industrial dispersion will take place. In other words, we cannot specify a null hypothesis. Unfortunately, such careful interpretation is usually lacking, because the analytical limitations of these various models are typically overlooked, as are the empirical requirements needed to substantiate them.

Our paper will therefore argue that the microeconomic foundations of industrial location theory must now be reconsidered. The basis of traditional industrial location models needs to be reconciled with many of these newer developments mentioned above so that a more comprehensive framework can be provided within which observed location behaviour can be discussed. Similarly, many of these newer developments must be reconsidered in the light of the insights of traditional location theory models, many of which are as relevant today as they were when the models were first developed. In order to do this, it is first necessary to understand both the principal insights and the methodological issues raised by classical location models and then to relate these to many of the various locational approaches which are currently popular. As we will see, this requires us to reconsider the transactions-costs features and the inter-firm relations assumed by the modern models of industrial clustering, because it is also these features which determine the empirical data requirements needed to confirm or reject such models. This type of analytical clarification, which was fundamental in classical location theory but which is now largely absent in modern discussion of clustering, is essential in order to determine the situations in which industrial clustering or dispersion is advantageous.

On the basis of these discussions we will argue that the directions in which future location theory must be developed are three-fold. Firstly, we must take seriously the transactions-costs features and inter-firm relations assumed by the various models of industrial clustering in order to understand their analytical strengths and limitations. This will also us to determine the appropriate empirical methods of analysis in different locational contexts. Secondly, a review of the available evidence on spatial transactions costs suggests that while in some respects spatial transactions costs have fallen

over time, there are other respects in which spatial transactions costs appear to have actually increased over time. When building analytical models we must therefore treat spatial transactions costs in a rather more considered manner than has previously been typical. In particular, we will see that a simple dichotomy between transport costs and information costs is not sufficient, and that broader definitions of distance costs are required. Thirdly, the issue of the environment must be more coherently integrated within location theory. As we will see, however, this also requires us to consider the first two future directions of location theory. A movement towards a fusion of these various approaches will hopefully provide theories which are more appropriate in a world of falling trade barriers, rapid information flows and greater environmental awareness.

The paper is organized as follows. In the next section we outline the major insights of classical location theory and explain the approaches used to generate these insights. In the third section we discuss the explicit assumptions and methodological approach of this tradition, and then in the fourth section we explain the equivalent implicit assumptions embedded in modern theories of clustering. It will be argued that this is essential in order to understand the ways in which such theories should be developed or improved upon. In the fifth section we outline three specific ways in which location will need to be developed in order to address the kinds of spatial economic questions arising in the coming decades.

THE LOCATIONAL INSIGHTS OF CLASSICAL AND NEOCLASSICAL LOCATION THEORY

The classical fathers of location theory such as VON THUNEN, 1826; LAUNHARDT, 1885; and WEBER, 1909; came out of a very particular nineteenth century tradition of economic analysis, in which the focus of analysis was primarily on the nature and characteristics of the production process. As such, the essential questions all concerned the ways in which factor inputs were transformed into physical commodities, and this tradition aimed to understand the determining features of this transformation process. Within this broad production-focused tradition, the distinguishing feature of the early location work was an attempt to reveal the relationship between issues of geography and production. Here, the notion of geography was understood in terms of that of land and land use, and the tradition investigated the unique role which land as a factor input plays in determining the characteristics of the transformation process.

The general approaches employed in these classical analyses can be split broadly into two types. The first type of classical locational analysis was based on a Ricardian land use tradition in which the conditions underpinning the profitable cultivation of land were the major focus of analysis. Here the work of Von Thunen extended the original Ricardian system by differentiating between land, not on the basis of differing quality, but on the basis of differing location. By integrating the themes of place-specific market prices, transportation costs and rental payments, Von Thunen arrived at the classic conclusion of concentric rings of different types of cultivation around a central (urban) market location. Within mainstream microeconomic analysis, however, the key conclusion of Von Thunen's concentric-circle result is that an upward-sloping supply curve for a product can be generated simply by way of changes in land use. In particular, higher market prices for a good will allow for higher potential rental payments, which consequently encourage larger areas of land to come under cultivation, thereby producing larger output quantities. An upward-sloping supply curve is therefore generated simply by issues of land use, without any explicit modern notions of factor substitution or marginality conditions.

The second type of classical locational analysis can be characterized in modern terms primarily as a form of production function type of analysis. In this tradition of LAUNHARDT, 1885, and WEBER, 1909, explicit relationships are defined between the quantity of outputs produced and the quantity of inputs required. The location problem is then defined specifically with respect to these fixed input–input and fixed input–output relationships, and is conceived of primarily as a transportation problem. Here, the focus is on the cost minimizing conditions required in order to produce and ship a particular quantity of a good from the production location to the market, while accounting for the transportation costs involved in the delivery of the requisite inputs. In these models there is, once again, no real explicit notion of either factor substitution or marginality conditions.

There are two major differences between these two location theory traditions. The first difference is in terms of the real-world motivations for the objects of their questioning. In the Ricardo-Von Thunen tradition of the early nineteenth century, the object of the analysis was the development of agricultural hinterlands around urban market centres. At this time, the German states still exhibited primarily an agrarian economy with an inherited spatial industrial structure which was largely unchanged from the middle ages. Agriculture and the behaviour of the (urban-located) agricultural markets was still the dominant feature of the Germanic economies, and this location-theory tradition therefore attempted to understand the interactions between land use, commodity prices and agricultural transformation processes. By the second half of the nineteenth century, however, the industrial transition of Germany was well under way. In the Launhardt-Weber tradition, the primary motivation for the location-theory analyses was therefore the need to understand the relationship

between the behaviour of nineteenth century manufacturing enterprises, and the rapid growth of the German urban-manufacturing centres. As such, the focus of this location tradition was on the dominant aspects of the transformation processes which took place within manufacturing rather than in agriculture.

The second major difference between these two classical location theory approaches is in terms of the notions of geography they employ. In the Ricardo-Von Thunen tradition, the notion of geography employed is one of land use and the variations in the productivity of land within basically, a one-dimensional spatial framework. Distance-transportation costs are introduced here simply in order to differentiate between land on the basis of the location-specific rental payments achievable at each location. As such, location-specific variations in residual land rewards are the natural outcome of geographical space. These one-dimensional conclusions are then translated directly into two-dimensional results. On the other hand, in the Launhardt-Weber tradition, the notion of geography employed is primarily that of two-dimensional space, in which geography is understood within an explicitly geometric framework. Here, land is largely undefined and variations in the productive and qualitative characteristics of the land are ignored. As such, in this production-function type of tradition, geography is not understood in terms of its effects of land use factor rewards in the way it is in the Ricardo-Von Thunen tradition. Rather it is understood in terms of the effects of spatial transactions costs on the overall profitability and factor rewards, with no explicit role for land *per se.* In this Launhardt-Weber type of analysis, location-specific labour inputs and land are viewed largely as a composite input, whereas in the Von Thunen tradition, land and all non-land inputs are treated as a composite input.

The methodological approach in this classical location theory approach is to distinguish between the costs incurred by the firm which are explicitly distance-related from those costs which are location-specific. In other words, different notions of geography are used here to distinguish between various cost components, in that some costs are specified as functions of distance while others are specified as functions of location. The costs which are distance-related are assumed to be transport costs and the location-specific costs are the local labour and land costs. By splitting up the various geographical aspects costs in this manner and then defining the cost relationships in terms of a production function we can then proceed to solve the location problem. In a fixed-coefficients framework with given market locations and prices, all the possible areas of profitable production are identified,[1] and within this subset, the maximum profit location can be calculated. Moreover, we are also able to define the conditions under which the location of a firm will change. However, the fact that this is so is precisely because the classical location models explicitly distinguished between the costs incurred by the firm which are distance-related from the costs which are location-specific.

Classical location theory therefore provides us with our first fundamental principle of location-theory:

> *Location theory principle* 1. In situations where both locational coordinates and production technology exhibit stable and identifiable fixed-coefficients relationships, if transportation costs are known or can be calculated, then the fixed-coefficients assumptions embedded in all these models allows the equilibrium factor conditions to be determined. Alternatively, for any given set of factor prices, where production technology is governed by stable and identifiable fixed-coefficients relationships the conditions under which firms will be mobile can also be determined.

The classical location tradition evolved without any explicit notions of either factor substitution or marginal productivity conditions, because such concepts were themselves only gradually emerging within late nineteenth century economics (BLAUG, 1968). As such, modern microeconomic principles such as variable factor proportions and well-behaved factor–factor relationships are almost entirely absent from these analyses. Consequently, the relationships between inputs and outputs and between inputs and inputs are all specified in fixed terms which are independent of the level of output. In modern terminology, we are in a world of fixed coefficient production functions. On the other hand, there were some principles of marginality embedded in these constructs, but these are slightly different in emphasis in that they relate primarily to the levels of profits generated and the spatial dimensions or limits at which such profits fall to zero. In essence, we can therefore consider these notions of marginality primarily in terms of defining the spatial boundary conditions within which profitable production takes place.

The classical location theory models were not given a genuine neoclassical twist until the middle of the twentieth century. The work of ALONSO, 1964; MUTH, 1969; MILLS, 1970; and EVANS, 1973, reinterpreted the fixed coefficients world of Von Thunen within a framework of mutual factor substitutability. The variable proportions principles and well-behaved factor relationships employed within neoclassical microeconomics were now incorporated within a Von Thunen-type framework in order to build a neoclassical model of land use. In these new models of location and land use factor substitution is assumed to take place between land and a composite factor, capital, which represents all non-land inputs. In that sense, these models share exactly the same approach as the original Von Thunen model. Meanwhile, neoclassical developments also greatly extended the analytical insights of the classical production-function type of location theory.

Following on from the initial work of ISARD, 1956, the fixed coefficients model of Laundhardt-Weber was greatly extended within a substitution framework by the production-location model of MOSES, 1958, and its subsequent developments (MILLER and JENSEN, 1978; ESWARAN *et al.*, 1981). The Moses fusion of a neoclassical production function with a Weber-type location problem provided for a completely new set of locational conclusions, the specific outcomes of which depend on the nature of the firm's production function.

As we have seen, in classical location models the technical factor relationships are exogenously fixed, and this specification defines the nature of the firm. On the other hand, in the neoclassical models this requirement is lifted. Instead, the various *alternative* factor relationships are specified in the neoclassical framework, as defined by the nature of the firm's production function. The neoclassical location theory models therefore differ fundamentally fro the classical location theory models in the way that they specify technical relationships. As before, the neoclassical location theory approach is to distinguish between the costs incurred by the firm which are explicitly distance-related from those costs which are location-specific. Once the technical substitution relationships are defined along with the input–output relationships by the production function specification, then the location problem can be solved accordingly.

The most general and fundamental insight of this neoclassical tradition provides us with our second principle of location-theory:

> *Location theory principle* 2. In situations where production technology is governed by principles of factor substitution, all location-optimization problems can be considered to be production-optimization problems, and all production-optimization problems can be considered to be location-optimization problems. In other words, the various technical relationships do not need to be stable, but do need to be identifiable, for such models to be tractable. Only the locational coordinates need to be both identifiable and stable.

A fundamental feature which is common to both of these classical and neoclassical location theory traditions is the fact that these analytical approaches generally accept the market location and market prices as given. Neither microeconomic approach makes any real attempt to provide an explanation of the nature of market locations or market prices. Nor do they make any real attempt to explain why many activities are often grouped together in geographical space. Rather, the industrial or urban clustering of activities is simply assumed to exist primarily for historical reasons or because of topographical features.[2] The arguments underpinning agglomeration behaviour (MARSHALL, 1920; HOOVER, 1948) are generally treated as being exogenous to the micro-location problem. Once these assumptions are made and prices are assumed to be given, then the microeconomic location issues can be worked out accordingly on the basis of the exogenously-specified technical factor substitution and input–output relationships. Both of these approaches therefore fail to provide any reasons why firms will group together in space, except for the situation where independent behaviour generates the same optimum locations for similar firms (MCCANN, 1995a). Yet, for such geographical grouping of firms to be maintained the firms must be willing to tolerate increased local land and labour prices, in situations where the classical and neoclassical models would unambiguously point to dispersion as rational behaviour. Why would this be so? On this particular point, microeconomic location theory provides no answer, and it is necessary to turn to the literature on agglomeration and clustering discussed below. However, where such clustering does exist, the third strand of classical microeconomic location theory does provide a fundamental insight into the conditions under which such a phenomenon is or is not possible. This third strand of classical location theory is known as 'market area analysis'.

This third strand of location theory, namely market area analysis, arose during the early part of the twentieth century out of the literature concerning inter-firm oligopolistic rivalry. The market area work of HOTELLING, 1929, and PALANDER, 1935, provided the basis for the development of a range of models which investigated the role played by location and space in determining the nature of inter-firm competitive behaviour in situations in which the firms are interdependent. These models defined markets explicitly in spatial terms by assuming a geographical distribution of customers, rather than by simply defining a market as a single point in space. The outcomes of these models suggested that the spatial characteristics of a market itself contributes to the nature of the competitive behaviour, in that there are dynamic interrelationships between the spatial nature of the market, the type of industrial structure and the various strategies employed by the competitors. Such models have tended to be developed by means of a game theoretic approach in which we explicitly specify the interrelationships between the firms. From the perspective of location theory, one key outcome has arisen from this approach. This was demonstrated by D'ASPREMONT *et al.*, 1979, who showed that there is no stable Hotelling-type solution in an environment in which prices are flexible. In other words, in situations where price adjustments can be considered to be one of the competitive strategies available to the oligopolistic firm, geographical clustering is not a natural outcome, because of the Bertrand problem of downward price spirals.

We can turn this analytical conclusion on its head, however, in order to understand one of the necessary conditions under which geographical co-location

would indeed be tenable. Geographical proximity and industrial clustering are only possible in situations in which price competition for identical products is not a major feature of the market behaviour.

Spatial competition and market area analysis therefore provides us with our third and fourth fundamental principles of location theory:

> *Location theory principle* 3. The specific reasons for, and outcomes of, industrial co-location or clustering, cannot be understood without a consideration of the cost interrelationships between the co-located firms.

> *Location theory principle* 4. Industrial co-location or clustering can only take place in markets characterized primarily by non-price competition, product heterogeneity, or the existence of transactions costs which militate against efficient price competition.

Taken together, our four location theory principles provide profound insights into not only the nature of firm location behaviour, but also the theoretical–analytical requirements necessary in order to discuss such behaviour. Moreover, the validity of these analytical principles is independent of the actual industrial sector or geography being analysed. However, for our purposes, the important point is that these fundamental insights are generated because of the specific way in which the analytical outcomes are arrived at. In classical and neoclassical location theory, the way in which these principles are derived is by explicitly specifying the production characteristics of the firm, and thereby implicitly defining the organizational characteristics of both the firm and the production-consumption hierarchy within which the firm is placed (McCann, 1995a). All that is required to do this is to define a production function for the firm and then to exogenously specify the location or spatial definition of the input and output markets. Once this production–organizational system has been defined, the geographical alternatives can then be specified and the location problem can be solved. On the other hand, in the case of the market area models, the competitive or collaborative relationships between the interdependent firms are specified, along with the technical cost conditions. Once these are defined, then the geography-revenue-profit conditions are worked out. Yet, what is common to each of these modes of analysis is that both the inter- and intra-organizational relationships are explicitly specified within a geographical setting either by means of a production function of by means of the game theory conjectures. As long as these internal or external relationships are specified, then the geographical behaviour of the firm can be analysed and understood. Without such specifications, the models would be intractable because both the production relations and geographical relations would have to be determined simultaneously.

Within each particular geography-production-organization frame of reference, this classical methodological approach is therefore both coherent and consistent. The reason is that by defining the geography-production-organization frame of reference of each model structure, it is theoretically possible to specify a null hypothesis which can be tested empirically. In other words we can make specific predictions about the nature of micro-location behaviour which in principle at least is testable. However, these methodological principles also raise questions and problems which are more generally applicable to all types of microeconomic industrial location analysis. The reason for this is that many modern models of economic geography aim to discuss location behaviour without explicitly asking questions about the nature of the production or organizational characteristics of the firms or the relations between the firms. This is particularly problematic in the recent burgeoning literature on industrial clustering. These problems are discussed in the following two sections.

THE METHODOLOGICAL INSIGHTS OF CLASSICAL AND NEOCLASSICAL LOCATION THEORY

If we combine the principles of classical and neoclassical location theory with those of spatial competition models, we can see that all three types of models are able to produce insights which are generally applicable to all location problems. The reason for this is that these models are all based on an implicit methodological principle: the position of the firm within the production hierarchy (Williamson, 1975) is exogenously given. The location problems are then specified simply in terms of optimizing the relationship between the firm and the spatial transactions costs it faces. The definition of the firm itself, in terms of determining whether the transactions it undertakes should be internal or external to the firm, is not discussed; rather, these issues are simply assumed to be given in terms of the production function relationships of the firm or the game conjectures. Once these assumptions have been made, then what we can explicitly do in location theory is discuss the optimization of these relationships with respect to exogenously given geographical locations of input sources or markets is discussed (McCann, 1995a, 1999). The reason why such hierarchy issues are not discussed in all three location theory traditions, is that the location or production optimization problems can only be solved if at least either the locational relationships or the production relationships are exogenously determined. In the case of classical location theory, both the exact production relationships and the relationship between the geographical coordinates are exogenously specified; in the case of neoclassical location theory only the exact geographical coordinates are exogenously specified,

while the various production function alternatives are specified; in the case of market area analysis, only the geography is specified. In none of these three model types is the definition of the firm itself a decision variable. In other words, there is nothing related to the markets-and-hierarchies problems (WILLIAMSON, 1975) in microeconomic location theory, because if such problems were included, then the models would become intractable. The fundamental problem this raises is therefore that if neither the geographical relationships nor the production relationships are exogenously specified, then each of these three strands of location theory falls apart, and it becomes impossible to specify either a location problem or an empirical assessment of location behaviour.

This requirement of an analytical reference framework, of either a geographical or structural-hierarchical nature, is the essential difference between microeconomic models of location theory and various other less specific theoretical approaches to location analysis, such as the Marxist, structuralist, or 'post-Fordist' traditions. Yet, we would argue that the methodological requirement of a reference structure, irrespective of whether this reference is structural or geographical, provides an analytical rigour to the field of microeconomic location theory. Adopting this approach means that we can make predictions about the location behaviour of individual firms which, at least in principle, ought to be testable. Moreover, adopting this approach also means that the empirical requirements for testing such predictions can be specified. In the case of classical, neoclassical and market area location theories, the empirical requirements for testing the theories relate to the calibrating the relationship between transport costs and the technical production relationships, because transport costs are assumed in these models to be the spatial transactions costs faced by the firm. In more modern situations, however, it may well be that the spatial transactions costs faced by the firm are rather more complex than simply transport costs, and that broader or more sophisticated definitions of spatial transactions costs must be adopted. Yet this does not mean that such approaches are somehow out-of-date, because the methodological requirement of a geographical or structural reference framework is still essential in order to analyse firm location behaviour. Without such a reference framework it becomes impossible to speak about the relationship between the production behaviour of the firm and the spatial transactions costs it faces. And yet, only with such a discussion can we identify whether observed location behaviour is optimal or even rational.

The fact that these location theory methodological and analytical insights are generally applicable to all aspects of industrial location also reflects a broad movement within the overall science of location analysis. Early classical location models were developed in order to discuss the behaviour of particular sectors, whereas more recent neoclassical models have been adapted in a largely analogous manner to the industrial, commercial or residential sectors, in the hope that a single analytical toolkit can be used for a range of different questions. For example, in the neoclassical framework, the Von Thunen-type of logic is applied primarily to the case of urban land development rather than to agricultural development; here, the Von Thunen notion of an agricultural hinterland is replaced with the notion of an urban hinterland, in which 'hinterland' here refers to the phenomenon of suburban development. In a world in which the majority of people live in urban areas this was a major breakthrough, in that this schema provided for the possibility of understanding the economic nature of the living environments most people construct for themselves or find themselves in. Similarly, in the neoclassical location-production models, the actual nature of the inputs and outputs included in the production functions remain largely undefined. This is implicitly accepted so as to provide a more general framework for understanding the location behaviour of all types of firms. While there are some problems with this approach (MCCANN, 1999), the neoclassical location-theory research programme explicitly attempts to provide a common framework for analysing the location behaviour of all of the various sectors of the economy.

This movement towards providing an analytical framework which is intended to be generally applicable to all sectors and locations has been most evidently been observed in the recent work of the 'new economic geography' research programme (FUJITA *et al.*, 1999) and also the recent literature on industrial clustering (PORTER, 1990, 1998). Both of these analytical frameworks are based on the crucial assumption that there are many situations in which it is advantageous for firms to group together in space. However, unlike classical or neoclassical theory, what neither of these frameworks is able to do is to explicitly distinguish between the situations in which it is advantageous for an *individual* firm to locate close to other firms from the circumstances in which it is disadvantageous for the individual firm to do so. Nor does either framework provide a means by which we can identify which particular type of firm will benefit from which particular clustering behaviour. The reason for this is that neither the new economic geography nor the clustering literature explicitly discusses the production function nature of the individual firm;[3] nor does either literature distinguish between the nature of the spatial transactions costs faced by the firm from those production costs which are location-specific. Finally, in tandem with both classical and neoclassical location theory, neither the new economic geography nor the clusters literature raises the issue of the hierarchical definition of the firm, and the problem of whether transactions should be external or internal to the firm. Yet, in discussing the questions of industrial clustering from a

microeconomic perspective, these are actually the central issues. As we have seen, one of the methodological lessons from classical and neoclassical location theory is that only by exogenously defining the production and organizational nature of the firm and its relations can we begin to discuss the geographical aspects of the firm. Alternatively, by exogenously defining the geography of the firm we can begin to discuss the production, organization and relational issues faced by the firm. What we cannot do is leave *both* the geography of the firm and its technical production nature undefined, because such an analytical problem cannot be solved. Only by asking such questions can we hope to understand and distinguish the microeconomic conditions under which clustering is advantageous from the situations in which it is disadvantageous.

The apparent lack of microeconomic definition and rigour in both the new economic geography and the clusters literature means that there is at present a whole range of microeconomic problems which current thinking on agglomeration and clustering is largely unable to deal with. The result of this is that terms such as 'agglomeration', 'cluster', 'milieux', 'localization, 'complex' and 'network' are frequently used in much of the modern geographical literature in a more or less interchangeable manner, without any real consideration of the microeconomic underpinnings of these different concepts. It is simple assumed *ex ante* that clustering or co-location is beneficial, and this has lead to such arguments being viewed as a panacea for all regional problems (MARTIN and SUNLEY, 2001). As such, regional policies formulated on the basis of these arguments, which are explicitly intended to influence firm location behaviour, are often built on very weak analytical frameworks. Yet, underlying these various analytical models there are indeed specific assumptions concerning the nature of the firm, its production and organizational characteristics and the nature of the transactions costs it faces. Unfortunately, these assumptions are largely unspecified, yet from the perspective of location theory, it is necessary to make these assumptions explicit in order to define the terms of reference within which the location problem can be set. For these reasons, we would argue that the methodological approaches and insights of the classical and neoclassical microeconomic traditions have profound, and often overlooked, implications for our understanding and interpretation of both of these modern literatures.

If we follow the approach of classical and neoclassical location theory in order to help correct for these problems of a lack of microeconomic rigour, it is first necessary to make explicit the implicit assumptions which underlie the different theoretical descriptions of industrial clustering or spatial grouping that we have at our disposal. In particular we must specify the implicit assumptions regarding the organizational nature of the individual firm and the nature of the interrelations between the co-located firms. Secondly, we must explicitly define the nature and behaviour of the spatial transactions costs faced by modern firms. Only by doing this can we understand the both the applicability and limitations of the location models we are using. Without such an understanding our spatial conclusions will be neither verifiable nor refutable. Thirdly, in order to point to the possible future developments required in location theory, we must consider how both the interdependencies between firms have changed over recent years, as well as how the nature of the spatial transactions costs faced by the firms have changed over recent years.

In the next section we explain the implicit assumptions which underlie the different theoretical descriptions of industrial clustering, with particular reference to the nature of the organization of the firm, the interdependencies between firms and the inter-firm transactions costs they face. In the subsequent sections we then discuss the recent changes in modern spatial transactions costs in order to point to possible ways in which future location theory should be developed.

THE MICRO-FOUNDATIONS OF INDUSTRIAL CLUSTERING

In order to explain the implicit assumptions concerning firm interdependencies and spatial transactions costs which underlie the various theoretical descriptions of industrial clustering which are currently popular, it is necessary to adopt a transactions costs approach of a type proposed by WILLIAMSON, 1975. However, in our geographical-structural framework it is not sufficient to distinguish between the transactions which are internal to or external to the individual firm, as in the original Williamsonian framework. In adopting the methodological approach of classical location theory it is necessary to distinguish between the transactions which are location-specific from those which are not. In other words, we must explicitly distinguish between the transactions which are internal to, and external to, the spatial industrial cluster. In order to do this we can present three stylized sets of geography-firm-industry organizational relationships evident in situations where firms are clustered or co-located (SIMMIE and SENNET, 1999; GORDON and MCCANN, 2000; MCCANN, 2001a) which we term the *pure agglomeration model*, the *industrial complex model* and the *social network model*. These three stylized characterizations of industrial clusters are distinguished in terms of the nature of firms in the clusters, the nature of their relations, and transactions undertaken within the clusters. What we are not doing here is presenting a new theory of clustering. Rather, our reason for this characterization is to categorize the different transactions cost and firm behavioural assumptions which are implicit in each of the various analytical frameworks which are currently

Table 1. *Industrial clusters*

Characteristics	Pure agglomeration	Industrial complex	Social network
Firm size	Atomistic	Some firms are large	Variable
Characteristics of relations	Non-identifiable	Identifiable	Trust
	Fragmented	Stable and frequent trading	Loyalty
	Unstable		Joint lobbying
	Frequent trading		Joint ventures
			Non-opportunistic
Membership	Open	Closed	Partially open
Access to cluster	Rental payments	Internal investment	History
	Location necessary	Location necessary	Experience
			Location necessary but not sufficient
Space outcomes	Rent appreciation	No effect on rents	Partial rental capitalization
Example of cluster	Competitive urban economy	Steel or chemicals production complex	New industrial areas
Analytical approaches	Models of pure agglomeration	Location-production theory	Social network theory
		Input–output analysis	(Granovetter)
Notion of space	*Urban*	*Local or regional but not urban*	*Local or regional but not urban*

popular. In reality, all spatial clusters or industrial concentrations will contain characteristics of one or more of these ideal types, although one type will tend to be dominant in each cluster. Understanding what are the dominant features of a particular cluster is essential in order to work out how we are to analyse it either theoretically or empirically.

The characteristics of each of the cluster types are listed in Table 1, and as we see, the three ideal types of clusters are all quite different. Firstly, in the model of pure agglomeration, inter-firm relations are inherently transient. Firms are essentially atomistic, in the sense of having no market power, and they will continuously change their relations with other firms and customers in response to market arbitrage opportunities, thereby leading to intense local competition. As such, there is no loyalty between firms, nor are any particular relations long term. The external benefits of clustering accrue to all local firms simply by reason of their local presence. The cost of membership of this cluster is simply the local real estate market rent. There are no free riders, access to the cluster is open, and consequently it is the growth in the local real estate rents that is the indicator of the cluster's performance. This idealized type is best represented by the MARSHALL, 1920, model of agglomeration, as adopted by the new economic geography models (KRUGMAN 1991; FUJITA *et al.*, 1999). The notion of space in these models is essentially urban space, in that this type of clustering only exists within individual cities.

Secondly, the industrial complex is characterized primarily by long-term stable and predictable relations between the firms in the cluster, involving frequent transactions. This type of cluster is most commonly observed in industries such a steel and chemicals, and is the type of spatial cluster typically discussed by classical (WEBER, 1909) and neoclassical (MOSES, 1958) location-production models, representing a fusion of locational analysis with input–output analysis (ISARD and KUENNE, 1953). Component firms within the spatial grouping each undertake significant long-term investments, particularly in terms of physical capital and local real estate, in order to become part of the grouping. Access to the group is therefore severely restricted both by high entry and exit costs, and the rationale for spatial clustering in these types of industries is that proximity is required primarily in order to minimize inter-firm transport transactions costs. Rental appreciation is not a feature of the cluster, because the land which has already been purchased by the firms is not for sale. The notion of space in the industrial complex is local, but not necessarily urban, and may extend across a sub-national regional level. In other words, these types of complexes can exist either within or far beyond the boundaries of an individual city, and depend crucially on transportation costs.

The third type of spatial industrial cluster is the social network model. This is associated primarily with the work of GRANOVETTER, 1973, and is a response to the hierarchies model of WILLIAMSON, 1975. The social network model argues that mutual trust relations between key decision-making agents in different organizations may be at least as important as decision-making hierarchies within individual organizations. These trust relations will be manifested by a variety of features, such as joint lobbying, joint ventures, informal alliances and reciprocal arrangements regarding trading relationships. However, the central feature of such trust relations is an absence of opportunism, in that individual firms will not fear reprisals after any reorganization of inter-firm relations. Trust relations between key decision makers in different firms are assumed to reduce inter-firm transactions costs, because when such trust-based relations exist, firms do not face the problems of opportunism. As such, these trust relations circumvent many of the information issues raised by the markets

and hierarchies dichotomy (WILLIAMSON, 1975). Where such relations exist, the predictability associated with mutual non-opportunistic trust relations, can therefore partially substitute for the disadvantages associated geographic peripherality. Inter-firm cooperative relations may therefore differ significantly from the organizational boundaries associated with individual firms, and these relations may be continually reconstituted. All of these behavioural features rely on a common culture of mutual trust, the development of which depends largely on a shared history and experience of the decision-making agents.

This social network model is essentially aspatial, but from the point of view of geography, it can be argued that spatial proximity will tend to foster such trust relations over a long time-period, thereby leading to a local business environment of confidence, risk-taking and cooperation. Spatial proximity is thus necessary, but not sufficient to acquire access to the network. As such, membership of the network is only partially open, in that local rental payments will not guarantee access, although they will improve the chances of access. In this social network model, space is therefore once again local, as with the complex, but not necessarily urban, and often extends over a sub-national regional level. Once again, in this case, both information transactions costs and transportation costs may play a role in determining the importance of geographical peripherality. The major geographical manifestation of the social network is the so-called 'new industrial areas' model (SCOTT, 1988), which has been used to describe the characteristics and long-term growth performance of areas such as the Emilia-Romagna region of Italy (PIORE and SABEL, 1984; SCOTT, 1988).[4]

FUTURE DIRECTIONS FOR LOCATION THEORY

On the basis of the previous sections we argue that there are three primary directions in which future location theory must develop:

Direction 1: *Acknowledging the micro-foundations of clusters.* Using a transactions-costs approach in order to understand and define the spatial-industrial nature of an industrial cluster is the first major problem which future location theory will need to address. This approach is entirely in keeping with that of the classical location approaches, the only difference being that the context of analysis is rather different. As we see, each of these different model types has a different notion of geography, a different notion of the production-organizational arrangements of the firms, and a different notion of the transactions costs faced by the firms. It is important for us to be aware of these fundamental differences, because the techniques that we should adopt to analyse each of these cluster types are quite different. Although in reality many observed spatial groupings of activities will exhibit characteristics of more than one of the stylized characterizations here, it is still necessary for us to decide what is the dominant nature of the cluster, because this will determine the appropriate empirical technique for us to use. For example, the financial services of the City of London exhibits primarily the characteristics of the pure agglomeration model, although there are also some secondary characteristics associated with the social network model (GORDON and MCCANN, 2000). In this case, real estate capitalization techniques and aggregate production function methods will probably be the most appropriate tools of empirical analysis, supported by various case study techniques. A second example of this concerns the PORTER, 1990, 1998, description of clustering. Although Porter assumes that the dominant competitive effects of clustering are mediated by information flows between firms and individuals within the urban sphere, the primary effect of which is to stimulate local competition by increasing the transparency associated with competitive improvements, he also acknowledges that such information flows are mediated by strong interpersonal networks which may well extend beyond the urban scale in situations where such trust-relations exist. As such, the clustering description of PORTER, 1990, 1998, can also be argued to fit primarily into the pure agglomeration model but also appears to exhibit secondary characteristics associated with the social network category. The empirical techniques used to analyse the Porter model will therefore be similar to those used to analyse the City of London. On the other hand, a third example is the case of Silicon Glen in Scotland. In this particular spatial grouping of firms, the dominant characteristics of the cluster are that of the industrial complex, with some secondary aspects of the social network model embedded within the system. Here, traditional location theory techniques and input–output analysis will be most appropriate, allied with some case study work. In each of these cases, although we acknowledge that the spatial industrial groupings are highly complicated, what we are doing is to determine the dominant characteristics of the inter-firm relationships in order to determine how we are to most appropriately analyse the group empirically. Using an industrial complex model based on input–output techniques in the first two cases will actually tell us very little about the nature of either of these particular clusters; similarly, using a pure agglomeration model in the third example will also tell us very little. In situations in which we have widespread case study evidence, the choice of model will be made relatively easy. However, in situations where case study evidence is lacking, we can use the framework outlined in Table 1 to help us interpret the empirical data available so we can understand the nature of the industrial cluster, without necessarily ascribing a particular model structure to it *a priori* (GORDON and MCCANN, 2000). Adopting this classical type of

methodological approach will at least allow us to say something testable.

Direction 2: Redefining the nature of modern spatial transactions costs. The second problem facing future location theory is to reconsider how modern spatial transactions costs are incorporated into location models. There have been fundamental changes in the nature and levels of spatial transactions costs which have become very evident over recent years. Yet, rather than reducing the importance of location problems it can be argued that in many cases these changes have actually increased the importance of location as an economic issue. Therefore it is necessary for these changes to be incorporated into future models. To see this, we can split up modern spatial transactions costs into two major groups, because the spatial transactions costs faced by firms are primarily of two types: information transmission costs; and transportation costs.

Since the 1980s we have seen dramatic improvements in the ability of decision makers and planners to coordinate activities across space. The primary reasons for these improvements have been the enormous technological developments in information technology, and also the advent of widespread usage of these technologies. These developments have meant that complex operations can now be managed both more efficiently and effectively than was previously possible. There are two aspects to these developments.

Firstly, the new information technologies have reduced the real costs of communicating across distance, allowing us to more efficiently control existing spatial arrangements of activities (*The Economist*, 1999a). This is a common observation in industrial sectors and activities where physical commodities are being moved across large distances, such as in the management of international importing and exporting supply chains (*Financial Times*, 1999) or the coordination of multinational manufacturing activities (*The Economist*, 1999b). Analogous arguments also exist for the case of the service sectors, in situations where information rather than physical goods is being transferred across space. Secondly, the existence of these new information technologies also allows decision makers to undertake the coordination of spatial arrangements of activities which were previously not possible. On the other hand, however, there are some other arguments which suggest that over time the development of these information technologies is actually leading to increases in the costs of transmitting information across space, thereby increasing the relative importance of geographical centrality. The argument here is that an increase in the quantity, variety and complexity of information produced itself increases the costs associated with transmitting this information across space. This is because much of the information will be of a non-standardized tacit nature, and the transmission of this type of information essentially requires face-to-face contact. The opportunity costs involved in not having face-to-face contact will consequently increase with the quantity, variety and complexity of the information produced (McCann 1995b; Gaspar and Glaeser, 1998). The effects of this will be to increase the costs of doing business across large geographical distances.

Meanwhile, transportation technologies have improved dramatically over recent years. Obvious examples of this include the growth in roll-on roll-off trucking, containerization, rapid-turnaround shipping, and the increased efficiency and frequency of airline services. On the other hand, the quantity, variety and complexity of market information generated in the modern economy is increasing. This also implies that in many industries which involve the production or shipping of goods across space, the variety and complexity of the logistics operations being undertaken will also increase. As the demand for delivery speed increases, the associated opportunity costs of lead-times also increase, and the average inventory levels maintained will fall. Analytically, the effect of this is to increase the transactions costs associated with shipping goods over any given distance.[5] The most extreme example of this trend towards more frequent shipments, is the application of just-in-time (JIT) manufacturing and distribution techniques, the influence of which has pervaded all areas of modern production, distribution and retailing. In the new JIT production and distribution arrangements (Nishiguchi, 1994; Schonberger, 1996), it is necessary to control the flows of goods between firms to a very high degree, in order to ensure the timeliness of deliveries. New information technologies allow firms to coordinate logistics activities across huge geographical areas in a very sophisticated and timely manner. Both household and industrial consumers now expect goods to be delivered JIT. As such, the nature of demand for transactions across space has changed dramatically. As such, there is a direct parallel with the argument regarding information costs only, in this case, the opportunity (time) costs of goods shipments are tied up in the levels of inventory being held rather than the opportunity (time) costs of not having face-to-face contact.

There is a range of empirical evidence which suggests that the spatial transaction costs involved in the shipping of goods have indeed increased over the last two decades, because of this demand for more frequent deliveries. Firstly, the average inventory levels for almost all manufacturing and distribution sectors in the developed world have fallen dramatically since the 1980s, relative to the value of output (Shonberger, 1996; *Financial Times*, 1998). This implies that the average lead times of goods-shipments have fallen over recent years, with a concomitant increase in goods-shipment frequencies. Secondly, by carefully disentangling the various components of transport costs it becomes clear that the proportion of global output

which is accounted for by logistics and transportation activities in the economy has not fallen over recent decades (*Financial Times*, 1997; HUMMELS, 1999). Thirdly, while the transportation cost component of bulk materials has indeed generally fallen, in the case of manufactured goods, there is evidence that this proportion has actually increased over recent decades, in spite of the improvement in transportation and logistics technologies (HUMMELS, 1999). Fourthly, industries which are very dependent on JIT shipments have tended to reorganize their trade patterns in favour of geographically close suppliers and customers (REID, 1995; MCCANN, 1998). Moreover, this behaviour is even evident in industries in which the product value-weight ratios are extremely high (MCCANN and FINGLETON, 1996). In other words, such localization behaviour is present in the very industries which traditional Ricardian trade theories would have ruled out.

GLAESER, 1998, argues that taking a broad view of all the empirical evidence indicates that the aggregate share of total output accounted for by transportation costs has fallen markedly over time. However, it appears that most of the evidence points to falling geographical transactions costs for existing types of activities. The sectors in which spatial transactions costs have indeed fallen significantly over recent decades, are generally the sectors in which the nature of the spatial transactions undertaken have not changed fundamentally over time, in terms of the required frequency of interaction. This is typically the case in many raw material, agricultural or extraction industries, and in industries producing manufactured products at a mature stage within their product cycles (VERNON, 1966). This is also the case in service sector industries in which the nature of the information being transacted is rather standardized. On the other hand, in production sectors in which the demand lead-times have fallen dramatically, or in industries in which the variety and complexity of information generated has increased significantly, spatial transactions costs would appear not to have fallen over recent decades, and in some cases will actually have increased.

On the basis of these arguments it is clearly necessary for location theory to develop models which better incorporate issues such as time, financial costs, delivery speed and frequency into the location-production problems than do previous models. One approach is to attempt to explain the structure of spatial cost and revenue functions within an explicit frequency-optimization framework, without assuming that the shapes of these function *a priori*. If we adopt such an approach, it is possible to demonstrate that many of the convex (MCCANN, 1995b) and concave shapes (MCCANN, 2001b) which are familiar in location models can be generated without necessarily resorting to neoclassical assumptions of factor substitution. In the built environment, where putty-clay capital-stocks and labour hysteresis are major features, these observations have profound implications for the justifications we adopt for our production function specifications.

Within this general frequency-optimization framework, one direction of this research has focused on the role played by the total *logistics costs* faced by firms producing goods, rather than simply the total transport costs that they face (MCCANN, 1993, 1997, 1998). This logistics costs approach incorporates transport costs plus all of the industrial costs associated with inventory holding within the classical Weber-Moses type of framework, and allows for a whole series of new location-optimization conclusions to be generated. In particular, we see that the costs of geographical distance are far higher than has previously been supposed. This is because the inventory capital and space costs faced by firms are co-determined within the transport cost optimization problem (MCCANN, 2001b). An outcome of this is that firms which produce very high value-weight ratio products are actually *more* sensitive to locational issues than are firms producing low value-weight ratio products. This conclusion is completely opposite to the perceived wisdom which assumes that location is largely irrelevant for firms producing very high value-weight ratio goods. Moreover, the logistics-costs framework also shows that industrial clustering is often a natural outcome for firms producing high value goods, even in situations in which information spillages play no role whatsoever. This conclusion throws serious doubt on the simple transport cost-information cost dichotomy assumed by many high profile commentators (KRUGMAN, 1991; GLAESER, 1998), and provides a whole range of possible future research agendas.

Direction 3: *Introducing the environment into location theory.* A third key direction for future location research concerns the issue of the 'environment'. Here we are referring specifically to the nature of the local, or location-specific environment. The evaluation of environmental variations has been almost entirely absent within location theory models, yet within real estate and urban economics, environmental valuation is assumed to be a central feature underlying modern consumption and choice behaviour. Until now, the major analytical developments in this field have centred on the use of hedonic models (SHEPPARD, 1999) as a means of understanding how environmental preferences are included in the pricing and valuation of land. Moreover, adopting this approach also allows us to consider how such variations and preferences contribute to location choices. Yet, from the perspective of microeconomic location theory, it is necessary to consider how such environmental variations and pricing outcomes affect the location behaviour of firms. In order to do this, we must consider exactly what is meant by the term 'environment'; this could refer simply to the local natural environment, or the local built environment, the local business environment, or

various combinations of all three definitions. As yet, this is a problem which remains largely un-discussed by location theory. To the extent that these issues are discussed, they are generally subsumed within the overall clustering literature outlined above, and as we have seen, the problem with this is that by framing location problems in this clustering language tends to leaving such things largely un-discussed. However, the difficulties involved in integrating the nature of the local environment into location theory models must not be underestimated. The reason for this is that the issues concerning the hierarchical definition of the firm, and the intra- and inter-organizational features of the firms and their relations, which are largely ignored by the clustering literature, are themselves the issues that largely define the nature of the local business environment.

CONCLUSIONS

In this paper we have sought to outline the key analytical insights generated by industrial location theory and to point out the major methodological issues raised by this field of research. As we have seen, classical and neoclassical location models pose fundamental problems which are largely un-addressed by the modern clustering and agglomeration literatures, and yet which are essential in order to further our knowledge of spatial behaviour. In response to this, we have argued, firstly, that it is essential to develop a transactions-costs understanding of the internal and external technical and organizational relations of the firm. On the basis of this reference framework it then becomes possible to ask theoretical location questions which are empirically testable. Secondly, we must reconsider how spatial transactions costs are discussed and understood in the modern world. In particular, frequency-optimization models such as the logistics-costs model are required, which explicitly go beyond the adoption of a simple dichotomy between transport costs and information costs. Thirdly, it is essential for location models to properly integrate notions of the environment into the production-location frameworks.

NOTES

1. Classical location theory, irrespective of whether the analytical focus was on agricultural land use or manufacturing production, concerned itself primarily with the determination of what later became known as 'spatial margins of profitability' (RAWSTRON, 1958).
2. The exceptions here are the work of LOSCH, 1954, and CHRISTALLER, 1933. However, while the urban systems work of Losch was more advanced in terms of its treatment of microeconomic behaviour than that of Christaller, these 'ideal-landscape' approaches were quite different to the microeconomic approaches outlined here in that they were not explicitly considering the optimization behaviour of the individual firm.
3. New economic geography models ascribe a particular form of production function to the firm which is developed from a DIXIT and STIGLITZ, 1977, framework. The microeconomic limitations of these functions are discussed in NEARY, 2001.
4. Both the industrial clustering model of PORTER, 1990, 1998, and the 'new industrial areas' model of SCOTT, 1988, are therefore much less specific than the urban agglomeration about the particular spatial dimension which is critical in terms of information transactions costs. In cases where there are small-firm industrial structures, the spatial extent over which such trust relations operate will tend to be over small sub-national regional scales (SCOTT, 1988; PORTER, 1990). On the other hand, in industrial structures characterized by large vertically-integrated firms, such trust relations may operate over much larger regional spatial scales, and in the case of contiguous small-area nations, these regional scales may extend beyond the individual country boundaries (CASSON and MCCANN, 1999; ARITA and MCCANN, 2000; CANTWELL and IAMMARINO, 2000).
5. As with the case above, the (envelope) result turns out to be a non-linear square root function of all the cost variables (MCCANN. 1993, 1998, 2001b).

REFERENCES

ALONSO W. (1964) *Location and Land Use*. Harvard University Press, Cambridge, MA.

ARITA T. and MCCANN P. (2000) Industrial alliances and firm location behaviour: some evidence from the US semiconductor industry, *Appl. Econ.* **32**, 1,391–403.

BLAUG M. (1968) *Economic Theory in Retrospect*. Cambridge University Press, Cambridge.

CANTWELL J. A. and IAMMARINO S. (2000) Multinational corporations and the location of technological innovation in the UK regions, *Reg. Studies* **34**(4), 317–32.

CASSON M. and MCCANN P. (1999) Globalization, competition and the corporation: the UK experience, in WHITMAN M. (Ed) *The Evolving Corporation: Global Imperatives and National Responses*. Group of Thirty, New York.

CHRISTALLER W. (1933) *Die Zentralen Orte in Suddeutschland*, Fischer, Jena (translated by BASKIN C. W. (1966) *Central Placed in Southern Germany*, Prentice-Hall, Englewood-Cliffs, NJ).

D'ASPREMONT C., GABSZEWICZ J. J. and THISSE J. F. (1979) On Hotelling's stability in competition, *Econometrica* **47**(5), 1,145–50.

DIXIT A. K. and STIGLITZ J. E. (1977) Monopolistic competition and optimum product diversity, *Am. Econ. Rev.* **67**(3), 297–308.

Economist, The (1999a) The net imperative: a survey of business and the Internet, 26 June.

Economist, The (1999b) The world in your pocket: a survey of telecommunications, 9 October.
Ellison G. and Glaeser E. L. (1997) Geographic concentration in US manufacturing industries: a dartboard approach, *J. Pol. Econ.* **105**, 889–927.
Eswaran M., Kanemoto Y. and Ryan D. (1981) A dual approach to the locational decision of the firm, *J. Reg. Sci.* **21**(4), 469–89.
Evans A. W. (1973) *The Economics of Residential Location.* Macmillan, London.
Financial Times (1997) Survey: logistics, 7 October.
Financial Times (1998) Survey: supply chain logistics, 1 December.
Financial Times (1999) Survey: supply chain logistics, 17 June.
Fujita M., Krugman P. and Venables A. J. (1999) *The Spatial Economy: Cities, Regions and International Trade.* MIT Press, Cambridge, MA.
Gabaix X. (1999a) Zipf's law and the growth of cities, *Am. Econ. Rev.: Pap. & Proc.* **89**(2), 129–32.
Gabaix X. (1999b) Zipf's law for cities: an explanation, *Quart. J. Econ.* **114**(3), 739–67.
Gaspar J. and Glaeser E. L. (1998) Information technology and the future of cities, *J. Urban Econ.* **43**, 136–56.
Glaeser E. L. (1998) Are cities dying, *J. Econ. Perspectives* **12**(2), 139–60.
Granovetter M. (1973) The strength of weak ties, *Am. J. Sociol.* **78**, 1,360–89.
Gordon I. R. and McCann P. (2000) Industrial clusters, complexes, agglomeration and/or social networks?, *Urban Studies* **37**, 513–32.
Henderson J. V., Kuncoro A. and Turner M. (1995) Industrial development in cities, *J. Pol. Econ.* **103**, 1,067–85.
Hoover E. M. (1948) *The Location of Economic Activity.* McGraw-Hill, New York.
Hotelling H. (1929) Stability in competition, *Econ. J.* **39**, 41–57.
Hummels D. (1999) Have international transportation costs declined?, Department of Economics Working Paper, Purdue University.
Isard W. (1956) *Location and the Space Economy.* John Wiley, New York.
Isard W. and Kuenne R. E. (1953) The impact of steel upon the Greater New York-Philadelphia industrial region, *Rev. Econ. & Statist.* **35**, 289–301.
Krugman P. (1991) *Geography and Trade.* MIT Press, Cambridge, MA.
Launhardt W. (1885) *Mathematische Begrundung der Volkwirtschaftlehre.* Leipzig.
Losch A. (1954) *The Economics of Location.* Yale University Press, New Haven, CT.
Marshall A. (1920) *Principles of Economics*, 8th edition. Macmillan, London.
Martin R. and Sunley P. (2001) Deconstructing clusters: chaotic concept or policy panacea?, paper presented at the Regional Studies Association Conference, 21 November 2001, London.
McCann P. (1993) The logistics-costs location-production problem, *J. Reg. Sci.* **33**(4), 503–16.
McCann P. (1995a) Rethinking the economics of location and agglomeration, *Urban Studies* **32**(3), 563–77.
McCann P. (1995b) Journey and transactions frequency: an alternative explanation of rent-gradient convexity, *Urban Studies* **32**(9), 1,549–56.
McCann P. (1997) How deeply embedded is 'Silicon Glen'? A cautionary note, *Reg. Studies* **31**(7), 695–703.
McCann P. (1998) *The Economics of Industrial Location: A Logistics-Costs Approach.* Springer, Heidelberg.
McCann P. (1999) A note on the meaning of neo-classical location theory and its usefulness as a basis for applied research, *Pap. Reg. Sci.* **78**(3), 323–331.
McCann P. (2001a) *Urban and Regional Economics.* Oxford University Press, Oxford.
McCann P. (2001b) A proof of the relationship between optimal vehicle size, haulage length and the structure of distance-transport costs, *Transp. Research A* **35**, 671–93.
McCann P. and Fingleton B. (1996) The regional agglomeration impacts of just-in-time input linkages: evidence from the Scottish electronics industry, *Scot. J. Pol. Econ.* **43**(5), 493–518.
Miller S. M. and Jensen O. W. (1978) Location and the theory of production, *Reg. Sci. & Urban Econ.* **8**, 117–28.
Mills E. S. (1970) *Urban Economics.* Scott, Foresman, Glenview, IL.
Moses L. (1958) Location and the theory of production, *Quart. J. Econ.* **78**, 259–72.
Muth R. F. (1969) *Cities and Housing.* University of Chicago Press, Chicago.
Neary J. P. (2001) Of hype and hyperbolas: introducing the new economic geography, *J. Econ. Lit.* **39**, 536–61.
Nishiguchi T. (1994) *Strategic Industrial Sourcing: The Japanese Advantage.* Oxford University Press, Oxford.
Palander T. (1935) *Beitrage zur Standortstheorie.* Almqvist & Wiksells Boktryckeri, Uppsala.
Piore M. J. and Sabel C. F. (1984) *The Second Industrial Divide: Possibilities for Prosperity.* Basic Books, New York.
Porter M. (1990) *The Competitive Advantage of Nations.* Free Press, New York.
Porter M. (1998) Clusters and the new economics of competition, *Harv. Bus. Rev.* **76**(6), 77–90.
Rawstron E. M. (1958) The principles of industrial location, *Trans. Pap. Inst. Brit. Geogr.* **25**, 132–42.
Reid N. (1995) Just-in-time inventory control and the economic integration of Japanese-owned manufacturing plants with the county, state and national economies of the United States, *Reg. Studies* **29**(4), 345–55.
Schonberger R. J. (1996) *World Class Manufacturing: The Next Decade.* Free Press, New York.
Scott A. J. (1988) *New Industrial Spaces.* Pion, London.
Sheppard S. C. (1999) Hedonic analysis of housing markets, in Cheshire P. and Mills E. S. (Eds) *Handbook of Regional and Urban Economics III.* North-Holland, Amsterdam.
Simmie J. and Sennett J. (1999) Innovative clusters: global or local linkages?, *Nat. Inst. Econ. Rev.* **170**(October), 87–98.

VERNON R. (1966) International investment and international trade in the product cycle, *Quart. J. Econ.* **80**(2), 190–207.
WEBER A. (1909) *Uber den Standort der Industrien* (translated by FRIEDRICH C. J. (1929) *Alfred Weber's Theory of the Location of Industries*, University of Chicago Press, Chicago).
VON THUNEN J. H. (1826) *Der Isolierte Staat in Beziehung auf Landschaft und Nationalokomie*, Hamburg (translated by WARTENBERG C. (1966) *Von Thunen's Isolated State*, Pergamon Press, Oxford).
WILLIAMSON O. E. (1975) *Markets and Hierarchies*. Free Press, New York.