Methods of Interregional and Regional Analysis

WALTER ISARD
IWAN J. AZIS
MATTHEW P. DRENNAN
RONALD E. MILLER
SIDNEY SALTZMAN
ERIK THORBECKE

Ashgate

Aldershot • Brookfield USA • Singapore • Sydney

Published by
Ashgate Publishing Limited
Gower House
Croft Road
Aldershot
Hants GU11 3HR
England

Ashgate Publishing Company
Old Post Road
Brookfield
Vermont 05036
USA

British Library Cataloguing in Publication Data
Methods of interregional and regional analysis
1.Regionalism
I.Isard, Walter
304.2'3

Library of Congress Catalog Card Number: 98-071955

Printed and bound in Great Britain by
MPG Books Ltd, Bodmin, Cornwall

6. Gravity and spatial interaction models

Walter Isard

6.0 Introduction

A system of regions has an intricate structure. Only some of the major strands which interconnect people, households, firms, social groups, governmental agencies, and a variety of other operating and decision-making units have been isolated and subjected to analysis in the preceding chapters. The balanced scholar will therefore be uncomfortable, especially as he/she observes, perhaps at some distance in space, the geographic pattern in which human beings and their physical structures are massed in metropolises — metropolises which vary widely in size, configuration, and intensity of activity, such intensity tending to diminish in most all directions from the core. True, the fine-stranded generalized interdependence schemes as crystallized in interregional input-output and programming are powerful analytical tools. True, we have shown how urban complex analysis can be embodied in nonlinear programming and thereby capture some of the spatial juxtaposition (or agglomeration) economies and diseconomies. However, as yet, have they even in a small way, been able to cope with these economies? As already indicated, urban complex analysis has a very long way to go before it can fruitfully attack the complexities of metropolitan areas and the surrounding system of central places. An observer, impressed by the phenomena of human massing within any system of industrialized regions and the intricate spatial structures they possess, might ask: Is not society more than a matrix of finely detailed connections among units? Is not the structure of a system of regions more than the sum of the interactions of sets and patterns of units or sectors as conceived by interregional input-output and programming? Are there not other over-all forces pertaining to masses which pervade society and confine the multitude of possible interactions among its innumerable units? These questions motivate the analyst to explore fresh approaches, even quite different ones. One such approach is that of the gravity model and the diverse spatial interaction frameworks that have evolved or been associated with it.[1]

In the gravity, potential, and spatial interaction models — which for short we shall term gravity models whenever we speak generally of these models — the region is conceived as a mass. The mass is structured according to certain principles. These principles govern in an over-all fashion the range of behavior of the individual particles, both constraining and initiating their action. Interregional relations may be thought of as interactions among masses. Again, general principles may be said to govern the frequency and intensity of such interactions; and by so doing they influence the behavior of individual units (particles) within each mass. This approach is a macro one that resembles an approach frequently used by physical scientists. For example, Boyle's classic studies of the effects of pressure and temperature on the volume of gases were essentially investigations into the behavior of masses of molecules; the movement of any individual molecule was not a matter of inquiry.

6.1 A simple probability point of view

To develop the basic concepts underlying gravity models it is useful to start with a rather *simple and loose* probability point of view.[2] (A rigorous probability statement will be discussed in section 6.4.) Suppose there is a metropolitan region with population P. Let the region be divided into many subareas. Let there also be known the total number of internal trips taken by the inhabitants of this metropolitan region. We represent this number by the constant T. Further, let there be no significant differences among subareas in the tastes, incomes, age distributions, occupational structures, etc. of their populations and individuals within their populations.

Now, suppose we wish to determine the number of trips which originate in, let us say, subarea i, and terminate in, let us say, subarea j. Assume, for the moment, that no costs and no time are involved in undertaking a trip from one area to another, that is, that the friction of distance is zero. For this hypothetical situation we may expect that for a representative individual in subarea i the per cent of his journeys terminating in subarea j will be equal to the ratio P_j/P, the population of subarea j divided by the total population of the region, *ceteris paribus* — assuming that interaction opportunities are proportional to population size. That is, if the total population of the metropolitan region is 1,000,000 and that of subarea j 100,000, we may expect the individual to make 10 per cent of his/her trips to j. Additionally, since a representative individual in subarea i is by our homogeneity assumptions identical with a representative individual in any other subarea, and since his/her transport time cost is zero, we may estimate the number of trips he/she undertakes as the average number of trips per capita for the entire metropolitan region. This average is equal to T/P. Designating this average by the letter k, we find that the absolute number of trips which a representative individual in subarea i makes to subarea j is $k(P_j/P)$. That is, if 10 percent of the total population resides in subarea j, the individual in subarea i will tend to make 10 percent of his trips to subarea j; if the average number of trips per individual is 20, the individual will make, on the average, two trips to j.

This reasoning applies to one representative individual in i. But there are P_i individuals residing in subarea i. Therefore, the number of trips to subarea j which these P_i individuals will make will be P_i times the number of trips to subarea j which the representative individual in i makes. That is,

$$T_{ij} = k(P_i P_j / P) \tag{6-1}$$

where T_{ij} designates the total number of trips taken by individuals in i (i.e., originating in i) which terminate in j. In this manner we can estimate the expected total number of trips for every possible combination of originating subarea and terminating subarea. Thus we obtain for the metropolitan region a set of expected or hypothetical trip volumes (total number of trips) between subareas.

Our next step is to determine the possible effect of the actual distance separating a pair of subareas on the number of trips occurring between them. First, for a typical metropolitan region we obtain actual data on the number of trips between every pair of its subareas. We let I_{ij} represent the actual trip volume between any originating subarea i and any terminal subarea j. We divide this actual number by the expected or hypothetical trip volume T_{ij} to derive the ratio of actual to expected trip volume, that is, I_{ij}/T_{ij}. We also note the distance d_{ij} which separates i and j. Finally, we plot on a graph with a logarithmic scale along each axis both the ratio I_{ij}/T_{ij} and distance d_{ij} for this particular pair of subareas. For example, in Figure 6.1 where the vertical axis measures the ratio of actual to expected trips and where the horizontal axis measures distance, we may note point L. Point L refers to a pair of subareas approximately 3.6 miles apart for which the ratio of actual to expected trips is approximately 0.4.

In similar manner, for every other combination of originating subarea and terminating subarea we plot the set of data on the ratio of actual to expected trips and intervening distance. Suppose our data are as indicated in Figure 6.1. They suggest a simple relationship between the log of the

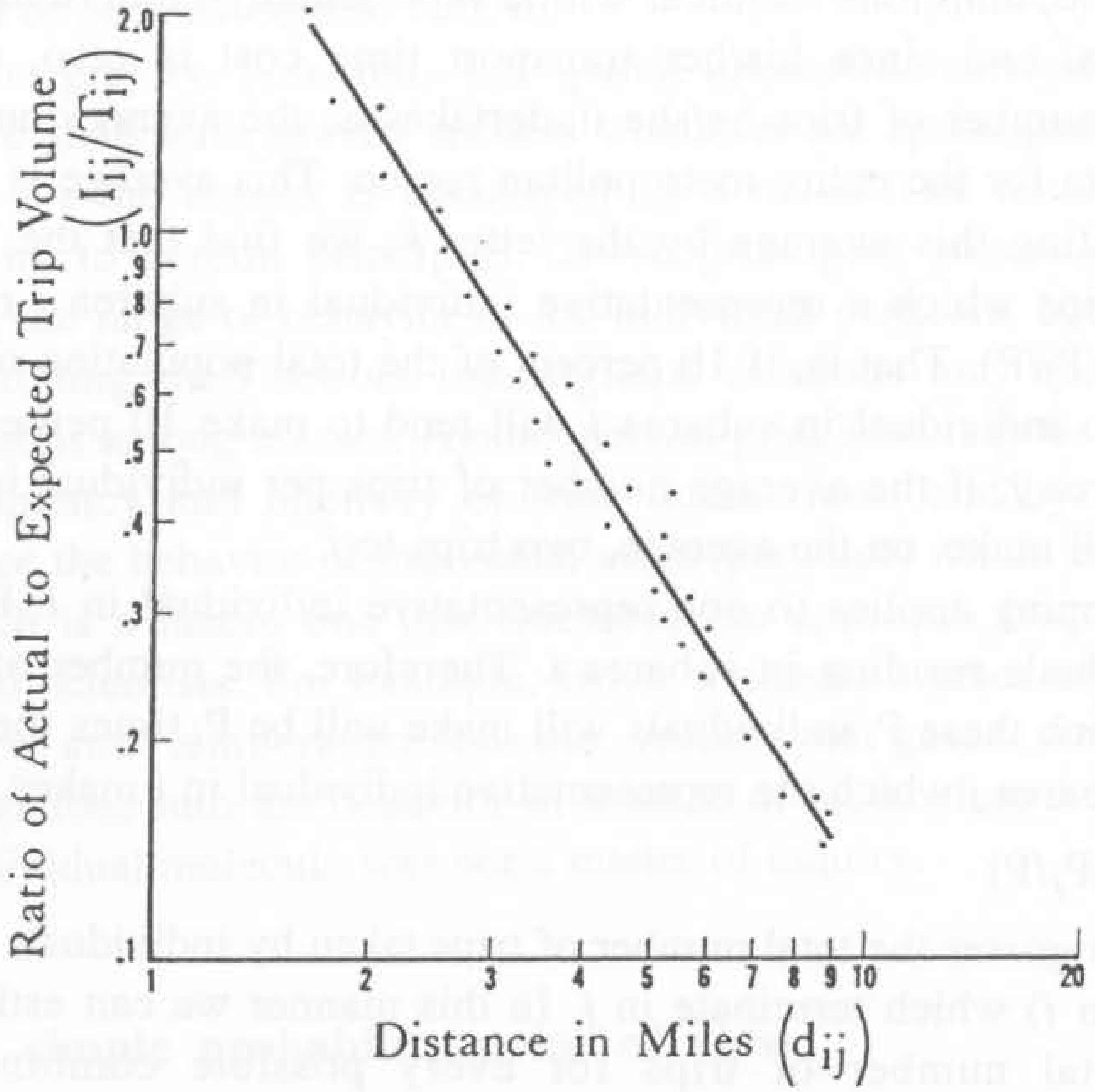

Figure 6.1 Relation between distance and the ratio of actual to expected person trips (hypothetical data)

ratio of actual to expected trip volume on the one hand and log of distance on the other hand. A straight line may be fitted to the plotted data by least squares or by other methods. Since our variables are the log of the ratio of actual to expected trip volume (the dependent variable) and the log of distance (the independent variable), the equation of the line is

$$\ln(I_{ij}/T_{ij}) = a - b \ln d_{ij} \tag{6-2}$$

In this equation a is a constant which is the intercept of the straight line with the Y axis, and b is a constant defined by the slope of the line.[3] Removing logs from equation (6-3) and letting c equal the antilog of a, we have

$$I_{ij}/T_{ij} = c/d_{ij}^{b}$$

or

$$I_{ij} = cT_{ij}/d_{ij}^b \tag{6-3}$$

Substituting in equation (6-3) the value of T_{ij} as given in equation (6-1), and letting the constant $G = ck/P$, where c, k, and P are constants as defined earlier, we obtain

$$I_{ij} = G(P_iP_j/d_{ij}^b) \tag{6-4}$$

This simple relationship may then be taken to describe roughly the actual pattern of trip volumes within the metropolitan region, *ceteris paribus*. That is, it depicts the interaction of people within the metropolitan region as a function of the populations of subareas and the distance variable when this interaction is reflected in trips.

Suppose we study the relationship of actual to expected magnitude on the one hand, and distance on the other hand, for a number of other phenomena reflecting the interactions of people within the metropolitan mass and *among* metropolitan masses. We might examine telephone calls, telegraph messages, express shipments, money flows, migration, commuting and shopping patterns, etc. Suppose that for all these phenomena we find, as in Figure 6.1, a close linear association between the log of the ratio of actual to expected volume and the log of distance.[4] We might then conclude that the relationship in equation (6-4) reflects a basic principle underlying the structure of metropolitan areas and systems of metropolitan areas — namely that, all else being equal, the interaction between any two populations can be expected to be related directly to their size and inversely to distance. This relationship derived from a probability point of view is essentially the gravity model of physics where P_i and P_j stand for masses M_i and M_j, the exponent b takes the value of 2, and I_{ij} represents the gravitational force F.

Additionally, it is to be noted that equation (6-4) can be converted into another useful form. Suppose we are interested in the interaction between a single subarea i and all other subareas. We would therefore derive the interaction of i with the first subarea (i.e., I_{i1}) plus the interaction of i with the second subarea (i.e., I_{i2}) plus the interaction of i with the third subarea (i.e., I_{i3}) plus . . ., and finally plus the interaction of i with the nth subarea (i.e., I_{in}). From equation (6-4) we find values for each of the interactions, $I_{i1}, I_{i2}, I_{i3}, \ldots, I_{in}$. By addition we obtain

$$I_{i1} + I_{i2} + I_{i3} + \ldots + I_{in} = G(P_iP_1/d_{i1}^b) + G(P_iP_2/d_{i2}^b) + G(P_iP_3/d_{i3}^b) + \ldots + G(P_iP_n/d_{in}^b) \tag{6-5}$$

or

$$\sum_{j=1}^{n} I_{ij} = G(\sum_{j=1}^{n} P_i P_j / d_{ij}^b) \tag{6-6}$$

Since P_i may be factored from the right-hand side of the equation (6-3), we derive, after dividing both sides by P_i

$$\sum_{j=1}^{n} I_{ij}/P_i = G(\sum_{j=1}^{n} P_j / d_{ij}^b) \tag{6-7}$$

Note that the numerator of the left-hand side of equation (6-7) is the total interaction of i with all areas including itself,[5] which when divided by the population of i, namely P_i yields interaction with all areas *on a per capita basis* or more strictly *on a per unit of mass basis*. Interaction on such a basis has been designated *potential* at i, for which we employ the symbol ${}_iV$. By definition, then,

$${}_iV = \sum_{j=1}^{n} I_{ij}/P_i \tag{6-8}$$

and, from equation (6-7) we have

$${}_iV = G(\sum_{j=1}^{n} P_j / d_{ij}^b) \tag{6-9}$$

It is to be noted that this equation can be derived directly from a simple probability approach as was equation (6-4). Equation (6-9) is the basis of *potential* models and as developed is a variation of equation (6-4), the basic gravity model. It is effectively employed in spatial choice models, for example, in the choice of that one region among many to which an individual might migrate (or firm relocate) or of that shopping center or medical complex an individual might patronize.

Extensive literature exists on the history and development of the use of gravity models and related concepts by social scientists. (See Isard et al. (1960), Fotheringham and O'Kelly (1989), Sen and Smith (1995) and others cited therein.) For purposes of this chapter, we start with the most advanced framework reached in year 1960. At that time much thought was given to the possibility that weights, w_i and w_j, should be applied to the respective masses M_i and M_j. For example, should population be weighted by its per capita income to help explain the level of airline trips between the populations (masses) of metropolitan regions i and j? Moreover, because of different agglomeration (deglomeration) economies and externalities that might be associated with the masses, should exponents α and β be applied to the masses M_i and M_j, respectively.

With regard to distance, should physical, economic (transport cost), or travel-time distance be employed? Furthermore, should the exponent b (the distance sensitivity factor) be 2 (as in gravitational force) or unity (as in the mutual energy concept of physics) or some other, particularly to reflect the different intensities of the 'falling-off' effects that are observed with different kinds of trips (for example, to grammar schools in contrast to higher educational institutions).

Thus there evolved the more general formulation (where P_i and P_j of equation (6-4) are replaced, respectively, with M_i and M_j):

$$I_{ij} = G[w_i(M_i^{\alpha})w_j(M_j^{\beta})]/d_{ij}^{b} \qquad (6\text{-}10)$$

where at times G_{ij} was suggested as a replacement for G to reflect complementarity of resources and other attributes of M_i and M_j and where at times other functional forms were employed in which distance (d) entered into the denominator in a different way than simply as a variable raised to some power.

Since 1960, a very large array of gravity and spatial interaction models has evolved. They relate to many different kinds of interactions, some purely theoretical but much more frequently interactions concerned with everyday problems and planning for which useful applications have been sought. Among others, these kinds of interactions have pertained to railway and airline trips; mail and telephone calls; large volume commodity movements; railway express shipments between urban areas; journey-to-work trips; trips to shopping centers, museums, libraries, recreational areas, hospitals, schools, universities and other cultural and educational sites; trade among nations; migration among regions; and in other fields of study such as magnitude of church attendance; social visits and marriages among neighborhood populations; and extent of gang warfare, conflicts among nations and competition among teams. These models are used to understand current and past interactions. However, they have also been fruitfully employed in many studies to forecast (project) the impacts of various changes in variables and policies that affect these variables — such as the impacts upon traffic of a new industrial area, a major residential development, a major shopping center construction, a new superhighway, the reorganization of the transportation system, and so forth. But such projection has been of a comparative statics type in which in most studies there is a shift from one stationary or equilibrium state to another, and wherein no significant structural change has taken place. Even in some studies, where state transitions have been permitted, the spatial interaction patterns that have been derived are stationary or equilibrium patterns, and do not involve a dynamical process that is in general pertinent to a gravity model.

Because there are so many types of applications and theoretical analyses that are capable of being embraced, or said to be embraced, by gravity and spatial interaction models, one cannot establish any particular models best for general use. Hence, in what follows, we shall discuss: (1) the various ways each variable can be defined and measured; (2) alternative distance measures and spatial separation and related functions that can be used; and (3) alternative hypotheses and theories that can be set forth as background for analysis. We leave to the researcher the problem of designing a specific model for the particular situation he/she is interested in, although we will have occasion to present a limited number of interesting formulations and applications of particular frameworks.

6.2 Definition and measurement of mass

What is a relevant mass? As in the above example, it can be population in an area taking trips to destinations whose masses are other population. But the mass of a destination area can be: number of jobs; square footage of retail space in a shopping center adjusted for the quality and diversity of its stores; number of hospital beds; amount of marina facilities; size of university; size of a region's economy (Gross Regional Product); its labor force; its income level; its total wholesale and retail sales; its level of consumption; its economic opportunities in general; its value added in manufacture; its investment in infrastructure; its newspaper circulation; its car registrations and a host of other of its magnitudes.

In the literature, population is often associated with a set of actors at one or more areas (points) of origination who are behaving in a conscious manner. In making decisions to undertake (produce) movement (travel, trips, migration), they are reacting to or being propelled to exploit opportunities, attractive attributes, and/or drawing forces elsewhere, in particular destination areas. However, these opportunities, attributes, and forces can be many in real life. So can the attributes of actors or elements of the originating mass which generate interaction. Hence, there can be innumerable pairs of a combination of actor attributes and a combination of destination attributes. But taking into account the specific attributes in a pairing in the modelling of the interaction within that pairing would diminish the effect of spatial separation and assign much more weight to the particularities of the attributes of the originating and destination elements and their complementarities. This would run counter to what is conceived in physics as gravitational interaction, namely the spatial interaction of the aggregate activity of the innumerable molecules comprising each mass. And *in general* it has been found in regional science and other social studies that the gravity model is most applicable when it pertains to the spatial interaction of large masses (aggregates), less applicable to that of their subaggregates, still less to that of the parts of these subaggregates, and even highly questionable to that of small groups and individuals. In brief, gravitational interactions are found to be much more pertinent on the *macro* level of analysis, where the effect of any specific attribute in the many diverse pairings of attributes averages out, than at the micro where there is no averaging.

To be specific, consider migration. Among others, significant attributes of destination areas may be number of unfilled jobs, wage levels, intensity of the drug problem, the crime rate, the level of interracial conflict, congestion, pollution, presence of cultural facilities, climate, probability of earthquakes, and other natural disasters. Significant attributes of originating areas may be unemployment rates, quality of the educational system, lack of social welfare programs, the conservativeness of a political regime, information level about opportunities, and a host of others including many, if not all, of those noted for destination areas. For a given individual or subclass of migrants some of these attributes may be pertinent in the decision to migrate to a new location and much more important than mere spatial separation. For others, different sets of attributes may dominate. However when the total of the migrations of all subclasses and individuals of a large national population is investigated, it will often turn out that spatial separation is one of the leading, if not dominant factor, that accounts for the aggregate pattern of migration. Obviously the extent to which there exists homogeneity of migrants and of destination areas, the more clearcut will be the spatial separation effect, and the smaller a sample of migrations will be required to establish the particular nature of the spatial separation effect.

Nonetheless, as we shall note below, modifications of the simple gravity model of equation (6-10) and its simpler forms have been found to be useful in particular situations where these modifications combine the gravitational interaction with the particular attribute complementarities of less aggregated masses — where subaggregates are designed (stratified) as much as possible to minimize the within-class variance with regard to each subaggregate's relevant elements.

In brief, all this is saying in another way that the gravitational interaction is more clearly discerned the fewer the differences among the elements of each mass and the larger the masses involved.

Up to now we have treated aggregates of behaving units (actors) as constituting the masses designated as origins. However, behaving units can comprise the masses of destinations, and elements like jobs and qualities of an urban region as origin attributes can set in motion interaction. For example, the impact of the development of a large industrial complex in a new district may be examined in terms of the area in which its labor force may come to take up residence. In brief, what is designated as origins and destinations depends on the objectives and other aspects of a particular study. Too, destinations and origins can often be viewed as interchangeable.[6]

Moreover, in some interaction, for example marriages among separated populations, there may exist mutual (bilateral) attractions (or opportunities) like mutual energy in physics. Similarly, in social correspondence, contract agreements and communications among firms, discussions among nations to reduce conflict, and so forth.

One way to handle the problem of heterogeneity among masses may be to employ weights. For example, take the effect of spatial separation upon first-class airline traffic among urban areas. It is reasonable to expect that, *ceteris paribus,* an area with high per capita income will generate a larger volume of such travel than an area of equal population but lower per capita income. One way to handle this particular heterogeneity is to multiply the population of each generating subarea i by its average per capita income — that is by applying the weight w_i noted in equation (6-10). Weights, w_j, might also be applied to destination areas to recognize the different quality of say recreational facilities when such air traffic is recreationally oriented.

Additionally, the investigator may wish to employ more than one type of weight to adjust for heterogeneity within and between masses, perhaps to take into account differences among behaving units in educational level, age-sex composition and other factors. Then each weight w_i might be a composite weight constructed and applied in an appropriate manner.

Processes internal to masses of an agglomerative, cross-catalytic nature, which affect the forces emanating from them, may be considered relevant, and may be captured in the exponents α and β in equation (6-10). In one sense when we set $\alpha = \beta = 1$ we assume a zero net effect on that basis. The α and β become parameters to be estimated in the log linear statement of equation (6-10) and in a sense provoke interpretation when in calibrations (in statistical estimation) they turn out to be other than unity.

To sum up, within each mass, M_i, $i = 1,...,n$, whether an origin or destination, many elements may be incorporated. From here on we shall designate the mass of an area as O_i when it plays the role of an originating area and as D_j when it plays the role of a destination area. O_i and D_j may be scalars, as when they represent the number of individuals, P_i and P_j, respectively, or when they constitute weight functions modified by agglomerative (cross-catalytic) effects of attributes captured in the α and β exponents of equation (6-10).

6.3 Definitions and measures of distance

In the literature, distance has on a number of occasions been defined physically along a straight line connecting two masses in terms of miles or other standard unit. However, if a metropolitan traffic study is being conducted, distance measured in terms of travel time may be considered more appropriate, or at times some combination of miles and travel time when both peak and off-peak travel are to be considered. In other studies, other measures of distance may be employed, for example: economic distance as measured by transport or travel cost, or the number of links in a transport route or communication channel. In these cases the symbol d_{ij} for distance is often replaced with the symbol c_{ij} to represent an economic (cost) distance. Also, when good estimates of social distance, political distance, ideological distance, psychological distance, or other distance (perhaps a cognitive-type) exist, they may be employed. Usually, however, such distances are inferred and estimated when for a given situation the number of interactions and size of the masses involved are known.[7]

6.4 Functional forms for spatial separation

While deterministic theories of the gravity model have been set forth, none have found wide acceptance and of practical significance. Theories that would view a mass as a composite of micro units each maximizing utility subject to a budget constraint to yield for each unit a demand function for 'interaction with spatial opportunities,' or which would associate disutility with transport and travel costs (travel time), have confronted difficulties. Proponents have not been able to conduct effective testing of relevant hypotheses. And while subjective notions such as spatial discounting may have some appeal (see Isard, 1975), rigorous probabilistic theories to explain spatial interaction phenomena noted above have found substantial acceptance; their hypotheses have been extensively tested and widely applied.

Among probabilistic theories, two forms, each with different desirable properties, have found extensive use. Each generally takes weights w_i and w_j, and exponents α and β as statistical parameters to be estimated. They differ with respect to the deterrence, attenuation or falling-off effect of distance. One views the deterrence function $F(d_{ij})$ as $(d_{ij})^{-b}$, namely as a *power* deterrence function, as in equation (6-4). The other views it as $\exp(-bd_{ij})$, namely as an *exponential* deterrence function, where b is a positive distance sensitivity parameter. The latter view yields the simple model form

$$I_{ij} = M_i M_j \exp(-bd_{ij}) \tag{6-11}$$

The power deterrence function is the one that derives directly from and takes the same form as the gravity model of physics. It has a framework that is suited for application across many studies, especially those involving forecasts. Specifically, it possesses a homogeneity property where estimated parameters are independent of the scale of a system and the units in which distance (cost) is measured. For example, if in a journey to work study, the number of opportunities at destination areas were to be doubled, then relative attraction would remain the same, that is for any two destinations, $j = r,s$, from equation (6-4) where M replaces P, we have

$$\frac{(2M_r)^\beta}{(2M_s)^\beta} = \frac{2^\beta M_r^\beta}{2^\beta M_s^\beta} = \frac{M_r^\beta}{M_s^\beta} \tag{6-12}$$

Similarly, if distance were to be measured in kilometers instead of miles. According to Fotheringham and O'Kelly (1989), this property, in contrast, is not possessed by an exponential deterrence model. As they note in discussing the situations for which a power or exponential function is most suited, 'a model with an exponential cost function calibrated with traffic flows from a major city could not be used to forecast traffic flows in a medium or small urban area' (p. 11).

However, this invariance property of power deterrence functions under similarity transformations is very questionable for very small distances or cost values, that is when d_{ij} or $c_{ij} \to 0$. Under these circumstances, the predicted (expected) spatial interaction would take on exceedingly large values — values that are not observed. However, this problem is easily exaggerated, being more theoretical than actual. For in a real situation, no two individuals can occupy the same space, and, in a significantly large aggregation d_{ij} or c_{ij} cannot approach zero. (After all black holes are not present in society.) A non-negligible average distance (and cost of spatial interaction) can be claimed to separate them. Alternatively, one can posit the existence of some ε to reflect a start-up, information gathering, or some terminal cost, especially given the fact that individuals are not truly homogeneous. Accordingly, d_{ij} would need to be replaced by

$$(\varepsilon + d_{ij})^{-b} \text{ or } [\varepsilon + d_{ij}{}^{b}]^{-1}$$

or some other expression. In contrast, this problem of overestimating low cost, short distance movements of an unmodified power deterrence function does not exist with an exponential deterrence function.

Another issue in evaluating the relative desirability of these two types of deterrence functions concerns treatment of expected cost increases. As discussed by Fotheringham and O'Kelly (1989, p. 11),

> Suppose in an analysis of passenger flows on public transit within a major city, costs are to be increased along certain routes. Two possible consequences of this action can result: the selective increase in costs will alter the whole trip matrix; or the trip matrix will remain stable. Also, two types of cost increase are possible; each fare can be increased by a constant multiple or a constant amount. This produces four scenarios under each of which one of the two spatial separation functions is appropriate and the other is inappropriate. For instance, if a multiplicative cost increase is to be applied and this is expected to alter the trip matrix, an exponential cost function should be employed. Conversely, if the cost increase is to be additive, a power function is more appropriate.[8]

The exponential deterrence function is not one that derives from the gravity model of physics. It developed from the early pioneering work of Wilson, 1970, 1974, who initially linked it to entropy analysis and statistical mechanics. This function also derives from the rigorous, independent work of Sen and Smith (1985) — namely a behavioral interpretation when probabilistic variations in interactions at the individual (micro) level are assumed to depend only on average interaction costs and activity levels and when individual interactions are assumed to be statistically independent.

Other issues exist. Fotheringham and O'Kelly conclude based on the kinds of interaction models they have studied and those with which they are familiar, there exists 'a reasonably widespread consensus that the exponential function is more appropriate for analyzing short distance interactions such as those that take place within an urban area. The power function, conversely, is generally held to be more appropriate for analyzing longer distance interaction such as migration flows' (pp. 12–13).[9]

Before concluding this section, we should make explicit the independence assumptions when an advanced formal behavioral interpretation of the gravity model is adopted, as in Sen and Smith, 1995. As they emphasize, each flow T_{ij} between every origin i and every destination j is to be considered a random variable. Thus these random variables can be converted into a set of random variables which are the estimators or estimates. These estimates, augmented by additional variables, yield other random variables which are the forecasts. Hence, the resulting mean (average) frequencies of spatial interactions of behaving units with respect to various origin-destination pairings imply an assumption of *locational independence.* That is, within any given interaction pattern involving a significant number of interactions the likelihood of any given interaction is assumed to be uninfluenced by the properties of the other realized interactions. It is posited for example that there are no congestion effects, such as would be realized and affect that likelihood if more individuals would want to shop at a store with insufficient capacity to serve them. It also is posited that there are no contagion or bandwagon effects, as when friends, shopping together, leads to identical interaction choices. A second assumption implied is that of *frequency independence.* That is, the realized value of each interaction frequency is assumed to be unaffected by the realized value of any other interaction frequency.

Practical application of probabilistic interpreted models must relax these independence assumptions. Their use requires that the masses at i and j are sufficiently large so as to minimize (1) the influence of an interaction associated with any individual behaving unit (or opportunity) on the likelihood of any other unit's interaction (or upon the effect of any other opportunity) and (2) the many types of frequency dependencies among the interaction types existing at the micro level. (See Sen and Smith, 1995 for further discussion.)

It should be noted that other advanced research with gravity models allows for the possibility of multiple measures of separation. If these models are power deterrence ones, the interaction would be a function of a set of positive cost (or distance cost) profiles, each raised to its own power. If these models are exponential deterrence ones, the logarithmic form of them would be linear combinations of the several separation cost measures, each with its own cost sensitivity parameter.

Another direction for advanced research concerns *threshold* models in which there is a genuine possibility that any given interaction (a migration, shopping trip, social visit) considered by an individual will not in fact be taken. Basically, individual behavior is postulated to involve in these models an implicit two-stage process in which a variety of potential interaction situations arise and are either acted upon or not, depending on the individual's current attitudes toward spatial interaction. In these models it is hypothesized that a given interaction will occur if and only if the anticipated travel costs (time, effort, stress) do not exceed the individual's current tolerance levels, designated as his interaction *threshold levels.*

Still another set of advanced gravity models motivated by the need (first forcefully stressed by Stouffer, 1940) to recognize the effect of 'intervening opportunities' on spatial interaction between locations. This has led to significant research on search processes and spatial choice behavior. In these models, each of a set of actors originating at one of a set of locations and attracted by opportunities distributed among a set of destinations, searches among these opportunities until (when there is no stopping rule) he/she identifies one meeting his/her needs (or none at all). These models need to specify for each origin a *search scheme*, namely the order in which spatially identified opportunities are evaluated by each of the relatively homogeneous actors at that origin. The search is concluded once a satisfactory opportunity is identified, all opportunities at a given destination being explored before a new destination is considered. (See Sen and Smith, 1995, for additional discussion.) Further research in the broad subject of spatial choice behavior, as it reflects spatial interaction, is presented in Fotheringham and O'Kelly (1989).

Gravity-type models can also be used to treat (1) interaction behavior of an hierarchical nature when first a relevant opportunity cluster is identified by a behaving unit and then a specific destination within the cluster is chosen,[10] (2) cases where prominence of a particular attribute, as perceived by actors, exists, (3) situations where actors possess limited information, and so forth.

6.5 Constrained gravity (spatial interaction) models

As already noted, the early development of exponential deterrence models was stimulated by the pioneering work of Wilson (1970, 1974) and his associates (see Fotheringham and O'Kelly, 1989). These models have often been designated maximum entropy and/or information-minimization models. The presentation of the basis for these designations is beyond the scope of this chapter as is their behavioralistic axiomatic formulation by

Sen and Smith (1995). From these models there has emerged a number of practical applications, particularly for situations where constraints on interaction exist or are provided exogenously. Where the constraints in a model specify the number of flows emanating from (or outflows produced at) each originating mass O_i (as defined above) (for example, the journey to work trips of residents at each i), the model has been designated *production-constrained.* Where these constraints specify the number of flows terminating at (or inflows attracted to) each destination mass D_j (for example, the number of jobs at each industrial area j), then the model has been termed *attraction-constrained.*

Where both types of constraints are specified the model has been termed *doubly-constrained* — in contrast to the two types of models just noted, which are often termed *singly-constrained.* At this point it is instructive to see how constraints affect the outcomes of a model. To do so, we use the example of a doubly-constrained model developed by Masser (1972) and reported upon in Haynes and Fotheringham (1984, pp. 24-29).[11]

Given O_i and D_j ($i,j = 1,2,3$), the task is to derive the T_{ij} for the cells in the O-D matrix below.

	D_1	D_2	D_3
O_1	T_{11}	T_{12}	T_{13}
O_2	T_{21}	T_{22}	T_{23}
O_3	T_{31}	T_{32}	T_{33}

Let the model we use be a most simple one, namely

$$T_{ij} = G(O_i D_j / d_{ij}) \tag{6-13}$$

where T_{ij} represents journey-to-work trips from origin i to destination j and where we set α, β, b and all weights as unity. What is first desired is an initial estimate of the outflows (residents taking journey-to-work trips) from the three origins, O_1, O_2, and O_3 whose number of residents are, respectively, 160, 450, and 180. This number for each origin is indicated in the total outflows column (col. 4) of Table 6-1 below.

The destinations are three, D_1, D_2, and D_3, whose inflows (workers to perform the jobs to be done at each) are 200, 370 and 220, respectively. The total inflow at each destination is indicated in the total inflows row at the bottom of Table 6-1. Let the distances (in miles) between these origins and destinations be as indicated in Table 6-2.

Table 6-1 Origin and destination masses and first round estimates of interaction (trips) (figures rounded)

	D_1 (1)	D_2 (2)	D_3 (3)	Total Out-flows (4)	Total Estimated Outflows (5)	Ratio: col. 4/col. 5 (6)
O_1	16,000	3,947	7,040	160	26,987	$0.005929 = \bar{A}_1$
O_2	6,000	83,250	9,900	450	99,150	$0.004539 = \bar{A}_2$
O_3	7,200	6,660	19,800	180	33,660	$0.005348 = \bar{A}_3$
Total Inflows	200	370	220			

Table 6-2 Distances between origins and destinations (in miles or other appropriate spatial separation measure)

	D_1	D_2	D_3
O_1	2	15	5
O_2	15	2	10
O_3	5	10	2

What needs to be computed are G, a gravitational-type constant, and the T_{ij}, $i = 1,2,3$ and $j=1,2,3$. The steps are:

(1) Calculate as follows a first extremely crude estimate of each outflow from each origin to each destination, to be designated $\bar{T}_{ij}$ using equation (6-13). For example: for $\bar{T}_{11}$ we multiply the total outflows for O_1 by the total inflow of D_1 and divide by 2 (namely d_{11} of Table 6-2) to obtain 16,000; for $\bar{T}_{12}$ we multiply the total outflow for O_1 by the total inflow for D_2 and divide by 15 to obtain 3,947; and so forth for each of the other seven $\bar{T}_{ij}$ to be calculated for the cells of Table 6-1.

(2) For each row, sum the items in it to obtain a first crude estimate of total outflows, designated $\hat{T}_i = \sum_j \hat{T}_{ij}$. For example, the total for the first row is 26,987.

(3) For each origin, calculate the ratio of the given total outflow constraint (160 for O_1) to the total of the first round estimate of outflows (26,987 for O_1). This ratio for O_1 is 0.005929 as given in column 6 of Table 6-1 and is designated $\bar{A}_1$. The corresponding ratios for the other two rows are 0.004539 and 0.005348. They are designated $\bar{A}_2$ and $\bar{A}_3$, respectively.

(4) Multiply the elements in each row by the corresponding ratio in column 6 to obtain the first round of rowwise adjusted trips as recorded in Table 6-3.

Table 6-3 Adjusted first round estimates of outflows and resulting first round estimates of destination inflows (figures rounded)

		D_1	D_2	D_3	Total Adjusted First Round Outflow Estimates
(1)	O_1	94.86	23.40	41.74	160.00
(2)	O_2	27.23	377.84	44.93	450.00
(3)	O_3	38.50	35.62	105.88	180.00
(4)	Total Estimated Inflow	160.60	436.85	192.55	
(5)	Total Inflows Constraints	200	370	220	
(6)	Ratio of row (5) to row (4)	1.2453 $= \bar{B}_1$	0.8470 $= \bar{B}_2$	1.1426 $= \bar{B}_3$	

(5) Sum the items in each column to obtain for each destination the first-round estimated total of inflows from all origins as recorded in row 4 of Table 6-3.

(6) Record the total inflow constraint for each destination in row 5 of Table 6-3.

(7) For each destination take the ratio of total inflow constraint (row 5) to total estimated inflow (row 4) and record the ratio in row 6. These ratios are designated $\bar{B}_1$, $\bar{B}_2$ and $\bar{B}_3$.

(8) Multiply the elements in each column by the ratio for that column in row 6 to obtain second round estimates of inflows $\bar{\bar{T}}_{ij}$ as recorded in Table 6-4.

(9) Sum the items across each row in Table 6-4 to obtain the total adjusted second-round outflow estimates for each origin, as recorded in column 4 of Table 6-4.

Table 6-4 Second round estimates of inflows and outflows (figures rounded)

	D_1 (1)	D_2 (2)	D_3 (3)	Total Adjusted Second Round Outflow Estimates (4)	Constraint on Total Outflows (5)	Ratio of col. 5 to col. 4 (6)
O_1	118.14	19.82	47.69	185.65	160	0.8618 $= \bar{\bar{A}}_1$
O_2	33.91	320.02	51.34	405.27	450	1.1104 $= \bar{\bar{A}}_2$
O_3	47.95	30.16	120.97	199.08	180	0.9042 $= \bar{\bar{A}}_3$
Total Estimated Inflows	200.00	370.00	220.00			
Constraint on Total Inflows	200.00	370.00	220.00			

(10) List the constraints on total outflows in column 5 of Table 6-4. If for any origin the total adjusted second round outflows estimate differs from its respective constraint on total outflows, take the ratio of the latter (col. 5) to the former (col. 4) to derive a new set of adjustment factors $\bar{\bar{A}}_i$.

(11) Multiply the elements in each row of Table 6-4 by the corresponding ratio in column 6.

(12) Continue to repeat steps 7 to 11 until the ratios in column 6 and row 6 in the resulting tables approximate unity.

When the computation comes to an end, the T_{ij} will be as in Table 6-5. In the process, each row i will have been multiplied in succession by one or more ratio values $\bar{A}_i$, $\bar{\bar{A}}_i$, $\bar{\bar{\bar{A}}}_i$,... and each column by one or more values $\bar{B}_j$, $\bar{\bar{B}}_j$, $\bar{\bar{\bar{B}}}_j$,.... The equation (6-13) will have been modified to be

$$T_{ij} = (A_i B_j M_i M_j)/d_{ij} \tag{6-14}$$

where

$A_i = \bar{A}_i \times \bar{\bar{A}}_i \times \bar{\bar{\bar{A}}}_i$... is taken to designate that for each element a_i of A_i,

$a_i = \bar{a}_i \times \bar{\bar{a}}_i \times \bar{\bar{\bar{a}}}_i ....$

and

$B_j = \bar{B}_j \times \bar{\bar{B}}_j \times \bar{\bar{\bar{B}}}_j ...$ is taken to designate that for each element b_j of B_j,

$b_j = \bar{b}_j \times \bar{\bar{b}}_j \times \bar{\bar{\bar{b}}}_j ....$

Table 6-5 Projected trips with a doubly constrained gravity model

	D_1	D_2	D_3	Total Outflows
O_1	107	13	40	160
O_2	47	334	69	450
O_3	46	23	111	180
Total Inflows	200	370	320	790

In equation (6-14) A_iB_j may be viewed as a derived gravitational constant G_{ij} reflecting the complementarity of the resources and other attributes of the two masses.

The above represents one practical way trip patterns can be estimated given values of α, β and b, assumed or obtained from previous calibration studies. For practical applications regarding constrained models, see Haynes and Fotheringham (1984) and Fotheringham and O'Kelly (1989).

From this example of flow estimation for a doubly constrained gravity model, where only totals (sometimes called marginals) of outflows and inflows are given, we can immediately see how the flows of singly constrained models are derived. If the constraints are on outflows only, that is are the 160, 450, and 180 for O_1, O_2, and O_3 in Table 6-1, then the resulting inflows after steps 1 to 5 are taken are those in row 4 of Table 6-3. These total inflows for D_1, D_2, and D_3 are, after crude rounding, 161, 437, 192 which total to 790, the total of trips. If the constraints are on total inflows, that is the 200, 370 and 220 of D_1, D_2, and D_3 respectively, then the derived total inflows that are produced when the equivalent of Steps 2 to 5 are done, first with respect to the columns rather than the rows of Table 6-1,[12] the resulting total inflows for O_1, O_2, and O_3 would be the resulting row totals of the estimated T_{ij}.[13]

This process for obtaining consistent numbers for cells of a matrix when only the totals of rows and columns are given and when the row and column totals need to add up to the same overall total is designated the RAS procedure. It is one that is often used in input-output studies for updating production coefficients, or for deriving a set of consistent production coefficients given some relevant other set. See chapter 3, pp. 89–92.

At this point consider the case where a researcher finds it more desirable to specify the masses at O_i and D_j ($i = 1,...,m; j = 1,...,n$) in terms of a single attribute (college population, or professional jobs) rather than in terms of a more appropriate measure of the masses that reflects say relevant weights modified by agglomerative (deglomerative) effects and externalities of attributes captured in the α and β exponents of equation (6-10). He/she may then accompany the single attribute measures of masses at O_i and D_j with two matrices. One would be the composite v_{ij} of factors pushing out units at O_i to destinations in general. See the (3x3) **v** matrix below. The second would be the composite w_{ij} of attracting factors pulling in units at D_j from origins in general. See the (3x3) **w** matrix below.

The **v** Matrix

O_1	v_{11}	v_{12}	v_{13}
O_2	v_{21}	v_{22}	v_{23}
O_3	v_{31}	v_{32}	v_{33}
	D_1	D_2	D_3

The **w** Matrix

O_1	w_{11}	w_{12}	w_{13}
O_2	w_{21}	w_{22}	w_{23}
O_3	w_{31}	w_{32}	w_{33}
	D_1	D_2	D_3

Also, he/she may set up a matrix G_{ij} of factors indicating the complementarity of attributes for each pair of origins and destinations, to take into account, for example, how the skill composition of workers at a given origin i matches the skill requirement of jobs at a given destination j.

6.6 Calibration, tests and applications of spatial interaction models

The calibration — the derivation of the parameters of a gravity model from a set of interaction data — has several valuable uses. First of all, we in regional science, geography and related fields may want not only to test for the influence of distance in many different sets of socio-economic-political interaction data but also to know the intensity of that influence. Further, there is the interesting question: does a satisfactory calibration suggest that the concept of mutual energy as defined by physics is more appropriate than that of gravitational force for the study of interactions of interest to social scientists. Moreover, if a statistically significant influence is not found in calibration, is such an outcome an indication of: an inadequate definition or measurement of mass, specification of weights, definition or measurement of distance; or the use of an improper functional form; or some combination of these inadequacies? These are indeed meaningful questions for one concerned with the identification of commonalities among physical, biological and social (cultural) interactions.

Aside from greater understanding and more insights resulting from calibration, the derived parameters may be extremely valuable for observing and gaining further knowledge of changes in a system. This is possible, for example, when calibration is performed for comparable sets of interaction data for different points of time to identify changes in migration patterns, shopping behavior, international trade (commodity movements) and so forth. However, as already indicated in previous sections, the most extensive use of a set of the derived parameters has been as a set of base period data for forward forecasting — for example, the impact of new investments in roads, transportation pricing, rehabilitation of central city districts, etc. On occasion, a base period set of data has been used for forecasting a situation in the past (backward forecasting).

Calibration involves the use of statistical procedures such as those discussed in chapter 4 on regional econometrics, the ones having been found most helpful by gravity modelers are Least Squares including Ordinary Least Squares (OLS) (see section 4.2.2) and Maximum Likelihood (see section 4.2.4). The particular use of these procedures and their variations in gravity modelling is well discussed in Fotheringham and O'Kelly (1989) and Smith and Sen (1995) to which the reader is referred since discussion of them is beyond the scope of this chapter. However, we do wish to present a sketch of each of two interesting calibrations.

6.6.1 The use of OLS (ordinary least squares) to test the effect of distance, cooperation and hostility upon trade of nations

While theoretically the potential usefulness of the gravity model in helping to understand trade among nations was pointed to decades ago (Isard and Peck, 1954), empirical work to calibrate a model to test the influence of the distance variable has only taken place in recent years.

A small scale study largely designed to calibrate one pertinent model of trade among nations was undertaken with regard to the spatial interaction of Turkey with selected OECD (Organization for Economic Cooperation and Development) nations. The basic model employed was that of equation (6-4) to which were added two political variables:

$$T_{ij} = g[(GNP_i \times GNP_j)/d_{ij}^{b}]\,(C_{ij}/H_{ij}) \qquad (6\text{-}15)$$

where

- T_{ij} = trade between originating country i and terminating country j
- g = a gravitational-type constant
- GNP_i = economic mass of originating country i (a measure of potential supply of commodities)[14]
- GNP_j = economic mass of terminating country j (a measure of potential demand for commodities)[15]
- d_{ij} = effective economic distance (in terms of equivalent nautical miles) between i and j
- H_{ij} = level of hostility between i and j
- C_{ij} = level of cooperation between i and j
- b = an exponent to which the economic distance variable is to be raised

The data for GNP for the diverse countries were obtained from standard sources; they were converted to US$ and adjusted to the exchange rates and price levels of 1985. Use of airline distances between capital cities, frequently employed in studies, was highly inadequate for this investigation. Instead, for each pair of nations, the weighted average of distances between their major economic centers was used, following the practice generally considered best of treating each land mile as equivalent to two nautical models.[16] Levels of hostility and cooperation between each pair of nations were COPDAB (Conflict and Peace Data Bank) data, or data developed by Yaman (1994) using the standard COPDAB procedures which have found wide acceptance among quantitative international relations scholars.[17] The data on trade were developed by taking the 1985 OECD data (converted to US dollars) on current exports from and imports to Turkey as a base. To them were applied other OECD data for each year on annual growth rate of total real exports from and imports to Turkey backward and forward from 1985, while at the same time maintaining the inter-temporal differences of the trade shares of each country.[18]

For testing (calibration) purposes, several equations were specified for (exports) EX_{ij} and for imports IM_{ij} (rather than a single equation for net trade) in order to determine whether the explanatory variables have differential impacts on imports and exports. For estimation purposes in a first model, the two basic equations were (in log form):

Model 1

$$\begin{aligned}\ln(EX_{ij}) = G &+ \beta_1\ln(GDP_i) + \beta_2\ln(GDP_j) + \beta_3\ln(DIS_{ij}) \\ &+ \beta_4\ln(COP_{ij}) + \beta_5\ln(HOS_j) + U_{ij}\end{aligned} \qquad (6\text{-}16)$$

and

$$\ln(\mathrm{IM}_{ji}) = \mathrm{A} + \alpha_1\ln(\mathrm{GDP}_i) + \alpha_2\ln(\mathrm{GDP}_j) + \alpha_3\ln(\mathrm{DIS}_{ij}) + \alpha_4\ln(\mathrm{COP}_{ij}) + \alpha_5\ln(\mathrm{HOS}_j) + \varepsilon_{ij} \quad (6\text{-}17)$$

where $\mathrm{G} = \ln\beta_0$, $\mathrm{A} = \ln\alpha_0$, $\mathrm{U}_{ij} = \ln u_{ij}$, $\varepsilon_{ij} = \ln e_{ij}$ and u_{ij} and e_{ij} are random disturbances; and where subscript i represents Turkey in each equation.

For comparative purposes, and to estimate the effect of the distance variable alone, a second model was constructed whose equations were

Model 2

$$\ln(\mathrm{EX}_{ij}) = \mathrm{G} + \beta_1\ln(\mathrm{GDP}_i) + \beta_2\ln(\mathrm{GDP}_j) + \beta_3\ln(\mathrm{DIS}_{ij}) + \mathrm{U}_{ij} \quad (6\text{-}18)$$

and

$$\ln(\mathrm{IM}_{ji}) = \mathrm{A} + \alpha_1\ln(\mathrm{GDP}_i) + \alpha_2\ln(\mathrm{GDP}_j) + \alpha_3\ln(\mathrm{DIS}_{ij}) + \varepsilon_{ij}. \quad (6\text{-}19)$$

The data set consisted of pooled time series and cross-sectional data for Turkey and its OECD trading partners for the eleven-year period from 1980 to the end of 1990. There are at least three reasons for using a pooled data set in this study rather than separate time series or cross-sectional data sets. First, the pooled data set provides significantly more degrees of freedom than do either of the alternatives. Second, using only cross-sectional data would produce separate parameter estimates for each year, but within each year there would be no variation in the explanatory variable measuring Turkey's GDP. Third, although it would be possible using time series data to estimate import and export equations for Turkey and each trading partner individually, there would be no variation in the distance variable in each equation. Thus it would not be possible fully to test each equation's underlying hypotheses with either of the two alternative data sets. Ordinary least squares (OLS) then was used to estimate the parameters (which, in this case also represent the appropriate elasticities) of the two models.

Parameter estimates for the two models are presented in Table 6-6. The estimates for exports are followed by those for imports. M1 and M2 indicate, respectively, the model with cooperation and hostility variables included, and the model without the cooperative and hostility variables.

In both models, the GDP variables for Turkey and its trading partners have the expected sign (+) and are highly statistically significant. Similarly, for the distance (DIS) variable (but with an expected negative coefficient) in each model.

The results, however, are less sanguine when considering the cooperation and hostility variables. Although estimates of the coefficients

Table 6-6 Estimation results

			Constant	$LnGDP_i$	$LnGDP_j$	$LnDIS_{ij}$	$LnCOP_{ij}$	$LnHOS_{ij}$	R^2	Ad. R^2	F–Ratio
Exports											
1980–90	M1	Coeff.	9.395	3.534	0.927	−2.184	0.097	0.008	0.862	0.858	239.0
(N=198)		t-ratio	6.68	12	18.7	19.3	2.67	0.271			
		p-value	≤0.0001	≤0.0001	≤0.0001	≤0.0001	0.0082	0.7867			
	M2	Coeff.	9.275	3.754	1.011	−2.295	—	—	0.856	0.854	384.0
		t-ratio	6.62	13.1	26.1	−22.6	—	—			
		p-value	≤0.0001	≤0.0001	≤0.0001	≤0.0001	—	—			
Imports											
1980–90	M1	Coeff.	5.565	2.771	0.935	−1.245	0.092	−0.029	0.854	0.851	225.0
(N–198)		t-ratio	4.53	10.8	21.6	−12.6	2.9	−1.15			
		p-value	≤0.0001	≤0.0001	≤0.0001	≤0.0001	0.0042	0.2534			
	M2	Coeff.	5.165	2.945	0.991	−1.290	—	—	0.848	0.846	360.0
		t-ratio	4.21	11.7	29.2	−14.5	—	—			
		p-value	≤0.0001	≤0.0001	≤0.0001	≤0.0001	—	—			

for the cooperative (COP) variable have the correct sign (+) and are statistically significant in both the export and import equations, the same cannot be said for the hostility (HOS) index. In the import equation (M1), the coefficient of the hostility index has a negative sign as expected, but the evidence indicates that the null hypothesis (that is, where the coefficient is equal to zero) cannot be rejected at the usual confidence levels. In the export equation in model M1, the sign is positive, but again the evidence indicates that the coefficient is not statistically significantly different from zero.

In both the export and import equations in the M2 model, the coefficient estimates of both GDPs and distance are similar to their respective M1 model in terms of magnitude, sign and statistical significance. In addition, both R-squared measures are similar. The calculated elasticities estimate the relative impacts of each of the explanatory variables on both imports and exports, and in additon provide useful quantitative estimates for policy purposes. For example, Turkey's GDP has the greatest relative impact of any of these explanatory variables on Turkish exports, indicating an increase of 3.5 to 3.8 percent in Turkish exports for each percent increase in its GDP. The distance variable provides the next largest impacts (from –2.2 to –2.3) followed by the trading partner's GDP (from 0.9 to 1.0), and the political variables (from 0.008 to 0.097). A similar hierarchy of results was obtained with the import models.

Finally, there are differential impacts of the explanatory variables on both exports and imports, although these differences are, in some cases, quite small (see Table 6-6). Nevertheless, there is sufficient evidence to justify the use of separate models for exports and imports.

In this study, we have reported the results of an exploratory statistical analysis of two different models to explain separately the behaviour of imports and exports between Turkey and its more important trading partners in the OECD. The results of these preliminary analyses indicate that the traditional gravity model variables (economic mass and distance) are the most important in explaining the behaviour of both imports and exports. Although the cooperation variable is also statistically significant and of the expected sign in its model, the same cannot be said for the hostility variable. However, standardized coefficient analysis indicates that these political variables have only a relatively small impact on both the import-and export-dependent variables when compared to the impacts of the more traditional gravity model explanatory variables (economic mass and distance).[19]

6.6.2 The use of maximum likelihood in a policy oriented application with particular attention to characteristics of origins, destinations and types of distance (separation) measures

In a relatively advanced study, Lowe and Sen (1996) use a gravity model to analyze a hospital patient flow system within an urban area for the purpose of forecasting the impact of policies involving health care financing reform and hospital closure. Their model, involving an exponential deterrence relation, is of the general form

$$T_{ij} = M_i M_j F_{ij} \tag{6-20}$$

where

$$F_{ij} = \exp\left[\sum_{k=1}^{K} c_{ij}^{(k)} \theta^k\right] \tag{6-21}$$

and where

T_{ij} = hospital trips (disaggregated by type)[20]
M_i = characteristics of origins (disaggregatable by type of patients)
M_j = characteristics of hospitals (disaggregatable by type of function (or complex of functions) and
c_{ij} = is a vector of separation measures $c_{ij}^{(1)},\ldots,c_{ij}^{(k)}$ and θ^k are parameters to be estimated.

However, they convert the model to the potential form which can either (a) emphasize access of an individual (patient) at i to all relevant hospitals as in

$$_iV = \Sigma T_{ij}/M_i = \sum_j M_j F_{ij} \tag{6-22}$$

which is designated *patient access*, or (b) emphasize the access of a particular hospital j to all the individuals at relevant origins

$$_jV = \Sigma T_{ij}/M_j = \sum_i M_i F_{ij} \tag{6-23}$$

which is designated *hospital success* in the market. In their forecasting they estimate M_j and F_{ij} from base period data on T_{ij} [which is the number of trips from each subarea (zip code district) i to each hospital j]. They then use these M_j and F_{ij} in their model to forecast T_{ij}.

Recognizing that the problem of evaluating policy impact requires the disaggregation of hospitals by characteristics (e.g., Medical School Hospitals, Major Teaching Hospitals, Community Hospitals, etc.), they treat flows to each type and examine hospital success by type. Thus, their study illustrates the need, frequently encountered, to attend to characteristics of origins and destinations in the application of the gravity model, especially when addressing spatial choice behavior. Because the authors had access to extensive data, they were able to conduct useful disaggregations, that is examine spatial interactions with respect to different types of hospitals, while still meeting the large sample size requirement.

Of particular interest is their examination of several different forms of spatial separation and definitions of pertinent distance. They first distinguish between the appropriateness of physical distance and travel time as a measure of the first spatial component ($c_{ij}^{(1)}$) in equation (6-21) by running their model (based on equation 6-20) using (1) distance (physical), (2) log of distance, (3) the square root of distance, (4) travel time, (5) log of time and (6) the square root of time. They obtain the results of Table 6-7. They conclude that travel time (models 4, 5, and 6) provide better fits than distance (models 1, 2, and 3). They also consider the square root of time to be the best of the time measures suggesting that one longer hospital trip is proportionately less burdensome than more than one shorter trips. That is, a model with a falling-off effect with increase in time is to be preferred to one with the more extreme falling-off effect realized with the use of the log of time measure.

Next, they consider a second type of separation measure (the ($c_{ij}^{(2)}$), a social-economic type of separation). They recognize that the poorer patients without hospital insurance or funds from government sources are discouraged from using certain identifiable hospitals, which insured patients from wealthier zip code areas can use. Accordingly, they develop a measure which reflects affordability and the admitting practices of hospitals with regard to sources of payments. A payer compatibility index is constructed, ranging from 0 (no match) to 1 (perfect match) of a zip code i with a hospital j that comprises the second component ($c_{ij}^{(2)}$) of equation (6-21).

Further, to recognize the geographically dispersed markets of medical-school affiliated hospitals, they create an indicator (a dummy variable) γ_{ij} which they set at unity for all i if hospital j is a medical school, or zero otherwise. Thus, they obtain $c_{ij}^{(3)} = \gamma_{ij} c_{ij}^{(1)}$.

Finally, to account for the relatively concentrated local markets characterizing all hospitals they set $c_{ij}^{(4)} = \delta_{ij} c_{ij}^{(1)}$ where $\delta_{ij} = 1$ when zip code i is adjacent to hospital j, otherwise $\delta_{ij} = 0$. The results are recorded in Table 6-8. The standard errors are quite small when additional separation measures are added to the measure given by the square root of travel time while the chi-square ratio is significantly reduced. Hence the authors conclude that the 'payer match, medical school and adjacent trip

Table 6-7 Parameter estimates for different distance and travel time separation measures, 1987 data[a]

Measure	Model 1	Model 2	Model 3	Model 4	Model 5	Model 6
Distance[b]	-0.00054 (0.0000)					
Log Distance		-1.92852 (0.0020)				
Square Root Distance			-0.07429 (0.0001)			
Travel Time[c]				-0.13526 (0.0002)		
Log Time					-3.18650 (0.0031)	
Square Root Time						-1.37392 (0.0015)
Chi-Square Ratio	9591.66	23.24	18.28	148.27	15.66	15.66

[a]Standard errors in parentheses.
[b]Units — UIC Geography coordinate system.
[c]Travel time in minutes from UIC Geography and Urban Planning.

Table 6-8 Parameter estimates and standard errors for gravity models 6 through 9, all trips 1987 data[a]

Separation Measure	Model 6	Model 7	Model 8	Model 9
Travel Time[b]	-1.37392 (0.0015)	-1.28388 (0.0016)	-1.21444 (0.0019)	-1.16160 (0.0019)
Payer Match		4.85569 (0.0315)		3.87926 (0.0313)
Medical School			0.33386 (0.0033)	0.29462 (0.0033)
Adjacent Zip			0.18678 (0.0011)	0.16275 (0.0011)
Chi-Square Ratio	15.66	13.07	10.99	9.64

[a]Standard errors in parentheses.
[b]Square root of travel time.

code measures provide additional explanation of hospital trip behavior,' p. 450.

In many other ways the authors fruitfully conduct a careful and advanced analysis of an application of gravity models for policy purposes. For example, using estimated parameters of origin and destination characteristics and spatial separation, they forecast what the impact of a universal health measure would be if it were appropriate to set the payer match from model 9 to zero.

6.7 Conclusion

In this chapter we have treated the gravity and gravity-type models, still not widely recognized by social scientists, as reflecting basic spatial behavior of society. We have looked at several of these models, and have cited literature on many others. The possibilities for fruitful applications are numerous.

An excellent source book on many of these applications is Fotheringham and O'Kelly (1989). These authors extensively develop the rationales of spatial choice models, emphasizing the discrete choice (trade-off decision) problem of an individual behaving unit and taking into account both spatial awareness and spatial information processing. They also examine interregional and interurban migration studies, examining hierarchical information processing use and interesting variations of the strict gravity model. With regard to retailing, they go beyond single-stop trip modelling and extend the analysis of and cover applications to situations where multi-stop multipurpose trips are involved. Finally, they examine applications of location-allocation models wherein an optimization framework is employed to identify both (1) the optimal site for a single facility (hub) and (2) the best set of sites for a network of facilities (hubs) wherein forces for clustering and decentralization are examined.

Still there are innumerable other potentially fruitful applications stemming from the commonality of the gravitational effect, not only in the diverse social and natural sciences, but also in non-scientific fields. And for the regional scientist, the applications will certainly be many following the development of the synthesis of the gravity models (models much more capable of treating space and spatial interactions) with others such as input-output, social accounting, econometric, programming and applied general equilibrium discussed in this book. These models just mentioned deal by and large with regions as points of concentrated activities connected at best by a discrete number of transport and communication lines.

Endnotes

1 Development and applications of gravity models may be found in, among others, the references cited at the end of this chapter, especially Carrothers (1956). In order to facilitate the understanding of the mathematical terms used in this chapter, we have followed traditional notation on gravity models and have refrained from using a strange (and to some extent more complicated) notation which would be consistent with that of the preceding chapters and with any mathematical formulation of the fused frameworks of the chapters to follow. Sections 6.1 and 6.2 draw heavily upon Isard et al., 1960.

2 This point of view is still effective in applications.

3 In Figure 6.1, $a = 3.9$ and $b = 1.5$.

4 Studies depicting such close association will be cited at later points in this chapter.

5 For the moment we ignore a discussion of the interaction of subarea i with itself. This point is taken up later in section 6.5.

6 We do not discuss here problems connected with the use of existing data on originating and destination areas, or in the choice of them. Carrothers (1956) has investigated the degree to which a potential calibration of a gravity model actually represents what it purports to represent when different sizes and shapes of areas are involved as well as different internal distributions of relevant masses. He concludes that the best set of general-purpose areas tends to satisfy as closely as possible, among others, the following criteria: (1) absence of concentrations of mass on the peripheries of the area; (2) existence within each area of a definite nodal center of gravity of mass; (3) coincidence of the center of gravity of a relevant mass with the center of gravity of the physical area; and (4) regular geometric shapes for the physical area of each region.

7 Theoretically an *effective distance* d_{ij} between actor categories (places) i and j may be conceived as the dot product:

$d_{ij} = \mathbf{x}_{ij} \cdot \mathbf{w}_{ij}$

where $\mathbf{x}_{ij}$ is a vector in n-dimensional space, each component of which measures one type of distance just noted, and where $\mathbf{w}_{ij}$ is an n-dimensional vector indicating the respective weights to be applied to the several components.

8 In another trip matrix study, Choukroun (1975) has noted that the exponential function is appropriate when all trip makers are relatively identical. When they are not and where the distance-decay parameters for the individuals are distributed according to a gamma distribution, the power deterrence function is to be preferred.

9 A spatial separation model may be generalized as: $I_{ij} = GM_i(\cdot)M_j(\cdot)F(d_{ij})$ where (1) $M_i(\cdot)$ and $M_j(\cdot)$ are unspecified origin and destination functions corresponding to the O_i and D_j variables when they are weight functions incorporating cross-catalytic effects of their attributes, (2) $F(d_{ij})$ is a distance deterrence function, simple or generalized, and (3) G is the *universal* gravitational constant. For other possible generalizations, see Sen and Smith (1995), chapter 2.

A related general model, a theory of movement, considers a closed system of groups each composed of units wherein units move from one group to another, each group being both a possible origin and destination. (See Alonso, 1976, summarized in Anselin and Isard, 1979 and Fotheringham and O'Kelly, 1989.) Push-out factors of any group which induce units to leave are repulsive characteristics of the group (such as widespread poverty and high crime rates). Pull-in characteristics (low unemployment rates, low crime rates, etc.) are intrinsic attractions. The pull-in factors of groups, however, are attenuated by (1) the friction of distance in the broadest sense covering all forms of distance, inclusive of the affinity of a group i to any other group; and (2) the 'ease of entry' into any targeted group. The ease of entry may be related to congestion when many units attempt to enter existing groups, or to repugnance toward inmovers exhibited by existing members of an attracting group, etc. The push-out factors are also subject to attenuation. Aside from the friction of distance (again viewed in the most general sense), low responsiveness of dissatisfied units to attractions elsewhere, distrust of information disseminated by groups at potential destinations, and other elements may diminish the 'ease of exit' factor. Alonso's theory does yield as movement M_{ij} between group i and j the relation: $M_{ij} = k\,\overline{v}_i\,\overline{w}_j d_{ij}$ where $\overline{v}_i$ is the weighted sum of repulsive factors at group i adjusted for 'ease of exit,'

$\bar{w}_j$ is the weighted sum of attractive factors at group j adjusted for 'ease of entry' and k is a factor of proportionality.

10 Choice models of this sort involving the behavior of one or more individuals may be more appropriately designated 'competing destination' models. See, for example, Fotheringham (1991). See also Fotheringham and Pitts (1995) on the direction of distance variables.

11 We cite this example since the procedure for derivation of the outcome is somewhat different than the one adopted here. The reader is referred to the former procedure for further understanding.

12 Where the O_i are constrained,

$$A_i = (\sum_j B_j D_j / d_{ij})^{-1}$$

Where the D_j are constrained,

$$B_j = (\sum_i A_i O_i / d_{ij})^{-1}$$

13 For diverse uses of constrained models, see Fotheringham and O'Kelly (1989). From their experience these authors find that while the doubly-constrained model provides the highest quality of information, the singly-constrained model provides a larger amount of information though of lesser quality. The unconstrained model provides the most information, but this information is lowest in quality and consequently not generally considered to be acceptable.

These authors also discuss quasi-constrained interaction models, designated *relaxed spatial interaction models*. Such models have been developed to treat situations where because of limited information one or more of the marginal totals of the predicted interaction matrix is constrained to lie within a specified range of values. In addition, they discuss Tobler models that replace the traditional multiplicative framework with an additive one for achieving balance among the estimated T_{ij}.

14 Excluded were Iceland, New Zealand, Portugal and Greece. For reasons, see Yaman (1995), p. 4.

15 Admittedly, GNP is not as good a measure of mass of an exporting country (an originating area) as of mass of an importing country (a

destination area). Yet GNP is as least as good a general measure of productivity and export supply potential as any other.

16 See Yaman (1994) for details.

17 See Yaman (1994) for development of the levels of hostility and cooperation.

18 This procedure was designed to eliminate the effect on dollar data due to variation among the sample countries in their exchange rates and price indices. See Isard, Saltzman and Yaman (1997) and Yaman (1994) for further discussion.

19 Although the results of these preliminary experiments indicate that the theory behind the use of political variables to explain international trade finds little support in these data, it may be the case that these variables are, nevertheless, significant. For example, the data used to measure hostility and cooperation may not be wholly representative of the political phenomena which affect bilateral trade between nations. In addition, the defined hostility variable may not be a significant measure among a group of nations that are dedicated to cooperative efforts as are the members of the OECD. Clearly, more research remains to be done with these models in order to establish the importance of the political variables for explaining international trade.

The use of covariance models and/or error components models could help to sharpen our results by relaxing some of the assumptions inherent in the OLS models tested in this exploratory phase of the research. Also, more sophisticated model structures, such as those using simultaneous equations, could help to test and/or develop new theories about how economic, political and regional science variables interplay in bilateral trade.

20 Strictly speaking, for forecasting purposes T_{ij} is the expected number of trips, viewed as a Poisson random variable in a system where the locational and frequency independencies requirements are assumed by the authors to be approximately met — an assumption that some scholars may question with regard to the Chicago market which was the source of data.

References

Anderson, James E. 1979. 'A Theoretical Foundation for the Gravity Equation,' *American Economic Review*, 69, 106–16.

Beckmann, M. J., and T. F. Golob. 1972. 'A Critique of Entropy and Gravity in Travel,' in G. F. Newell (ed.), *Traffic Flow and Transportation.* New York: American Elsevier.

Bergstrand, J. H. 1985. 'The Gravity Equation in International Trade: Some Microeconomic Foundations and Empirical Evidence,' *Review of Economics and Statistics*, 67, 474–81.

Bergstrand, J. H. 1989. 'The Generalized Gravity Equation, Monopolistic Competition, and the Factor Proportions Theory in International Trade,' *Review of Economics and Statistics*, 71, 143–53.

Carrothers, G. A. P. 1956. 'An Historical Review of the Gravity and Potential Concepts of Human Interaction,' *Journal of the American Institute of Planners*, 22, 94–102.

Choukroun, J. M. 1975. 'A General Framework for the Development of Gravity-type Trip Distribution Models,' *Regional Science and Urban Economics*, 5, 177–202.

Fotheringham, A. S. 1991. 'Migration and Spatial Structure: The Development of the Competing Destination Model,' in J. Stillwell and P. Congdon (eds.), *Migration Models: Macro and Micro Approaches.* London: Bellhaven.

Fotheringham, A. S., and M. E. O'Kelly. 1988. *Spatial Interaction Models: Formulations and Applications.* Dordrecht, Netherlands: Kluwer Academic Publishers.

Fotheringham, A. S., and T. C. Pitts. 1995. 'Directional Variation in Distance-Decay,' *Environment and Planning A*, 27, 715–29.

Gasiorowski, M. J. 1986. 'Economic Interdependence and International Conflict,' *International Studies Quarterly*, 30, 23–38.

Gasiorowski, M., and S. W. Polachek. 1982. 'Conflict and Interdependence: East-West Trade and Linkages in the Era of Détente,' *Journal of Conflict Resolution*, 26, 709–29.

Haynes, Kingsley E., and A. Stewart Fotheringham. 1984. *Gravity and Spatial Interaction Models.* California: Sage Publications.

Hua Chang-i and F. Porell. 1979. 'A Critical Review of the Development of the Gravity Model,' *International Regional Science Review*, 4, 97–126.

Isard, W. 1975. 'A Simple Rationale for Gravity Model Type Behavior,' *Papers of the Regional Science Association*, 35, 25–30.

Isard, W., D. F. Bramhall, G. A. P. Carrothers, J. H. Cumberland, L. N. Moses, D. O. Price, and E. W. Schooler. 1960. *Methods of Regional Analysis: An Introduction to Regional Science*. Cambridge, Mass.: MIT Press.

Isard, W., and W. Dean. 1987. 'The Projection of World (Multiregional) Trade Matrices,' *Environment and Planning A*, 19, 1059–66.

Isard, W., and M. Peck. 1954. 'Location Theory and Trade Theory: Short-Run Analysis,' *Quarterly Journal of Economics*, 68, 305–20.

Isard, W., S. Saltzman, and A. Yaman. 1997. 'Trade and Conflict: A Gravity Model Reformulation,' in M. Chatterji and Kaizhong Yang (eds.), *Regional Science in Developing Countries*. London: Macmillan.

Linnemann, H. 1966. *An Econometric Study of International Trade Flows*. Amsterdam: North Holland.

Lowe, J. M., and A. Sen. 1996. 'Gravity Model Applications in Health Planning: Analyses of an Urban Hospital Market,' *Journal of Regional Science*, 36, 437–62.

Niedercorn, J. H., and B. V. Bechdolt. 1972. 'An Economic Derivation of the Gravity Law of Spatial Interaction: A Further Reply and a Formulation,' *Journal of Regional Science*, 12, 127–36.

Niedercorn, J. H., and J. D. Moorehead. 1974. 'The Commodity Flow Gravity Model: A Theoretical Reassessment,' *Regional and Urban Economics*, 4, 69–75.

OECD (Organization for Economic Co-operation and Development). 1987. *OECD Economic Surveys Turkey: 1985–86*. Paris: OECD.

OECD (Organization for Economic Co-operation and Development). 1991. *OECD Economic Surveys Turkey: 1989–90*. Paris, OECD.

OECD (Organization for Economic Co-operation and Development). 1994. *OECD National Accounts 1960–1992*. Paris: OECD.

OECD (Organization for Economic Co-operation and Development) (various years). *OECD Trade by Commodities Series C*. Paris: OECD.

Polachek, S. W. 1992. 'Conflict and Trade: An Economics Approach to Political International Interactions,' in W. Isard and C. H. Anderton (eds.), *Economics of Arms Reduction and Peace Process*. New York: Elsevier.

Pollins, B. M. 1989a. 'Does Trade Still Follow the Flag?,' *American Political Science Review*, 83, 465–80.

Pollins, B. M. 1989b. 'Conflict, Cooperation and Commerce: the Effect of International Political Interactions on Bilateral Trade Flows,' *American Journal of Political Science*, 33, 737–61.

Sayrs, L. W. 1989. 'Trade and Conflict Revisited: Do Politics Matter?,' *International Interactions*, 15, 155–75.

Sen, A., and T. E. Smith. 1995. *Gravity Models of Spatial Interaction Behavior*. New York: Springer-Verlag.

Sheppard, Eric C. 1984. 'The Distance-Decay Gravity Model Debate,' in G. L. Gaile and C. J. Wilmott (eds.), *Spatial Statistics and Models*. Dordrecht, Netherlands: Kluwer Academic Publishers.

Smith, T. E. 1976. 'Spatial Discounting and the Gravity Hypothesis, *Regional Science and Urban Economics*, 6, 331–56.

Smith, T. E. 1983. 'A Cost Efficiency Approach to the Analysis of Congested Spatial Interaction Behavior,' *Environment and Planning A*, 15, 435–64.

Stewart, J. Q. 1948. 'Demographic Gravitation: Evidence and Applications,' *Sociometry*, 11.

Stouffer, A. 1960. 'Intervening Opportunities and Competing Migrants,' *Journal of Regional Science*, 1, 1–20.

Summary, R. M. 1989. 'A Political-Economic Model of US Bilateral Trade,' *Review of Economics and Statistics*, 71, 179–82.

Van Bergeijk, Peter A. G. 1989. *Trade and Diplomacy: An Extension of the Gravity Model in International Trade Theory*, Research Memorandum No. 320, Institute of Economic Research, University of Groningen, Groningen, Netherlands.

Wilson, A. G. 1970. *Entropy in Urban and Regional Planning*. London: Pion Limited.

Wilson, A. G. 1974. *Urban and Regional Models in Geography and Planning*. London: Wiley.

Yaman, A. 1994. 'An Economic and Political Model of Turkish Bilateral Trade with the OECD Countries, 1970–1990: A Gravity Model Approach,' unpublished thesis, Ithaca, New York: Cornell University.