ISBN 0 86094 044 6

ISSN 0306 6142

POTENTIAL MODELS IN HUMAN GEOGRAPHY D.C.Rich D.C. Rich

POTENTIAL MODELS IN HUMAN GEOGRAPHY

D.C. Rich

CATMOG

(Concepts and Techniques in Modern Geography)

CATMOG has been created to fill a teaching need in the field of quantitative methods in undergraduate geography courses. These texts are admirable guides for the teachers, yet cheap enough for student purchase as the basis of classwork. Each book is written by an author currently working with the technique or concept he describes.

1. An introduction to Markov chain analysis - L. Collins
2. Distance decay in spatial interactions - P.J. Taylor
3. Understanding canonical correlation analysis - D. Clark
4. Some theoretical and applied aspects of spatial interaction shopping models - S. Openshaw
5. An introduction to trend surface analysis - D. Unwin
6. Classification in geography - R.J. Johnston
7. An introduction to factor analytical techniques - J.B. Goddard & A.Kirby
8. Principal components analysis - S. Daultrey
9. Causal inferences from dichotomous variables - N. Davidson
10. Introduction to the use of logit models in geography - N. Wrigley
11. Linear programming: elementary geographical applications of the transportation problem - A. Hay
12. An introduction to quadrat analysis - R.W. Thomas
13. An introduction to time-geography - N.J. Thrift
14. An introduction to graph theoretical methods in geography - K.J. Tinkler
15. Linear regression in geography - R. Ferguson
16. Probability surface mapping. An introduction with examples and Fortran programs - N. Wrigley
17. Sampling methods for geographical research - C. Dixon & B. Leach
18. Questionnaires and interviews in geographical research - C. Dixon & B. Leach
19. Analysis of frequency distributions - V. Gardiner & G. Gardiner
20. Analysis of covariance and comparison of regression lines - J. Silk
21. An introduction to the use of simultaneous-equation regression analysis in geography - D. Todd
22. Transfer function modelling: relationship between time series variables - Pong-wai Lai
23. Stochastic processes in one-dimensional series: an introduction - K.S. Richards
24. Linear programming: the Simplex method with geographical applications - J.E. Killen

continued on inside back cover

CONCEPTS AND TECHNIQUES IN MODERN GEOGRAPHY No. 26

POTENTIAL MODELS IN HUMAN GEOGRAPHY

by

D. C. Rich

(Macquarie University, Australia)

CONTENTS

Acknowledgement

I would like to thank Peter Matwijiw and two referees for their comments on a previous draft of this monograph.

Figure 1 is reprinted in adapted form from *Geographical Review,* Volume 37, 1947, with the permission of the American Geographical Society.

Figures 5,6 and 7 are reprinted in adapted form from A.R.Pred's *City systems in advanced economies,* with the permission of the author and the publishers, Hutchinson of London and John Wiley of New York.

I INTRODUCTION

The potential model has appeared frequently in the human geography literature as an index of the intensity of possible interaction between social or economic groups at different locations. This potential interaction may be of goods, telephone calls and other forms of telecommunications, migration, and a whole range of other social and economic contacts. If we examine the situation within a city, we could add to this list many other types of interaction, such as journeys to work, shop, school or entertainment. This concern with 'potential for interaction' is quite distinct from and should not be confused with 'potential for the future' in the sense of capacity for change.

The basic notion is, then, one of <u>flows</u> - of goods and services, of ideas and information, and of people. With this emphasis, the potential model has close conceptual, empirical and historical associations with the gravity model. Whereas the gravity model is concerned with analysing or predicting an observed pattern of spatial flows, the potential model is more concerned with the <u>opportunity</u> for interaction between groups, created by their sizes and locations, than with the interaction itself. The links between the two models are explored in more detail in section II.

All applications of the potential model share this same background. However, the model is used in a variety of guises and for a number of purposes; consequently, it is probably more appropriate to refer to 'potential models' than to a single model. Ignoring, for the moment, the details of the construction and application, the versions of the basic model vary in three main ways: in their interpretation; in the phenomena dealt with; and in the reasons for using them.

<u>a)</u> <u>Interpretation:</u> potential models have been interpreted variously as giving a measure of the 'influence' of one place or group on another, as a generalised measure of concentration or density (for example, of population), as an index of the nearness of groups to a point and, most commonly, as an indicator of the geographical position, or accessibility, of groups in different places relative to each other.

<u>b)</u> <u>Phenomena:</u> potential models have been used in connection with a variety of social, economic and demographic phenomena. Consequently, the term 'potential' has been qualified by a wide range of terms: ethnic, income, supply, demand, market, economic, agglomeration, polarisation and participation potentials have all appeared in the literature as variants of the basic population potential model.

c) <u>Function:</u> potential models have been used in three main roles. First, they have been used to fulfil a descriptive or illustrative function: maps of potential values have been used to illustrate the broad spatial distributions of such phenomena as total population, scientists, and the availability of information. Second, they have been used to provide explanatory variables in statistical analyses of the spatial distribution of phenomena as diverse as industrial location, per capita incomes and the price of onions. Finally, potential models appeared in one part of the debate in the 1950s and 1960s on the most appropriate methodology for geography, and were used to support calls for a more 'scientific' approach to the subject. For example, Stewart and Warntz (1958A) saw the status and progress of geographical research as being reduced by its concentration until that time on 'microscopic' studies of individual areas; they proposed an alternative 'macroscopic' methodology involving not the collection of more case studies nor the examination of larger areas, but rather a heightened level of abstraction and generalisation as a means of securing greater understanding of the processes operating to determine spatial relationships. They saw distance, time and numbers of people as being basic elements or dimensions of socio-economic systems, just as the physical sciences had long recognised a number of fundamental dimensions. They argued that a thorough appreciation of these basic factors is essential for an understanding of the workings of human society, and they put a lot of effort into trying to derive theoretical or empirical generalisations about the spatial structure of societies using these dimensions. They further argued that the progress of the macroscopic approach had hitherto been impeded by the absence of a sufficiently abstract and subtle measure of geographical position; the potential model was proposed as such a measure and consequently as a cornerstone of a new 'macrogeography'. It is rarely accorded such a central position in geographical research today, but it is still frequently used in its descriptive and explanatory roles. This monograph concentrates on the latter two functions, and an example of each type of application is discussed in sections V and VI.

The text is as far as possible self-contained. However, the discussion assumes a basic knowledge of simple statistical methods, particularly correlation and regression analysis. A background knowledge of the gravity model is also useful but not essential; an excellent introduction is provided by Taylor (1975) in this CATMOG series.

II SOCIAL PHYSICS

(1) Background

It is important to understand the nature of the simple gravity model in order to come to grips with the potential model, because their early development in the social sciences is closely intertwined. Both came to be used because of a belief, held by some, in the value of quantitative representations of aggregate relationships in the social sciences in general and in geography in particular. This belief led to a quest for uniformities in social behaviour which can be expressed in mathematical forms more or less corresponding to the known patterns of physical science' (Stewart, 1952, p 110); both gravity and potential models came to the social sciences from physics.

The gravity model suggests that two separate groups of people, say in two cities, generate a mutual interaction in proportion to the product of the sizes of the cities, and that this interaction is impeded by the frictional effect of the intervening distance over which it must take place. In other words, the volume of interaction between the two cities is a positive function of their population sizes and an inverse function of the distance between them. This relationship can be expressed algebraically as:

$$I_{ij} = \frac{f(M_i, M_j)}{f(D_{ij})} \qquad (1)$$

where I_{ij} = the volume of interaction between city i and city j;

M_i, M_j = the population (or 'mass') of i and j; and

D_{ij} = the distance between i and j.

Both gravity and potential models are based on two broad notions related to the D and M terms in this equation. The first is that the probable frequency of interaction between people in different cities is inversely proportional to the difficulty of movement between them. This difficulty is the result of a 'frictional' effect and is directly proportional to the intervening geographical distance between the two cities. The frictional effect is the product of many factors, such as transport costs (whether for people or goods), travel time, the availability of information in one place about the contact opportunities existing in the other, and the probability of making chance contacts. The frictional effect serves to inhibit contacts and means that spatial interaction commonly exhibits a distance decay pattern, with a high volume of contacts over short distances but progressively fewer over longer distances.

The second notion is that any person in a city may be considered to generate the same interaction as any other. Consequently, the volume of interaction between two cities is directly proportional to their respective population sizes: as either M_i or M_j increases, the level of interaction between i and j also increases. Clearly, individual people do not behave identically in their propensity to interact or in other ways; rather, the gravity and potential models are concerned with an average pattern from which the deviations of individuals are assumed to cancel each other out. Thus, both gravity and potential models consider aggregate patterns of interaction between groups of people, with the level of interaction being directly related to the sizes of the groups.

Both models rely on the argument that aggregate human behaviour is subject to certain laws and that these laws are the same as those which physicists found to govern the aggregate behaviour of molecules. Human beings are not like molecules, but the behaviour of human and physical aggregates are subject to similar laws of gravitation. There is, then, an analogy between physical and social phenomena, so that the study of human interaction has often been called 'social physics' - the application of physical laws to aggregate human behaviour.

The term 'social physics' was used as long ago as 1836 by Adolphe Quetelet, a social statistician, and the search for analogies between physical and social phenomena goes back even further. At first, the term meant little more than adopting in the study of social phenomena the quantitative approach that was already well established in the physical sciences. Probably the first attempt to view human interaction in gravity terms was made by Carey (1858-9), who argued as follows:

> Man, the molecule of society, is the subject of Social Science .... The great law of Molecular Gravitation is the indispensable condition of the existence of the being known as man .... the greater the number collected in a given space, the greater is the attractive force that is there exerted... Gravitation is here, as everywhere, in the direct ratio of the mass and the inverse one of distance.

(Carrothers, 1956). This is fundamentally the same formulation as that given in equation (1).

(ii) Gravity

The gravity concept made only sporadic reappearances in the social science literature during the 80 years after Carey wrote. Our modern interest in the gravity and potential models stems from the early 1940s when J.Q. Stewart, an astronomer at Princeton University, generalised previous work on human interaction and formalised it in terms of Newton's planetary gravitational theory. His interest was raised by an observation that students at Princeton came predominantly from the local region, with progressively fewer coming from successively further away (Stewart, 1941, 1942).

This observation led him (Stewart, 1948) to explore the analogy with Newton's theory. Newton's original statement suggested that there is a gravitational force, F, between two particles of mass, mi and m_j, at a distance, D_{ij}, apart. This force attracts the two masses together and has the magnitude

$$F = G \frac{m_i m_j}{D_{ij}^2} \quad (2)$$

where G is a universal constant, the so called gravitational constant. Stewart proposed that by replacing the physicists' concept of masses with population, M, we could derive laws of demographic gravitation. From (2) we thus have a demographic force, DF, between two cities:

$$DF = K \frac{M_i M_j}{D_{ij}^2} \quad (3)$$

where K is a constant to be found empirically. Similarly, Newton's concept of the mutual energy, E, of the two masses:

$$E = G \frac{m_i m_j}{D_{ij}} \quad (4)$$

has a direct equivalent, demographic energy, DE, where:

$$DE = K \frac{M_i M_j}{D_{ij}} \quad (5)$$

Stewart (1947) suggested that the number of students attending Princeton was related to the demographic energy between the university and each home

area. However, the equations for demographic force and energy differ only in the exponent to which distance is raised; later social scientists have not felt constrained by the exponent of 1 suggested by the physical analogy. Equations (3) and (5) have been combined and generalised into a single equation representing the interaction, I, between i and j:

$$I_{ij} = K \frac{M_i M_j}{D_{ij}^b} \quad (6)$$

where b is an exponent to be derived empirically. This equation is the basic gravity model used in the social sciences, which is the focus of Taylor's monograph in this series (Taylor, 1975) and is therefore not further dealt with here.

(iii) Potential

Lagrange added the notion of gravitational potential to Newton's concepts of force and energy. Whereas equation (4) represents the mutual energy created between masses i and j, gravitational potential represents the energy created at one mass by another. For example, the energy created by j at i is (M_j/D_{ij}). We are usually more concerned with the total potential created at a point by all the masses in a system; this is the sum of all the potentials created by the individual masses. If there are n discrete masses, then the total gravitational potential at i, V_i, is given by:

$$V_i = \frac{m_1}{D_{i1}} + \frac{m_2}{D_{i2}} + \frac{m_3}{D_{i3}} + \ldots + \frac{m_n}{D_{in}} \quad (7)$$

This can be more conveniently expressed as:

$$V_i = \sum_{j=1}^{n} \frac{m_j}{D_{ij}} \quad (8)$$

By analogy, population potential, P, is given by:

$$P_i = \sum_{j=1}^{n} \frac{M_j}{D_{ij}} \quad (9)$$

As with the gravity model, it is often argued that in social sciences it is unnecessary to follow the physical analogy precisely by using a distance exponent of 1 (although Stewart and Warntz (1958B) raised arguments against using values other than unity, which are discussed in section IV(v)). A more general version of population potential is therefore:

$$P_i = \sum_{j=1}^{n} \frac{M_j}{D_{ij}^b} \quad (10)$$

Population potential may be viewed in a number of ways. It is an index of the nearness of people to one another or a measure of the influence of people at a distance. Slightly more concretely, it indicates the intensity of possible contact between people at i and those at all other locations.

But most commonly, it is regarded as an indicator of relative position, or accessibility. Thus, equation (10) gives a measure of the accessibility of people in i to the people in all parts of the area being examined.

A single potential value is difficult to interpret, and it is much more useful to calculate the potentials of all the zones (subdivisions of the study area) under investigation and to compare the scores of each zone. (Potentials are often scaled by expressing them as a percentage of the highest score of any zone in the group, to aid this comparison.) By computing potentials for a number of zones, it is possible to construct a map showing lines of equipotential, such as Figure 1. The resulting potential surface gives quite a good visual impression of geographical variations in relative accessibility. Places with relatively high potentials are those with many opportunities for interaction with other places because they are highly accessible to the major concentrations of population; on the other hand, places with low scores have relatively poor opportunities for contact.

Potential is only one of a number of indices of accessibility (Neft, 1966). The various measures differ mainly in their weighting of the mass term. In the case of potential, the mass at j is weighted by the inverse of the distance between i and j (equation 10), so that distant masses have relatively little impact on the aggregate potential of i. It is also possible to weight mass by distance directly (in which case accessibility scores are dominated by distant masses) or to adopt one of a number of other procedures. The various measures generally indicate somewhat different patterns of accessibility, and even the point identified as being the most accessible is likely to differ. Nevertheless, of the relatively simple measures that are available, potential is the most useful in a number of respects (Neft, 1966; Rich, 1978).

III LINES OF DEVELOPMENT

The basic themes and concepts of social physics have been refined and applied in a great variety of ways. Much of the work in this sphere, and certainly most recent geographical work on potential, falls into one of two broad 'schools'. The first is really a linear extension and amplification of Stewart's work while the second builds on the work of Harris (1954) and uses potential to examine geographical patterns of industrial location and regional development. Rather than trying to describe all work done on potential models, this section concentrates on these two major lines of research. Further examples of recent research are described in sections V and VI.

(i) Population and income potentials

Population potential maps have been constructed for virtually every area in the world'. Probably the earliest was Stewart's map of population potentials in the United States in 1940 (Stewart, 1947), which is reproduced in Figure 1. The major peak on the surface is at New York City, indicating that this was the place most accessible to the American population in aggregate. In rural areas east of the Sierras, the potential surface slopes downwards from New York, reflecting generally declining accessibility. Superimposed on this trend are local peaks at each major city, but these are too small to show up on this relatively coarse-grained map. Rural areas near New York City

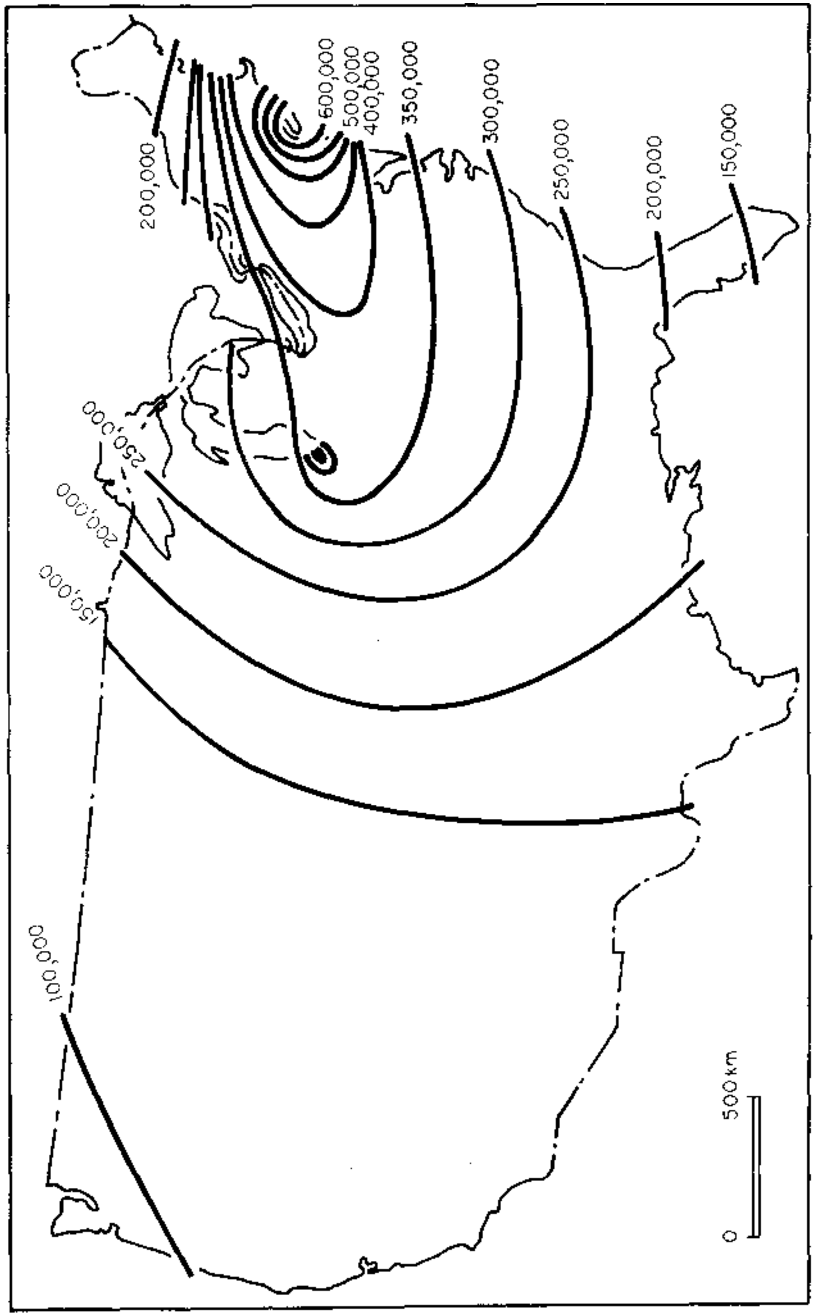

Fig. 1 Population potentials in the United States, 1940 (after Stewart, 1947)

have higher potentials than urban areas in the Mid West because of the low base potentials on which such cities stand. Consequently, as Stewart (1948, p 39) points out: you are closer to people of the whole United States on a farm in Hunterdon County, New Jersey, than you are at the center of Omaha, Nebraska'.

Another early concern was with the relationship between population potential and population density. If for the moment we regard people as being continuously distributed over an area instead of being gathered into discrete zones. potential is given by:

$$P = \int \frac{1}{D} \, SdR \qquad (11)$$

where S is the population density in an infinitesimal element of the area, dR. This formulation is of little practical value, because to quantify it requires knowledge of the precise location of each individual, but it does illustrate that potential is a sort of smoothed out density' (Stewart, 1942, p 71), where the smoothing takes place in accordance with an inverse distance formula. Population potential is therefore a weighted sum of the population densities of the surrounding areas (Craig, 1972). The two concepts are, however, complementary and not identical. Whereas density relates to an area, potential has point significance and this is reflected in the fact that (when a distance exponent of 1 is used) potential is measured in units of people per kilometre (or, more generally, mass per unit of distance), whereas density is measured in the more familiar notion of people per square kilometre. Potential is a scalar quantity; that is, it has no direction in space.

On an empirical level, potential was found to be 'very useful in describing the variations in density which occur between areas such as the counties of the United States' (Anderson, 1956, p 176). As earl: as 1941, Stewart pointed to the correlation between population potential and rural population density in 37 U.S. states east of Colorado. As well, a whole series of other variables are closely but non-linearly correlated with population potential, including rural non-farm population density, rural non-farm rents, farmland values, density of railway track, and death rates. The general form of these relationships was found to be:

$$Y = c.P^{z} \qquad (12)$$

where Y is the variable concerned, c is a constant, P is population potential and z is an exponent. This is a 'power' relationship and is linear when both Y and P are logged.

However, not all areas create the same intensity of interaction per head of population. People differ in their ability to interact, so that in calculating potentials, population is often weighted accordingly. Warntz (1957; Stewart and Warntz, 1958A) argued that weighting population by per capita income is the best means of representing each area's ability to interact. He at first called the resulting potentials 'gross economic population potentials' but the more usual term is <u>income potential.</u> Further analysis has shown that income potential has close power relationships with many variables such as the density of roads, rail tracks and telephone wires, and (inversely) with the average geographical sizes of farms, counties and states in the 48 conterminous states of the U.S.A. However, most interest has been focused on the relationships between income potential on the one hand and income density and income per capita on the other (e.g. Warntz, 1965, 1973; Coffey, 1978; Coffey and Matwijiw, 1979).

One interesting extension of this work involved attempts by Warntz (1957, 1959) and by Tegsjo and Oberg (1966) to use potential to explain the formation of prices of agricultural commodities, such as wheat, potatoes, onions, strawberries and eggs, in the U.S.A. and Sweden. The aim was to extend the simple notion that price is a function of supply and demand, by suggesting that geographical variations in prices are direct functions of spatial variations in demand and inverse functions of spatial and temporal variations in supply. Accessibility of any point to national demand was represented by income potential, while geographical variations in supply were measured by another variant of potential where the volume of local production of the commodity under investigation formed the mass term. Temporal variations in supply were represented by 'supply time potentials', a similar notion to geographical potentials except that distance is measured in days, weeks or months since the last harvest of the product. The three potentials were used as independent variables and in regression analyses were quite successful in statistically explaining prices. For example, in the case of wheat, Warntz (1959) formed the following regression equation:

$$Y = 203.97 + 0.51P' - 0.17P'' - 5.31P''' \qquad (13)$$

$$R^2 = 0.74$$

where Y = average farm price of wheat, in cents per bushel;

P' = income potential, in billions of dollars per 100 miles;

P'' = wheat supply space potential, in tens of millions of bushels per 100 miles;

P''' = wheat supply time potential, in hundreds of millions of bushels per month.

The values used were averages for the period 1940-49. All three explanatory variables were significant and the regression coefficients had the anticipated signs, thus supporting Warntz's hypothesis about the effects of supply and demand on prices.

Given this apparent success with historical data, Warntz suggested that his approach might be useful for projection purposes. For example, given a planned boost to production of a particular commodity, perhaps because of a new irrigation programme, what is the impact on the geographical pattern of prices for wheat? Any production change will alter the product supply space potential surface. Insertion of these new scores in the regression equation will allow a rough estimate of the new pattern of prices to be produced. However, the results so obtained would have to be used with considerable caution, because of difficulties with the statistical methodology and possible limitations of the potential concept in this context; alternative models of price formation are available (Chow, 1961) and may give superior results.

(ii) <u>Market potential and its derivatives</u>

Many people have suggested that potential is a useful tool in studies of industrial location and regional development. Harris (1954) was the first to use it in these fields. He argued that the distribution of the U.S. national market exerts a strong control over the location of American manufacturing industry. He adapted Stewart's basic concept to produce a measure of <u>market potential,</u> which aimed to represent the accessibility of each point to the national market, and thus to illustrate the relative attractiveness of alternative locations. Harris was content to produce a map of the market potential surface and discuss its implications; similar descriptive exercises have been carried out for other areas such as Florida (Dunn, 1956) and Southern Ontario (Kerr and Spelt, 1960). More recently, market potential has been used as an explanatory variable in regression analyses of the geographical pattern of industrial growth and decline in areas such as Southern Ontario (Ray, 1965), Nigeria (Olagbaiye, 1968; Abumere, 1978) and Great Britain (Keeble, 1976).

Harris's assumption that market access is the major determinant of industrial location has been increasingly questioned. In reality, businessmen may seek to attain a number of different goals besides sales maximisation or distribution cost minimisation, and a wide range of factors may influence location decisions (Smith, 1971). Many of these factors, such as labour supply, agglomeration economies and sources of components, are likely to be closely related to population size, as of course is the distribution of the final market. It therefore makes little sense to interpret potential simply as an index of market access. Rather, it represents the relative ease of interaction between a manufacturing plant and a whole range of activities (Clark et al., 1969; Rich, 1978) which collectively form not only the market for outputs but also the sources for many of the firm's inputs. Consequently, Clark preferred the term economic potential to Harris's earlier name. Section VI discusses an example of the use of economic potential as an explanatory variable in a regression analysis of industrial growth and decline in Scotland.

Economic potential has itself spawned a number of offspring. Blair (1976) used a very similar model, which he termed polarisation potential, to examine the extent to which local agglomeration economies in part of Philadelphia provide a favourable environment for nurturing new manufacturing firms. Richardson (1974) suggested another variant, agglomeration potential. He argued that the most important links between a firm and its economic environment vary between industries and even between firms, and that each of these links may have a different distance-decay pattern due to varying spatial frictions. The model therefore involves identifying each individual component of interaction and evaluating its particular influence on the firm's locational preferences. The data requirements of such an approach are clearly vast and the full model has not been used in practice.

Westaway (1974) developed the same basic theme in a slightly different direction. He argued that the prime factor determining the location of the headquarters of large firms, and thus a major influence on patterns of regional growth, is the availability of high quality specialist information. This information might relate to technical and commercial innovations, finance, the behaviour of competitors, government policy and the like. Much of this information cannot be adequately handled by telecommunications, so the opportunities for easy face to face contact between business executives and information sources is a major factor in locating company headquarters. Westaway devied a measure of contact potential, taking into account the geographical clustering of likely information sources, to represent spatial variations within Britain in the relative ease of information acquisition through personal contact.

IV OPERATIONALISATION

A number of decisions must be made when constructing and using a potential model, such as how best to quantify the distance and mass terms. Three, often conflicting, criteria might help to guide such decisions: the theoretical appropriateness of the particular choice; the usefulness of the resulting potential model in the empirical situation for which it is being constructed; and the ease and practical convenience of construction. Because of the wide range of circumstances in which potential models can be used, there is often no universal 'best' course of action. Therefore, the emphasis here is on outlining the range of choices available and the consequences of taking them, but first a simple worked example is given to illustrate the procedures used in computing potentials.

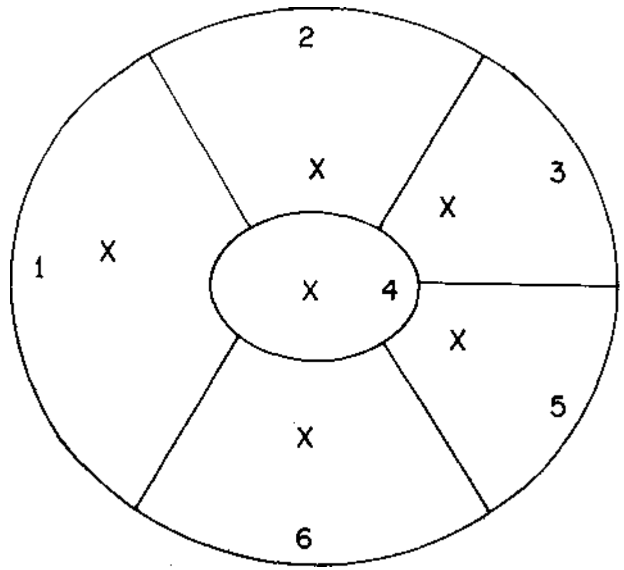

Fig. 2 Hypothetical universe and zones

(i) A simple example

Six stages are involved in the construction of a potential surface.

A. Define the universe over which potentials are to be calculated and divide it into a number of zones.

The six zones shown in Figure 2 are used in this example.

B. Obtain suitable data for the mass term, M.

C. Define a centroid for each zone and obtain a suitable measure of the distance, D, between each pair of centroids. In this example, the distance from a centroid to itself is taken to be an arbitrary, non-zero value;

the appropriate enumeration of this 'self-potential distance' term raises a number of practical problems which are outlined in section IV(v).

The centroids are marked with crosses in Figure 2.

The hypothetical mass and distance data are:

Zone i	Mass (population in '000s)	Distance (in kilometres) to zone j					
		j=1	j=2	j=3	j=4	j=5	j=6
1	5	1.5	4.2	6.3	3.8	6.7	5.0
2	20	4.2	1.5	2.3	2.4	4.2	5.1
3	12	6.3	2.3	1.5	3.2	3.1	5.1
4	15	3.8	2.4	3.2	1.5	3.0	2.7
5	8	6.7	4.2	3.1	3.0	1.5	3.4
6	7	5.0	5.1	5.1	2.7	3.4	1.5

D. Take each zone i in turn and calculate the potential exerted on it by all the zones in the universe, including itself.

In zone 1, the potential is given by:

$$P_1 = \frac{M_1}{D_{11}} + \frac{M_2}{D_{12}} + \frac{M_3}{D_{13}} + \frac{M_4}{D_{14}} + \frac{M_5}{D_{15}} + \frac{M_6}{D_{16}}$$

Using the data in the table, this is:

$$P_1 = \frac{5}{1.5} + \frac{20}{4.2} + \frac{12}{6.3} + \frac{15}{3.8} + \frac{8}{6.7} + \frac{7}{5.0}$$

$$P_1 = 3.33 + 4.76 + 1.90 + 3.94 + 1.19 + 1.40$$

$$P_1 = 16.52$$

Similar calculations give the total potentials of each zone (see table opposite)

E. Express the potential scores shown in the final column opposite as percentages of the highest score, (here 29.26).

Zone	Relative Potential (%)
1	56.46
2	100.00
3	89.30
4	97.95
5	74.40
6	67.84

F. Select a suitable contour interval and draw lines of equal potential on the map of the zones (Figure 3).

Potential at zone i	Potential created at i by zone j: j = 1	j = 2	j = 3	j = 4	j = 5	j = 6	Total potential at i
i = 1	$\frac{5}{1.5} = 3.33$	$\frac{20}{4.2} = 4.76$	$\frac{12}{6.3} = 1.90$	$\frac{15}{3.8} = 3.94$	$\frac{8}{6.7} = 1.19$	$\frac{7}{5.0} = 1.40$	16.52
i = 2	$\frac{5}{4.2} = 1.19$	$\frac{20}{1.5} = 13.33$	$\frac{12}{2.3} = 5.22$	$\frac{15}{2.4} = 6.25$	$\frac{8}{4.2} = 1.90$	$\frac{7}{5.1} = 1.37$	29.26
i = 3	$\frac{5}{6.3} = 0.79$	$\frac{20}{2.3} = 8.70$	$\frac{12}{1.5} = 8.00$	$\frac{15}{3.2} = 4.69$	$\frac{8}{3.1} = 2.58$	$\frac{7}{5.1} = 1.37$	26.13
i = 4	$\frac{5}{3.8} = 1.32$	$\frac{20}{2.4} = 8.33$	$\frac{12}{3.2} = 3.75$	$\frac{15}{1.5} = 10.00$	$\frac{8}{3.0} = 2.67$	$\frac{7}{2.7} = 2.59$	28.66
i = 5	$\frac{5}{6.7} = 0.75$	$\frac{20}{4.2} = 4.76$	$\frac{12}{3.1} = 3.87$	$\frac{15}{3.0} = 5.00$	$\frac{8}{1.5} = 5.33$	$\frac{7}{3.4} = 2.06$	21.77
i = 6	$\frac{5}{5.0} = 1.00$	$\frac{20}{5.1} = 3.92$	$\frac{12}{5.1} = 2.35$	$\frac{15}{2.7} = 5.56$	$\frac{8}{3.4} = 2.35$	$\frac{7}{1.5} = 4.67$	19.85

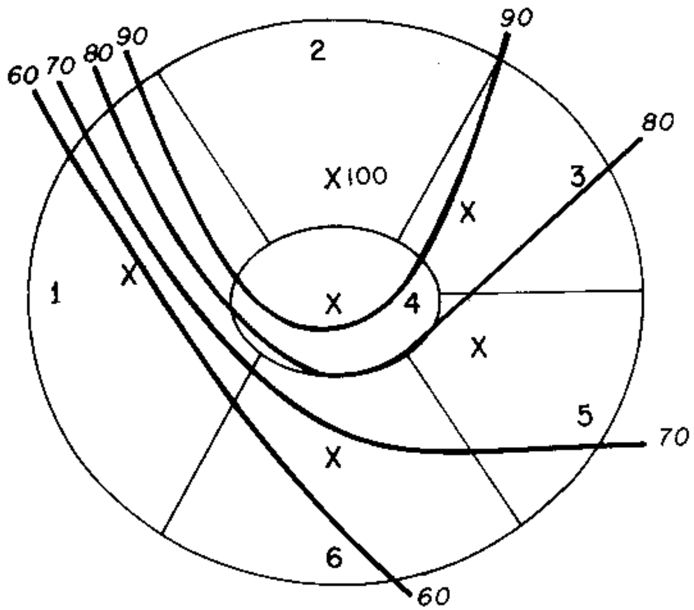

Fig. 3 Potential surface over the hypothetical universe

(ii) The universe

Every point on earth can be regarded as contributing to the potential of all others, but practical considerations usually mean that one must restrict the size of universe over which potentials are calculated. Any decision over what boundaries to choose is ultimately semi-arbitrary and this is seen by some as the central problem in all macroscopic models' (Lukermann and Porter, 1960, p 503). The extent of the universe selected will probably have some impact on the shape of the potential surfaces: Houston (1969) showed that the city identified as most accessible in Soviet Asia depends on the geographical limits put on the region. However, places of a given mass that are successively more remote create progressively less potential in the area of interest and so will have declining impact on the potential surfaces.

Selection of the appropriate universe requires some care. The important thing is that the universe should be a relatively closed social or economic system: the phenomenon under investigation should have relatively few, weak or unimportant links outside the system boundaries. For example, in a study of the social structure of a city, the appropriate boundaries might be the fringes of the metropolitan region concerned. Similarly, an industrial geographer might choose to calculate potentials over a complete nation state, even when he is interested in only one region of the nation; thus, in a study of manufacturing change in Outer South East England, Keeble and Hauser (1971, 1972) considered the whole of Great Britain to form a single universe, the argument being that the regional economy is significantly more open to flows of goods, capital and labour than the national economy. It is important to

consider each case on its merits and define the universe that most suits the problem in hand. For example, in studies of manufacturing location in Britain it might be increasingly useful to include the whole of the E.E.C. within the universe because of the growing importance of economic links between Britain and the mainland of Europe.

(iii) The zones

Conceptually, it might be possible to calculate potentials by treating each person separately. In this way, each individual j would contribute to the potential at i according to the reciprocal of the distance D. Besides the obvious practical problems involved, this may not even be a theoretically desirable approach, because below a certain level of spatial disaggregation, the distribution of population is irrelevant when using a macroscopic indicator such as potential (Craig, 1972).

Consequently, we need to divide the universe into a set of zones. Since we usually rely on using aggregate data collected for some other purpose to quantify the mass term, these zones often consist of statistical or administrative units of one sort or another. The choice is thus not so much one of selecting appropriate regional boundaries as determining the most suitable level of spatial disaggregation and the optimum number of, zones. Even here the choice is constrained by the level of disaggregation for which mass data are available.

However limited the choice of scales, it is useful to know the effect of changing the sizes and numbers of zones used. Various people have examined this problem empirically. For example, Stewart's first potential map of the United States (Figure 1) used only 24 zones; later versions used 115 (Stewart and Warntz, 1958A) and more than 3000 (Warntz, 1964). In Britain, Craig (1974) investigated the effect of scale on the East Riding of Yorkshire portion of the national population potential surface. Such experiments suggest that altering the number of zones within a given universe does not fundamentally alter the geographical pattern of the potential surface. Smaller units simply provide more detail on the surface; larger zones produce a more smoothed pattern which conceals urban peaks and rural troughs of potential. The problem when selecting the most appropriate set of zones is therefore one of balancing the amount of detail required on the potential surface against the rapidly escalating time and cost involved in using successively more zones.

(iv) Mass

The ability of areas to interact is a function of their volumes of socio-economic activity. The mass term should be a reasonable measure of the level of activity if potential is to give an adequate indication of the contact-making ability of different zones. The most appropriate definition will depend on the problem in hand, and especially on the types of interaction potential being investigated. In some studies, simple population totals have been used to quantify the mass term, but there have been many attempts to improve on this. One approach is to assign weights to the population (just as molecular weights are used in physics) so that potential is given by:

$$P_i = \sum_{j=1}^{n} \frac{W_j M_j}{D_{ij}^{\ b}} \quad (14)$$

where W_j is a measure of the relative economic or social weight of individuals in area j. This formation takes account of the fact that people in different circumstances may not all have the same ability to interact. Stewart used values of 2.0, 1.0 and 0.8 to weight population totals in the West, North and South of the United States, respectively. The implication of this is that, at the time of Stewart's research, each individual in the West had twice the capacity to interact of a person in the North and two and a half times that of one in the South. Stewart found that these estimated capacities were indeed consistent with evidence on various kinds of flows. Probably the most popular weight has been per capita income, used to give income potential surfaces. Others have suggested that population should be weighted by one or more of a whole series of variables such as occupation, age and education, as well as income, since all of these might influence the contact patterns of individuals.

A second possibility is to replace population with another variable that appears to give a good measure of local contact opportunities. For example, in his study of the market orientation of U.S. manufacturing, Harris (1954) used the value of retail sales to represent the size of the market. Similarly, Pred (1973, 1977) used the number of newspaper editions published locally each week in his study of the availability and circulation of information in pre-telegraphic America; this study is examined in more detail in section V.

A third approach, suggested by Anderson (1956) and others, is to raise the weighted or unweighted mass term to some exponent, a. Potential is then given by:

$$P_i = \sum_{j=1}^{n} \frac{M_j L^a}{D_{ij}^{\ b}} \quad (15)$$

This formulation aims to take account of factors, such as agglomeration economies, which might mean that an area's interaction capacity rises more than proportionally with its social or economic size.

Recently, the emphasis has been on simplicity in defining the mass term and many studies have used a single variable, such as zonal population, employment or income, thought to give a reasonable indication of interaction-generating capacity. There is generally less debate than in the past over which mass term to use since most of the readily available possibilities tend to be quite closely correlated and the choice among them has relatively little impact on the resulting potential surfaces (Houston, 1969). For many people, the distance term is more important and more interesting.

(v) Distance

The distance term represents the effects of the geographical separation of different masses in impeding the interaction between them. The simplest measure to use is the straight line distance between the centroids of each zone, which can be easily found using Pythagoras's theorem. When potentials are to be computed over large areas such as continents, the 'great circle'

distance is a slightly more complicated alternative. Both of these methods are easy to use and for many purposes they give adequate approximations to the effective distance separating centroids: Nordbeck (1964) found that when multiplied by a constant, or extension factor, straight line distance gives a very good statistical fit to shortest road distances within a city. The difficulty is that in some circumstances there are geographical variations in both the goodness of fit and the extension factor. For example, if we consider distances from places in the central core of England to other places in Britain, the extension factor is around 1.2, while if we look at distances from Anglesey and various peninsular areas the figure is over 1.4. Straight line distances therefore substantially underestimate the effective remoteness of many peripheral areas from the rest of the country.

For this reason, some authors have experimented with a variety of alternatives such as the distance along major transport routes. Carrothers (1958), for example, used railway distance, and Anderson (1956) suggested that it might be useful to take account of differences in the predominant mode of travel for particular types of interaction along particular routes. Many others have tried to get away from using such purely physical notions of distance and have turned to measures of economic or social separation, such as travel time or cost. A study of access to job opportunities in Atlanta (Bederman and Adams, 1974) used the average travel time between home and work zones. Harris (1954) used a measure of the costs of transporting goods, while a similar study in Western Europe (Clark et al, 1969) used a more complex cost index incorporating terminal charges, movement costs and (where appropriate) international tariffs.

A separate but related issue is the definition of the self-potential distance, D_{ii}. Self-potential is that part of the total potential of a given zone that is accounted for by the interaction of activities within the zone; it represents the portion of total interaction that is intra- rather than interzonal. The difficulty in defining D_{ii} is that if it is taken as zero then (M_i/D_{ii}) is infinity. Such a definition would, in any case, be inappropriate since it would imply that all activity in the zone is located at the centroid. A reasonable estimate of D_{ii} is important because self-potential often contributes a significant proportion of the total potential of particular zones.

Dii can be defined in a number of ways, though all are somewhat arbitrary. Applications of potential using transport cost rather than physical distance generally avoid the problem by using an estimate of intrazonal transport cost or by adding a significant terminal charge to all costs. Anderson (1956) tackled the problem by adding a constant to all distances, to avoid Dii values of zero; both Ray (1965) and Houston (1969) used an arbitrary fixed value to represent the average distance over which intrazonal contacts occur. The latter approach, though convenient, is difficult to justify because it takes no account of variations in the geographical sizes of the zones (which may be considerable if administrative units are being used). A preferable method, developed by Stewart (1947), uses the result that the self-potential of a uniform circular disc at its centre is equal to the mass divided by half the radius: the self-potential distance is taken to be half the radius of a circle of the same area as the zone concerned. Although this approach allows for variations in the size of zones, it does not allow for differences in their shape or in the distribution of the mass within them. Departures from circularity probably have a relatively minor impact onself-potential in most cases (Stewart and Warntz, 1958B), but the distribution of the mass may be more important. For example, the self-potential created by a uniformly distributed mass over a circle is only two-thirds that imposed by either a conical or a Gaussian distribution and less than half that imposed by a negative exponential distribution around the centroid (King, 1969, pp 96-97): the more concentrated the mass around the zone's centroid, the greater is the self-potential created. A partial solution to this might be to modify Stewart's approach by taking D_{ii} to be some fraction, less than half, of the zone's radius, to allow for the likely peaking of the mass around the centroid. However, at present there is no clear consensus on which definition of Dii is most appropriate.

A further issue is the selection of an appropriate exponent for the distance term. Newtonian theory suggests an exponent of 1.0, and Stewart (1947) drew faithfully on the analogy, arguing that higher exponents would render the resulting potential scores meaningless (Stewart and Warntz, 1958B). Olsson (1965, p 59) pointed out, somewhat cynically perhaps, that a variable exponent would complicate the search for social laws and constants into which Stewart and Warntz put so much effort. In contrast, most other authors have recognized the value of a more flexible approach in social and economic applications of potential. There has been no general agreement on any single exponent of general applicability, although values between 1.0 and 2.5 are commonly used.

Variations in the exponent represent differences in the effect of distance on the volume of interaction. High values indicate that distance is a major impediment to contact and that the bulk of contacts are highly localised. Conversely, low values indicate that interaction is little constrained by distance. Thus, different exponents are appropriate in different circumstances since some types of interaction are more significantly impeded by distance than others. Journey to work and to school may be much more localised than social or recreational travel and some groups of people may be more mobile than others, and the movement of goods is generally more localised than flows of information. Further, exponents seem to have shown some tendency to fall through time as transport and communication facilities improve and relative costs decline.

The value of the exponent has a significant impact on the nature of the potential surface. This is illustrated in Figure 4, which shows potential surfaces for Great Britain with exponents of 1.0 and 2.0. The lower value is associated with interaction that is little constrained by distance so that effective geographical variations in accessibility are relatively small; correspondingly, the potential surface is relatively flat, without major local variations in scores. Conversely, the larger exponent, indicating that distance is a major impediment to interaction, means that geographical variations in potential are substantial; the potential surface is strongly peaked around major population concentrations and more distant areas, by contrast, are seen to have a very low capacity to interact.

There remains the issue of how to select the distance exponent. Possible approaches are to choose a value arbitrarily or to duplicate a value used in another study. A more satisfactory method is to examine suitable interaction data, to fit a gravity model to them using a regression method described by Taylor (1975) and so derive an exponent empirically. Ray (1965) used this approach in his study of manufacturing in Ontario by regressing the tonnage

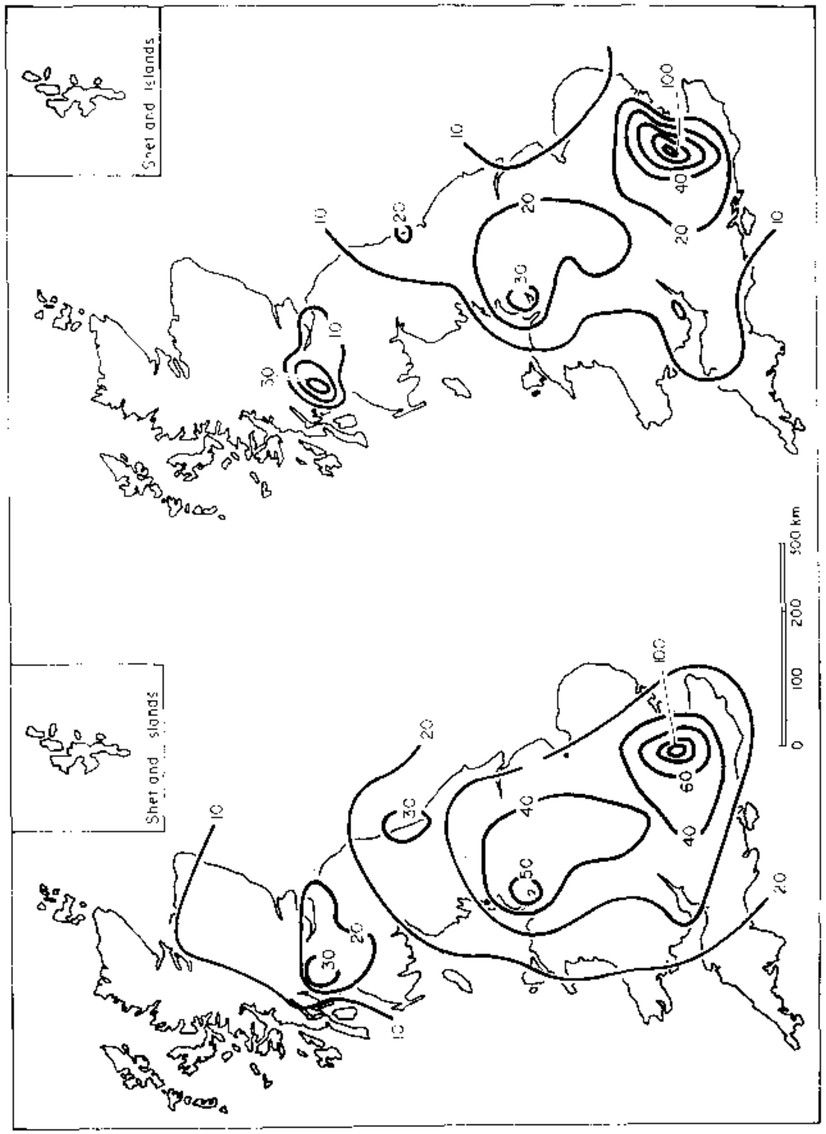

Fig. 4 Two potential surfaces for Great Britain, with distance exponents of 1.0 (left) and 2.0 (right)

of commodity flows on distance; he obtained a value of 1.42, although various technical and conceptual problems led Houston (1969) to question this result. A more pragmatic approach is to use a trial and error procedure to find the 'best' exponent in particular circumstances. When the potential surface is being constructed to act as an explanatory variable in regression analysis, 'best' may be defined as the surface providing the largest statistical fit to the dependent variable; this is the method used in the study described in section VI. This method indicates that there is no universally optimum exponent, although values much over 3.0 are rarely appropriate, and that quite small changes in the exponent may have quite a substantial impact on relative potentials and consequently on the statistical explanation achieved.

V CASE STUDY 1 - NEWS CIRCULATION IN THE UNITED STATES

One interesting application of potential by Pred (1973, 1977) used the model to illustrate geographical variations in the availability of business information in the United States between 1790 and 1840. Pred was concerned to understand the processes involved in the growth of cities, and particularly the mechanisms producing the long term stability in population size rankings of major cities that is a characteristic of many urban systems. He argued that the growth of the system of cities reflects an accumulation of decisions determining the location and size of public and private investment. Such decisions depend on the stock of relevant specialised information, since decision makers can only choose from among the alternatives of which they are aware. Commercial information, for example on costs, raw material supplies, demand and business opportunities, is almost never universally available. Rather, its distribution is spatially biassed because of the nature of the contact system through which news circulates, and because of the time lag involved in its geographical dissemination (particularly in the era before telecommunications). Investment decisions are thus made with incomplete knowledge of all the opportunities existing, and their locations reflect the information that is available. Information-rich areas tend to attract more investment, while the commercial links so created provide additional avenues for news dissemination, thereby further intensifying spatial biasses in the availability of information. Consequently, initial advantages in terms of information availability are translated, through a positive feedback process, into long term economic superiority. News of high quality, in the sense of being commercially relevant, accurate and up to date, is most readily available in large volumes in the major cities. The pattern of investment that results from this is a critical factor moulding the development of a system of cities.

In the United States of the post-independence, pre-telegraphic era, newspapers were the main means of disseminating commercial information, including advertising, shipping intelligence and wholesale prices. Further, in a period before the development of a news wire service and the appointment of non-locally based correspondents , the main source of domestic and foreign information was the reports carried in other journals. This practice of 'mutual journalistic plagiarism' (Pred, 1977, p 40) was well established as a major feature of the network by which news circulated. Consequently, a potential map of variations in accessibility to all U.S. newspapers illustrates spatial biasses in the availability of public commercial information.

Pred constructed such information accessibility maps for 1790, 1820 and 1840 (Figures 5, 6 and 7). Potential values were computed for all counties having newspapers, with the mass term being represented by the number of editions appearing weekly. Intercounty distances were measured between the population centroid of each county in 1960 (on the argument that any small shift in the population centre of gravity in the intervening period would have only a very slight effect on potential scores). Self-potential distance was derived by taking the mean of the particular county's two longest perpendicular axes; this had the effect of undervaluing the self-potential, and hence the total potential, of those counties, such as New York and Philadelphia, where there was a significant level of purely local exchange of news. The distance exponent was set to unity throughout. For ease of comparison, all potentials for 1790, 1820 and 1840 were converted to a single base, with New York's value in 1790 being set at 100.

The highest levels of accessibility to newspapers in 1790 occurred in the belt from Baltimore to Boston (Figure 5). Within this zone, accessibility peaked sharply at New York and Philadelphia. In contrast, the public-informational remoteness of the south-east and the western interior was marked, with large areas having potentials less than five per cent of that in New York, reflecting poor access to the nation's newspapers, the time lag involved in their circulation, and a low probability of contact with business information. Pred further underlined the extreme nature of geographical differences in information accessibility by comparing maps of information potential and population potential. Public information access declined very much more rapidly from New York than would be expected from the pattern of accessibility to population.

With a marked increase in the numbers of newspapers published in the thirty years to 1820, general levels of access to business information advanced considerably. Nevertheless, important spatial variations in information access remained. Whereas New York's potential had increased to an unmapped value of 302 (Figure 6), large areas of the country had failed to reach that city's 1790 score. New York was thus well on the way to establishing its pre-eminent position in the American news circulation network.

By 1840, most of the well-settled area west of the Appalachians had at last surpassed the public-information access that had been reached in New York 50 years before. Zones of moderately high potential had grown up around Cincinnati, Louisville, St. Louis and Pittsburg (Figure 7), which were already dominating the regional system of cities in the Ohio and Upper Mississippi valleys. However, by this time the dominance of New York (with an unmapped value of 825) over the circulation of domestic news in the U.S.A. was complete. The city's importance as a news source was further accentuated by its role as the primary point of entry of foreign news into the country. New York had thus achieved major competitive advantages over its commercial rivals in terms of access to business information. This advantage was a major factor allowing the city to emerge as the undisputed business capital of the U.S.A., which in turn further increased the volume of commercially useful information readily available there.

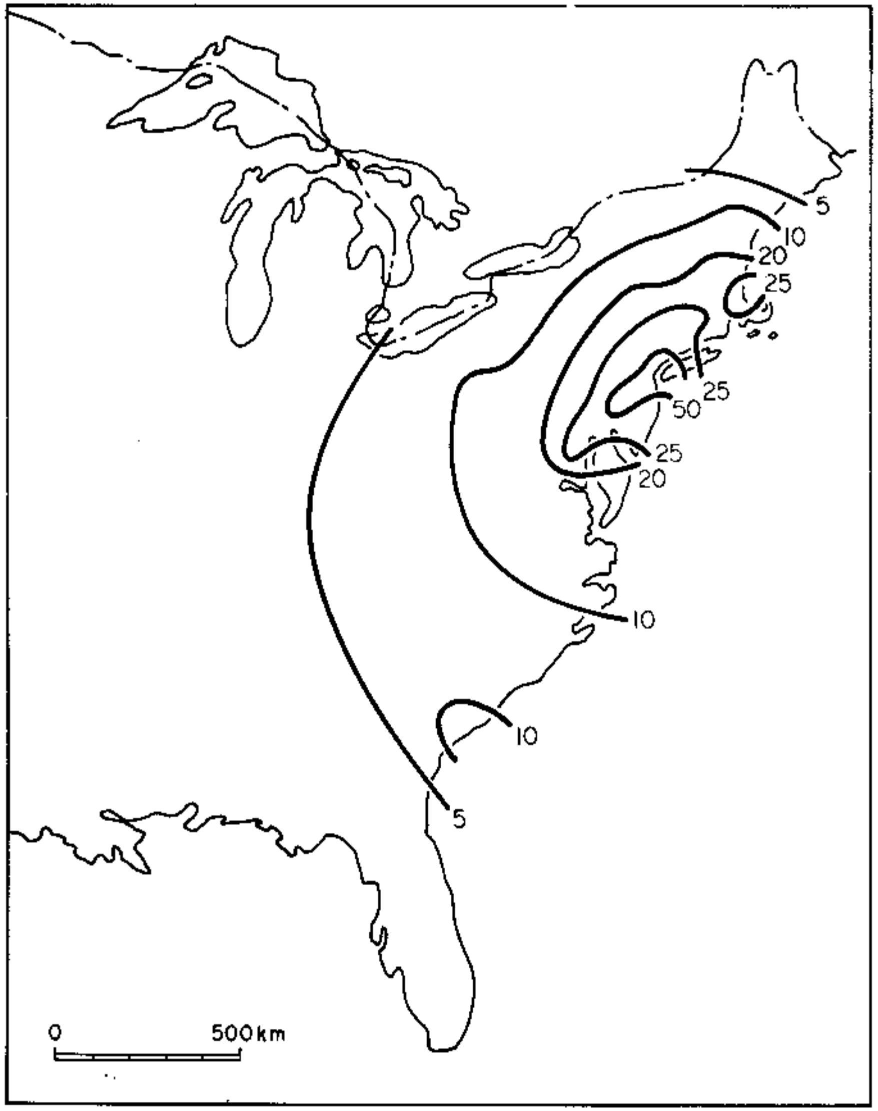

Fig. 5 Public-information accessibility in the United States, 1790
(after Pred, 1977)

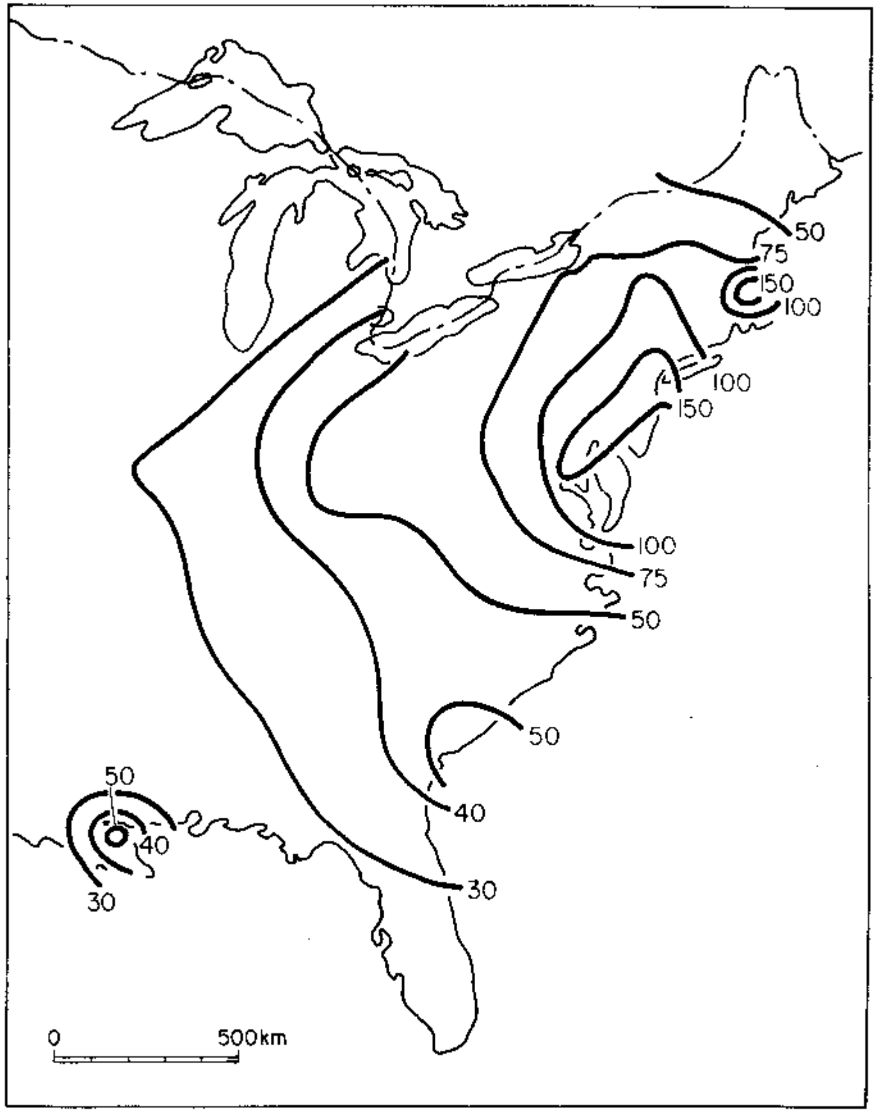

Fig. 6 Public-information accessibility in the United States, 1820
(after Pred, 1977)

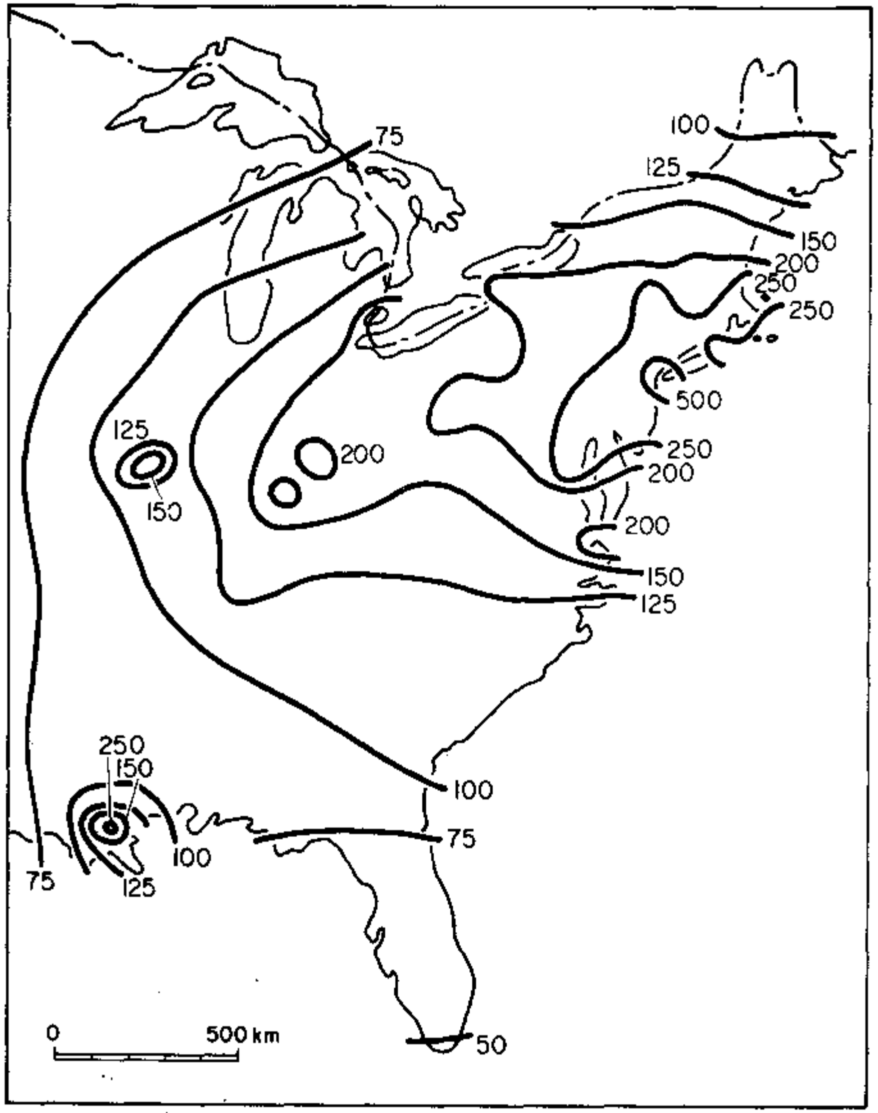

Fig. 7 Public-information accessibility in the United States, 1840
(after Pred, 1977)

VI CASE STUDY 2 - MANUFACTURING INDUSTRY IN SCOTLAND

The second example, illustrating the use of potential as an explanatory variable in statistical analysis, is a study of the geographical patterns of growth and decline in manufacturing employment in Scotland between 1961 and 1971. The work is part of an examination of the processes leading to the existence of the 'regional problem' in Britain. Two main research hypotheses were tested:

1) areas of Scotland that provide maximum accessibility to economic activity located within the region and in other parts of Britain are the most satisfactory locations for manufacturing; consequently, there is a positive linear relationship between accessibility and industrial growth; and
2) the fringes of the core areas of the Scottish economy experience the most rapid industrial growth and central areas fail to secure a share of new development proportional to their existing share of activity; consequently there is a quadratic relationship between accessibility and net industrial growth.

The idea behind the first hypothesis is that accessible areas are attractive to new industries and provide favourable environments for the expansion of existing activities, and that consequently they tend to have good records of employment growth compared with other areas. The possible advantages of centrality are many (Keeble, 1976, pp 46-71), but they may be conceived as resulting from the fact that the concentration of economic activities there provides manufacturing with a wide range of interaction opportunities. In general, firms seek locations offering a large volume of high quality potential contacts at a low cost, and because central areas offer this they have competitive advantages over remote regions. These advantages involve input and output links of all kinds; not only forward links with customers, but also backward links with sources of raw materials, components, labour, capital, information and more general urbanisation economies may be important.

The links influencing locational decisions will vary, but both inter- and intraregional contacts are likely to be relevant, so accessibility-seeking behaviour might occur at either scale in conditions approximating a free-market economy. However, British regional policy has long operated to offset the advantages of locating in the country's economic core, running from south-east to north-west England; consequently, accessibility-seeking behaviour might be more evident within regions, where public intervention might possibly have had less effect on broad patterns of location. This was one reason for testing the two hypotheses at the intraregional scale.

The contact advantages of accessible locations stimulate the economic growth of these areas. The advantages are cumulative, since by locating there firms add to the stock of interaction opportunities available to new locators. There is thus a situation of cumulative causation of the sort identified by Myrdal (1957), and by Pred (1973, 1977) in the study discussed in section V, whereby initial locational advantages are continually reinforced. There is a long term positive feedback relationship between centrality and regional growth. The first hypothesis concentrates on one part of this relationship by suggesting that in the short term, say five to ten years, growth is determined by patterns of accessibility; the feedback from regional growth to modify these patterns is ignored here.

The first hypothesis also ignores the probable emergence of spread effects which gradually develop (Hirschman, 1958; Friedmann, 1966) to modify the cumulative growth process. More concretely, the hypothesis overestimates the net benefits of locating in the most accessible areas of Scotland because of the peaking there - as in the cores of other regions - of land, labour and congestion costs. Further, the hypothesis ignores the gradual weakening of geographical constraints on the location of industry produced by economic and technical changes in industrial production and in transport and communications of all kinds. Hypothesis two thus suggests that conceptually there is a trade-off between the need to maximise potential contacts and the desire to avoid the diseconomies of excessive development.

Regression analysis was used to test the two hypotheses. The dependent variables examined were the employment changes (in jobs per 100 km^2) between 1961 and 1971 in manufacturing as a whole and in the 14 manufacturing orders (using the 1958 version of the Standard Industrial Classification). The independent variable was economic potential, scaled so that the highest value in Great Britain was 100.0. Great Britain was used as the universe over which potentials were calculated, although only the Scottish scores were used in the regression analysis, because it was felt that links both within Scotland and with other regions might influence industrial location patterns in Scotland. A variable exponent version of potential was used to represent as accurately as possible the contact-impeding effects of distance in different circumstances; the value of the exponent, b, was selected by trial and error to maximise the statistical fit of the regression model to each dependent variable. Two of the potential surfaces are shown in Figure 4.

The regression functions suggested by the two hypotheses are sketched in Figure 8. Hypothesis 1 implies that the greatest employment growth is in the most accessible regions of Scotland, namely around the Clydeside conurbation (see Figure 4). Hypothesis 2 suggests that the peak employment growth rates are found in near-central areas, such as the old counties of Ayrshire, the Lothians, Fife and Stirlingshire.

The actual pattern of employment change in manufacturing as a whole (Figure 9) does bear some resemblance to that suggested by hypothesis 2: the major employment losses were around Clydeside and the largest increases were in East and West Lothian, Ayrshire and Fife. On the other hand, a number of areas performed better (e.g. Aberdeenshire) or worse (e.g. Midlothian) than suggested. Regression analysis of these data provides little support for the linear hypothesis but shows that the quadratic model is of the form hypothesised and provides a reasonable fit to the data. The quadratic regression equation is:

$$Y = -179.2 + 63.4P - 3.1P^2 \qquad (17)$$

$$R^2 = 0.404$$

$$b = 1.3$$

Similar regression analyses showed that the patterns of change in 13 of the 14 individual industries were related to potential (Table 1) but, more interestingly, there is a clear distinction between the growing and the

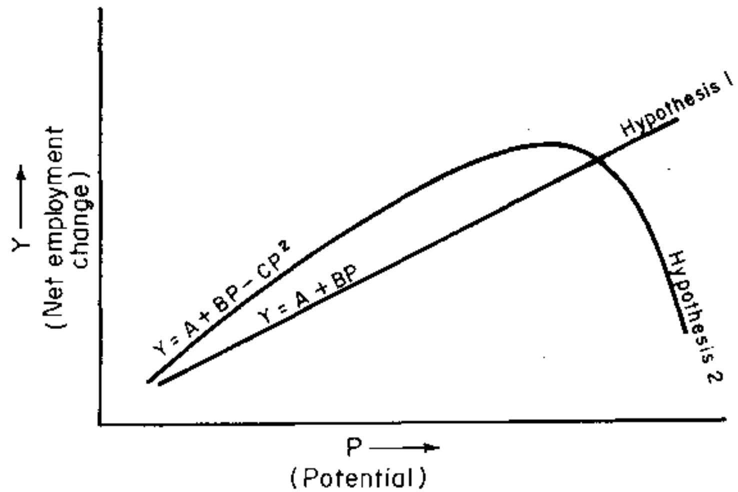

Fig. 8 Possible relationships between employment change and potential

declining industries. Six of the seven with expanding employment follow the quadratic pattern suggested by hypothesis 2, while the contracting industries exhibited much more varied patterns of change. Not one of the 14 industries supports the first hypothesis, although two have an inverse linear pattern of change, with their worst employment records in central areas.

The details of the results, and their implications for regional development, are not important here, but the analysis is consistent with the notion that, in growing industries, there is some sort of trade-off between centrifugal and centripetal locational pressures. However, the same geographical patterns could conceivably be produced by other processes. Geographical variations in the balance between nationally expanding and declining activities, and the spatially varying effects of regional policy are the two most obvious candidates; a closer examination indicates that although they do reinforce the growth of near-central areas, even when this is allowed for there is still a clear quadratic relationship between potential and employment change. This strongly suggests that the maximisation of contact opportunities and the minimisation of congestion costs are, in combination, important factors in moulding industrial change in growing activities and that economic potential can be used as a reasonable surrogate for the relative attractiveness of different areas on these criteria. On the other hand, the very different geographical patterns of change in industries with declining employment suggests that although these industries may be influenced by contact and congestion factors, other mechanisms are also important.

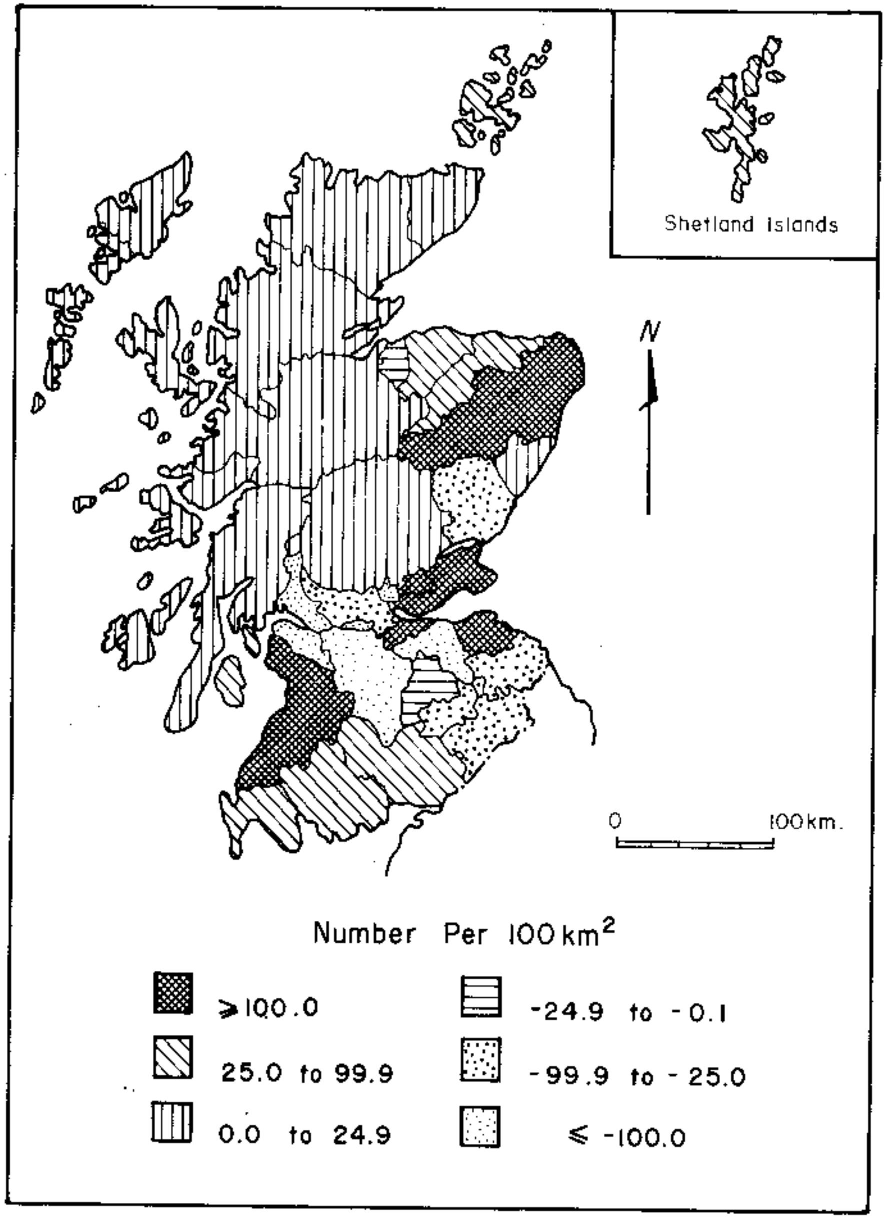

Fig. 9 Manufacturing employment density changes in Scotland, 1961-1971

Table 1: Summary of the regression relationships between economic potential and employment density changes in 14 manufacturing industries in Scotland, 1961-1971

Scottish Employment change	Type of regression relationship[a]				
	No significant relationship	Linear		Quadratic	
		B+	B-	B+C-	B-C+
Gain (+49550)[b]			Timber, Furniture etc (+170)	Food, Drink and Tobacco Engineering and Electrical goods Vehicles Metal goods not elsewhere specified Clothing and Footwear Bricks, Pottery, Glass and Cement (+49380)	
Loss (-67990)	Other Manufacturing Industries (-2990)		Shipbuilding and Marine Engineering (-22630)	Textiles (-22520)	Chemicals and Allied industries Metal Manufacture Leather, Leather goods and Fur Paper, Printing and Publishing (-19850)

Note: a) B and C are the linear and quadratic regression coefficients; hypothesis 1 suggests a positive B value, while hypothesis 2 implies that B is positive and C negative.
b) Figures in brackets represent the aggregate employment change of the industries in each category.

VII PROBLEMS AND PROSPECTS

(i) Outstanding problems

Potential models have been widely and often successfully used, but there are problems to be faced in their application. A number of technical decisions, outlined in section IV, have to be made during their construction. Some authors (e.g. Lukermann and Porter, 1960; Houston, 1969) have cited the difficulty of making these, and the semi-arbitrary nature of some of the decisions, as grounds for avoiding potential models if at all possible. However, these construction problems need not prevent potential models from making a useful contribution to geographical studies if they are used sensibly and carefully. But there are two other, perhaps more important, areas of difficulty.

The first is that by their very nature, potential models are used in the course of macro-level geographical work and suffer from the problems of all macro-analytical techniques. Simply, potential models deal with aggregates - of people, firms and the like - and consider average patterns: they have no direct behavioural input and tell us nothing about any one individual. Further, potential does not represent any single process, but is a fairly crude surrogate for a whole range of spatially varying phenomena. Consequently, it is often difficult to infer, and ultimately impossible to 'prove', the processes influencing individuals and producing observed aggregate patterns. For example, in the Scottish study just outlined, it is possible to say that patterns of industrial change are consistent with an hypothesis about the factors influencing individual firms and to discount other factors that might conceivably have produced similar patterns of change, but using potential as an explanatory variable it is not possible to go much further. Thus, the use of the potential model is not a substitute for detailed micro-level analysis of individual cases. Rather, the two approaches are complementary: while any use of potential must be firmly based on the results of relevant previous research, it is often useful in generating or improving hypotheses that can later be further tested by other methods.

The second problem area relates to the use of potential models in statistical hypothesis testing, especially their use as explanatory variables in regression analysis. Difficulties here mean that the conventional tests in inferential statistics are probably not appropriate. One aspect of the problem is that the so-called independent and dependent variables may not, in fact, be independent of one another. For example, because of the functional relationship between density and potential, some correlation is to be expected for any distribution of population potential and population density, so that the suggestion from t or F tests that the correlation is 'significant' may be entirely spurious. Testing procedures generally need to be much more conservative in this situation.

A different aspect of the same problem is that conventional inferential tests require random, independent samples (Gould, 1970), a requirement that runs counter to many geographical hypotheses. Potential is a highly spatially autocorrelated phenomenon: potential scores of neighbouring areas are, by definition, not independent of one another. This will tend to inflate the

statistical goodness of fit between potential and another spatially distributed phenomenon, and again means that conventional tests are inappropriate. One alternative approach, used in the Scottish study, is to test whether synthetic, spatially random data would be likely to have as strong a relationship with potential as that observed between potential and the dependent variable; if not, then one might be reasonably safe in suggesting some sort of structuring in the dependent variable of the type hypothesised. However, at present there is no consensus on how to deal with this difficulty and it is probably best to regard it as an unresolved problem.

(ii) Some recent advances

There is continuing interest in potential models and new applications appear fairly frequently. For example, potentials have recently been constructed to indicate the geographical distributions of scientists around the world (Inhaber, 1975), physicians in Seattle (Schultz, 1975), dentists in Newcastle-upon-Tyne (Bradley et al., 1978) and ethnic groups in Winnipeg (Matwijiw, 1979). Perhaps more interesting than these new applications of old models are attempts to refine the concepts and methods involved.

In the income potential field, the power relationships between potential and density and between potential and per capita income have been reinterpreted in entropy terms by Fein (1970), Warntz (1973) and others. This has stimulated renewed research. In particular, the social and economic processes underlying the spatial structure of the city have been explored by focussing on the varying 'integration' of different groups, that is their potential to interact with people and activities located elsewhere within the city. Boston, Toronto and Sydney have all been examined in this fashion (Coffey, 1977, 1978; Coffey and Matwijiw, 1979). The interpretation of the power relationship, and especially its exponent, are still subject to debate but it does seem likely that further investigation might lead to new insights about the processes moulding the structure of cities, and particularly the social and geographical distribution of income within them (Coffey, 1978).

The application of potential in industrial geography is continually being refined. Richardson's interesting proposal of an agglomeration potential model has already been mentioned. Tybout and Mattila (1977) have used a somewhat similar but slightly more simple approach to investigate the intrametropolitan location of 11 manufacturing industries in Detroit. They used multiple regression analysis to examine the distribution of employment in the industries in terms of a number of explanatory variables, including agglomeration economies, potential supply of inputs and potential demand for outputs for each firm. Each of these three explanatory variables was quantified using potential; the mass terms of the latter two were weighted by the relevant transaction coefficients of the U.S. national input-output table to take account of the relative importance to an industry of links with each particular type of activity. The use of input-output data to disaggregate the interaction between a firm and its environment in this way is an interesting line of research and promises to add considerably to the value of potential models in industrial location studies.

By way of conclusion, it is useful to point out that planners have also made a number of uses of potential models. Hansen (1959) developed a version of potential to examine opportunities for residential growth.

His model differed from simple population potential in that:

a) unlike population potential, it is used to predict likely population growth;

b) it includes a measure of the inherent attractiveness or physical capacity of each zone; and

c) whereas population potential measures the mutual accessibility of groups of people, Hansen's model represents the accessibility of residential areas to places of work, on the argument that work-place and journey to work are major influences on the location of residential development (Smith, 1974).

As an initial stage, Hansen argued that the likely residential development, Q, of zone i is proportional to the proximity of the zone to employment, M, and to the inherent attractiveness, or 'holding capacity', H, of the zone. In other words, potential scores are weighted by a measure of zonal attractiveness.

$$Q_i = H_i P_i \quad (18)$$

or

$$Q_i = H_i \sum_j \frac{M_j}{D_{ij}^{b}} \quad (19)$$

The development potential of i is then expressed as a proportion of the total development potential of all residential zones and multiplied by the total residential population, N, in the particular universe, so that:

$$Q_i = N \left[H_i \sum_j \frac{M_j}{D_{ij}^{b}} \right] \Big/ \left[\sum_i \left[H_i \sum_j \frac{M_j}{D_{ij}^{b}} \right] \right] \quad (20)$$

Hansen's model has been used frequently, particularly as the residential part of Lowry's 'model of metropolis' (Lowry, 1964). It has been successively revised and was eventually reformulated by Wilson (1970) in terms of observed rather than potential journey to work trips. The wheel had then gone full circle in that a potential model had been integrated back into an enlarged gravity model: More significantly, perhaps, in doing so Wilson gave the newly extended gravity model a much more rigorous theoretical base. This, like the analogy with Newton's theory, came from physics, but this time it was developed from the more modern concepts of statistical mechanics.

BIBLIOGRAPHY

Abumere, S.I. (1978), 'Multinationals, location theory and regional development : case study of Bendel State of Nigeria'. *Regional Studies,* 12., 651-664.

Anderson, T.R. (1956),''Potential models and the spatial distribution of population'. *Papers and Proceedings, Regional Science Association,* 2, 175-182.

Anscombe, F.J. (1960), 'Rejection of outliers'. *Technometrics,* 123-145.

BIBLIOGRAPHY

Abumere, S.I. (1978), 'Multinationals, location theory and regional development : case study of Bendel State of Nigeria'. *Regional Studies,* 12, 651-664.

Anderson, T.R. (1956), 'Potential models and the spatial distribution of population'. *Papers and Proceedings, Regional Science Association,* 2, 175-182.

Anscombe, F.J. (1960), 'Rejection of outliers'. *Technometrics,* 2, 123-145.

Bederman, S.H. and Adams, J.S. (1974), 'Job accessibility and under-employment'. Annals, *Association of American Geographers,* 64, 378-386.

Blair, J. (1976), 'Industrial polarization and the location of new manufacturing firms : an empirical application'. *Regional Science Research Institute, Discussion Paper Series,* No. 89.

Bradley, J.E., Kirby, A.M. and Taylor, P.J. (1978), 'Distance decay and dental decay : a study of dental health among primary school children in Newcastle upon Tyne'. *Regional Studies,* 12, 529-540.

Carey, H.C. (1858-9), *Principles of social science.* (Philadelphia : Lippincott).

Carrothers, G.A.P. (1956), An historical review of the gravity and potential concepts of human interaction'. *Journal American Institute of Planners,* 22, 94-102.

Carrothers, G.A.P. (1958), 'Population projection by means of income potential models'. *Papers and Proceedings, Regional Science Association,* 4, 121-152.

Chow, G.C. (1961) Review of Warntz, W. (1959), 'Toward a geography of price'. *Journal American Statistical Association,* 56, 209-210.

Clark, C., Wilson, F., and Bradley, J. (1969), 'Industrial location and economic potential in Western Europe'. *Regional Studies,* 3, 197-212.

Coffey, W. (1977), 'A macroscopic analysis of income regions in Metropolitan Boston'. *Professional Geographer,* 29, 40-46.

Coffey, W. (1978), 'Income relationships in Boston and Toronto : a tale of two countries?'. *Canadian Geographer,* 22, 112-129.

Coffey, W. and Matwijiw, P. (1979), 'Measures of demographic potential applied to the definition of income regions in Metropolitan Sydney'. *Tijdschrift voor Economische en Sociale Geografie,* 70, 98-109.

Craig, J. (1972), 'Population potential and population density'. *Area,* 4, 10-12.

Craig, J. (1974), 'How arbitrary is population potential?'. *Area,* 6, 44-47.

Dunn, E.S. (1956), The market potential concept and the analysis of location'. *Papers and Proceedings, Regional Science Association,* 2, 183-194.

Fein, E. (1970), 'Demography and thermodynamics'. *American Journal of Physics,* 38, 1373-1379.

Friedmann, J. *(1966),Regional development policy : a case study of Venezuela.* (Cambridge, Mass. : M.I.T. Press).

Georgescu-Roegen, N. (1971), *The entropy law and the economic process.* (Cambridge, Mass. : Harvard University Press).

Gould, P. (1970), 'Is 'Statistix inferens' the geographical name for a wild goose?' *Economic Geography,* 46, 439-448.

Hansen, W.G. (1959), 'How accessibility shapes land use'. *Journal American Institute of Planners,* 25, 73-91.

Harris, C.D. (1954), 'The market as a factor in the location of industry in the United States'. *Annals, Association of American Geographers,* 44, 315-348.

Hirschman, A.O. (1958), *The strategy of economic development.* (New Haven, Conn. : Yale University Press).

Houston, C. (1969), 'Market potential and potential transportation costs: an evaluation of the concepts and their surface patterns in the *U.S.S.R.'. Canadian Geographer,* 13, 216-236.

Inhaber, H. (1975), 'Distribution of world science'. *Geoforum,* 6, 231-236.

Inhaber, H. and Przednowek, K. (1974), 'Distribution of Canadian science'. *Geoforum,* 19, 45-54.

Keeble, D.E. (1976), *Industrial location and planning in the United Kingdom,* (London : Methuen).

Keeble, D.E. and Hauser, D.P. (1971), 'Spatial analysis of manufacturing growth in outer South-East England. 1: Hypotheses and variables'. *Regional Studies,* 5, 229-262.

Keeble, D.E. and Hauser, D.P. (1972), 'Spatial analysis of manufacturing growth in outer South-East England. 2: Methods and results'. *Regional Studies,* 6, 11-36.

Kerr, D. and Spelt, J. (1960), 'Some aspects of industrial location in Southern Ontario'. *Canadian Geographer,15,* 12-25.

King, L.J. (1969), *Statistical analysis in geography.* (Englewood Cliffs, N.J. : Prentice Hall).

Lowry, I.S. (1964), *A model of metropolis.* (Santa Monica : Rand Corporation).

Lukermann, F. and Porter, P.W. (1960), 'Gravity and potential models in economic geography'. *Annals, Association of American Geographers,* 50, 493-504.

Matwijiw, P. (1979), 'Ethnicity and urban residence: Winnipeg, 1941-1971'. *Canadian Geographer,23,* 45-61.

Myrdal, G.M. (1957), *Economic theory and under-developed regions.* (London : Duckworth).

Neft, D.S. (1966), 'Statistical analysis for areal distributions'. *Regional Science Research Institute, Monograph Series,* No. 2.

Nordbeck, S. (1964), 'Computing distances in road nets'. *Papers and Proceedings, Regional Science Association,* 12, 207-220.

Olagbaiye, J.A. (1968), 'Towards manufacturing locational analysis in Southern Nigeria : a population potential model'. *Nigerian Geographical Journal,* 11, 11-20.

Olsson, G. (1965), 'Distance and human interaction'. *Regional Science Research Institute, Bibliography Series,* No. 2.

Pred, A.R. *(1973), Urban growth and the circulation of information: the United States system of cities, 1790-1840.* (Cambridge, Mass. : Harvard University Press).

Pred, A.R. *(1977), City systems in advanced economies : past growth, present processes and future development options.* (New York : Wiley).

Ray, D.M. (1965), 'Market potential and economic shadow'. *University of Chicago, Department of Geography, Research Papers,* No. 101.

Rich, D.C. (1978), 'Population potential, potential transportation cost and industrial location'. *Area,* 10, 222-226.

Richardson, H.W. (1974), 'Agglomeration potential : a generalization of the income potential concept'. *Journal of Regional Science,* 14, 325-336.

Schultz, R.R. (1975), 'A space potential analysis of physician location'. *Proceedings, Association of American Geographers,* 7, 203-208.

Smith, D.M. (1971), *Industrial location : an economic geographical analysis.* (New York : Wiley).

Smith, P. (1974), 'The use of the potential surface technique in sub-regional planning'. *(University of Reading, Department of Geography, Geographical Papers,* No. 30.

Stewart, J.Q. (1941), 'An inverse distance variation for certain social influences'. *science,* 93, 89-90.

Stewart, J.Q. (1942), 'A measure of the influence of population at a distance'. *Sociometry,* 5, 63-71.

Stewart, J.Q. (1947), 'Empirical mathematical rules concerning the distribution and equilibrium of population'. *Geographical Review,* 37, 461-485.

Stewart, J.Q. (1948), 'Demographic gravitation : evidence and applications'. *Sociometry,* 11, 31-58.

Stewart, J.Q. (1952), 'A basis for social physics'. *Impact of Science on Society,* 3, 110-133.

Stewart, J.Q. and Warntz, W. (1958A), 'Macrogeography and social science'. *Geographical Review,* 48, 167-184.

Stewart, J.Q. and Warntz, W. (1958B), 'Physics of population distribution'. *Journal of Regional Science,* 1, 99-123.

Taylor, P.J. (1975), 'Distance decay models in spatial interactions'. *Concepts and Techniques in Modern Geography,* No. 2.

Tegsjo, B. and Oberg, S. (1966), 'Concept of potential applied to price formation' . *Geografiska Annaler,* 48B, 51-58.

Tybout, R.A. and Mattila, J.M. (1977), 'Agglomeration of manufacturing in Detroit'. *Journal of Regional Science,* 17, 1-16.

Warntz, W. (1957), 'Geography of prices and spatial interaction'. *Papers and Proceedings, Regional Science Association,* 3, 118-129.

- Warntz, W. (1959), *Toward a geography of price : a study in geo-econometrics* . (Philadelphia : University of Pennsylvania Press).

Warntz, W. (1964), 'A new map of the surface of population potentials for the United States, 1960' . *Geographical Review,* 54, 170-184.

Warntz, W. (1965), *Macrogeography and income fronts.* (Philadelphia : Regional Science Research Institute).

Warntz, W. (1973), 'New geography as general spatial systems theory - old social physics writ large?'. In: R.J. Chorley (ed.) *Directions in geography,* (London : Methuen).

Westaway, J. (1974), 'Contact potential and the occupational structure of the British urban system 1961-1966 : an empirical study'. *Regional Studies,* 8, 57-73.

Wilson, A.G. (1970), *Entropy in urban and regional modelling.* (London : Pion).

contd.

This series, Concepts and Techniques in Modern Geography is produced by the Study Group in Quantitative Methods, of the Institute of British Geographers.
For details of membership of the Study Group, write to the Institute of British Geographers, 1 Kensington Gore, London, S.W.7.
The series is published by Geo Abstracts, University of East Anglia, Norwich, NR4 7TJ, to whom all other enquiries should be addressed.