Functional Regions within the City Centre: A Study by Factor Analysis of Taxi Flows in Central London
Author(s): J. B. Goddard
Source: *Transactions of the Institute of British Geographers*, No. 49 (Mar., 1970), pp. 161-182
Published by: The Royal Geographical Society (with the Institute of British Geographers)
Stable URL: http://www.jstor.org/stable/621647
Accessed: 27-11-2017 18:07 UTC

Functional Regions within the City Centre:
a Study by Factor Analysis of Taxi Flows in Central London

J. B. Goddard, B.A.

(Lecturer in Geography, London School of Economics)

Revised MS. received 2 June 1969

ABSTRACT—This paper examines the problem of measuring the relationship between movement patterns and the location of activities within the city centre. Emphasis is placed on the problem of defining and analysing complex linkage systems. Data on taxi flows are presented as a single indicator of multi-faceted functional linkages within Central London. Cartographic analysis reveals a complex pattern of flows; correlation and factor analysis are therefore applied to the data to determine the underlying structure of the taxi-flow system. Traffic zones are compared in terms of the zones from which they receive trips; high factor loadings reveal groups of zones receiving trips from common origins; the related factor scores specify these origins. Linking groups of origins and destinations on a map reveals five functional regions. Hierarchical grouping of zones produces an objective classification of zones in terms of similarities in linkage characteristics. Comparison of the pattern of functional regions with familiar structural regions suggests a basic isomorphism between such regionalizations. More important, the analysis reveals the usefulness of the methods for investigating the structure of complex flow systems so as to specify the principal groups of observations and the linkages they have in common.

THE CLOSE connection between movement and the distribution of activities in spatial systems is a well-established geographical axiom. At the national level B. J. L. Berry has used multivariate statistical procedures to measure the relationships between the state of a system, expressed in terms of the distribution of productive activities, and the behaviour of that system, expressed in the form of commodity flows between various places.[1] He has structured these relationships within the framework of a general field theory of spatial behaviour, in which the pattern of flows, summarized as a set of functional regions, is held to be both a resultant and a determinant of the spatial distribution of activities, summarized as a set of uniform regions.[2] Berry's formulation is in some ways a more general statement of A. K. Philbrick's principle of 'areal functional organization'.[3] Similarly, O. Wärneryd, in his conceptual model of the urban system, has stressed the importance of functional interdependencies in explaining changes in the state of the system.[4]

It would be reasonable to assume that the interdependence of flows and spatial structure also holds within urban areas, in particular within large metropolitan centres. However, at this level the measurement of the relationships involved creates numerous conceptual, empirical and technical problems, many of which are insuperable at the present time. While several studies of city centre structure have emphasized the role of 'linkages' between highly specialized activities as a determinant of location, little progress has been made in measuring these linkages.[5] Most previous research has focused on the distribution of activities; linkages have then been inferred from locational associations among these activities.[6] For instance, within the financial core of the City of London, multivariate analysis of detailed employment data has highlighted the existence of distinctive spatial clusters of offices whose composition closely reflects the well-known functional organization of the City.[7] Within the broader compass of the London Central Area, floorspace data have suggested a number of sub-areas with distinctive land-use combinations.[8]

According to Berry's general field theory formulation, such pronounced spatial differentiation within city centres should simultaneously be determined by, and call forth, a correspondingly structured pattern of movement. Yet very few studies have been made of physical linkage patterns in the city centre, in relation to the activities that generate and attract these movements. In part this is because of data difficulties. While detailed information can be collected on the distribution of activities, a corresponding breakdown of movement patterns associated with each of these activities would require extensive surveys. More important, linkages frequently involve information flows that are extremely difficult to differentiate, that use a variety of communication channels and which often do not involve obvious physical movements.[9] Some of these difficulties can be by-passed by considering only physical movements of a given type, as a single indicator of multi-faceted functional linkages. But even then, observed movement patterns are extremely complex. It is therefore the object of this paper to suggest a method for analysing the complex linkage structures that are represented by movement patterns in the city centre. As the first stage in the scientific description of any complex pattern involves the development of a classificatory framework, numerical taxonomic procedures are applied to observed patterns of flows. This yields a more readily appreciated series of functional regions that can be compared with uniform regions derived from a similar analysis of activity distributions.

The Data and Study Area

For the purposes of this paper, the movements of taxis within Central London have been taken as the best available single indicator of functional linkages. Of the principal travel modes, the use of taxis is most characteristic of the Central Area; in fact, 55 per cent of all taxi trips in Greater London both begin and end in the centre.[10] However, taxis still represent only one type of communication and the subsequent analysis is therefore forced to ignore linkages maintained on foot, by private vehicles, public transport, telephone and post.

The Registrar General's definition of the Central Area for the 1961 census is taken as the limit for the study area, which for practical purposes has to be regarded as a closed system (Fig. 1). This area was subjectively extracted from the existing administrative framework by aggregating enumeration districts to provide an additional unit for collecting future census and planning data. Defined in this way, the Central Area includes an area of 28 km^2, which is much larger than the Central Business District defined using R. E. Murphy and J. A. Vance's criteria.[11] In fact, the Central Area includes both C.B.D. 'core' and 'frame' using E. M. Horwood and R. R. Boyce's terminology, thus containing the main-line rail termini, which are major generators of taxi traffic.[12]

The data were collected in 1962 as part of the London Traffic Survey. They are based on journey logs kept by an effective 10 per cent sample of all cabs registered in the Survey Area for one week in July.[13] As the logs were completed by taxi drivers, the data failed to differentiate trips according to passenger's journey purpose. An additional shortcoming of the data from the linkage standpoint is that non-fare paying trips (that is, trips made between setting down one passenger and picking up another) are not excluded. Because drivers tend to cruise back into the areas where most trips are generated, the major nodes in the system could be over-emphasized. Both of these limitations are overcome in a second survey of taxi usage conducted in July 1969 as part of a Home Office investigation into the London taxicab trade.

Trip origins and destinations determined from the logs have been assigned to seventy

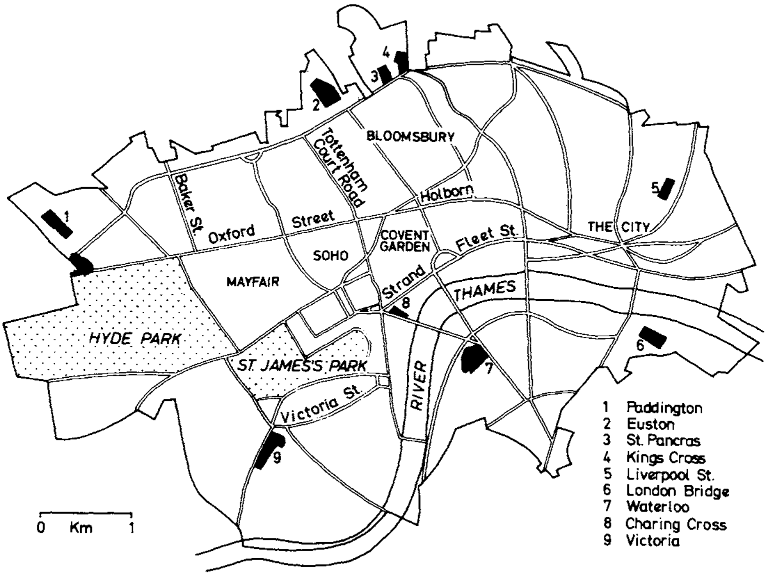

FIGURE 1—London Central Area

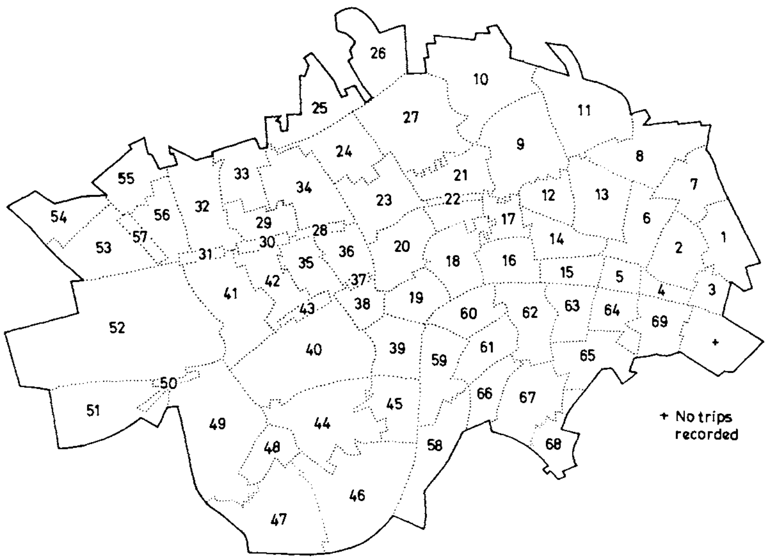

FIGURE 2—Traffic zones

traffic zones in the Central Area and averaged to give 24-hour weekday flows. All but eight of the zones consist of aggregations of enumeration districts delineated with due consideration of land-use characteristics. The remaining eight zones comprise sections of major shopping streets, for example, Oxford Street (three zones) and Piccadilly (two zones) (Fig. 2).

FIGURE 3—Taxi flows (24-hour average weekday)

The Pattern of Taxi Flows: Cartographic Analysis

Figure 3 is an attempt to describe the overall pattern of taxi flows between sixty-nine traffic zones; for one zone no trips are recorded. Flows of less than ten vehicles per day are excluded. This map at once illustrates the complexity of the linkage patterns and the limitations of cartographic analysis.[14] However, some points do emerge about the use of taxis in Central London. The most noticeable feature is that taxi flows reach a peak in the West End, particularly Mayfair and Victoria, while zones to the north and south of the City are only weakly connected, if at all, to the rest of the system. Apart from a minor node in the eastern part of the City, including the Bank, Liverpool Street and London Bridge stations, the level of taxi use is remarkably low in the financial centre. Because of the highly specialized nature of the City with its detailed clustering of linked activities, most commercial contacts are maintained on foot. In contrast, Mayfair, a recently colonized part of the commercial centre, while being characterized by broadly similar activities, does not show the same degree of clustering.[15] In addition to offices, the West End includes other major generators of taxi traffic, such as hotels, entertainments, shops and high-class residences. Such differences in the level of vehicular usage are one of the key characteristics distinguishing the 'core' from the 'frame' of the C.B.D.[16]

The other important feature of the pattern shown in Figure 3 is the main-line rail termini: some of the strongest flows are between Victoria, Waterloo, Euston, King's Cross/St. Pancras and Paddington stations. While inter-station flows, reflecting poor public transport connections, are of limited importance to the internal system, the links between the termini and particular parts of the Central Area are of major significance for activities with national control functions.

The diversity of the linkages and the limitations of cartographic methods indicate the need for some pattern-seeking technique that will eliminate the 'noise' of minor flows and concentrate on the basic elements in the system. The analysis proceeds from the belief that the complex network of linkages represented by the taxi system contains some order in the form of sub-systems with strong internal linkages but with weaker connections to other subsystems elsewhere in the Central Area.

Correlation Analysis

The flows shown in Figure 3 can be arrayed in a 69×69 origin/destination matrix, the leading diagonal representing intra-zonal flows. This matrix is not symmetrical since the outward movements from any one zone seldom correspond to the pattern of incoming trips. Several methods have been suggested for determining the basic structure of such flow matrices. In particular J. Nystuen and M. F. Dacey have suggested a method for defining nodal regions which considers flows as links in a directed graph.[17] However, this method is essentially hierarchical since it considers only largest flows; it is therefore unsuited to situations in which no theoretical hierarchy of regions can be assumed. Use is therefore made of correlation and factor analysis techniques first applied to binary connectivity matrices by W. L. Garrison, D. F. Marble and P. R. Gould but subsequently to matrices containing ratio data by Berry, S. Illeris and O. Pedersen.[18]

In the first stage of the analysis each column or destination vector in the matrix is correlated with every other, yielding a 69×69 matrix of coefficients.[19] Figure 4 illustrates how the destinations become the variables in the analysis and the origins become the observations over which each of these variables is measured. Surprisingly many of these relationships were near-linear and no transformations of the data were required.

In Figure 5, 'correlation bonds' have been drawn linking zones with highly significant similarities in the way they assemble their incoming trips.[20] The bonds do not represent functional linkages, but simply connect zones receiving their trips from similar origins. This first step adds little order to the cartographic depiction of flows. However, it does become increasingly obvious that most of the zones in the West End are very similar to each other and they have few significant associations with the City. Also, the pattern revealed in Figure 5 suggests a neighbourhood influence at work in the system, whereby the greatest similarities are to be found between adjacent zones.

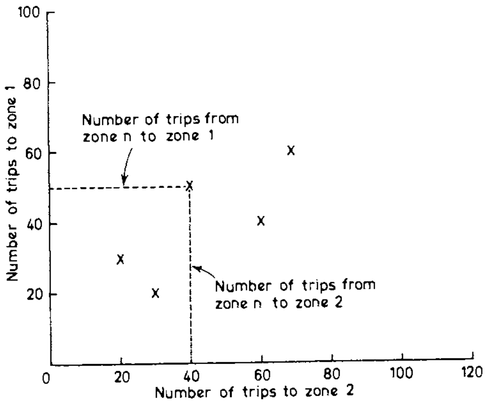

FIGURE 4—Derivation of correlation coefficients

Components Analysis

The next stage in the search for order in the taxi-flow system involved extracting characteristic values—latent roots or eigenvalues, together with their associated eigenvectors—from the correlation matrix.[21] Using the method of principal components, the sixty-nine correlated variables are transformed into sixty-nine uncorrelated variables or components in descending order of variability. Although as many components as variables can be extracted from the correlation matrix, this procedure is generally parsimonious as the leading components account for a substantial proportion of the total variance. In this instance the largest eigenvalue represents 26 per cent of the variance in the original data; together, the first five components account for 55 per cent of the variance and the first ten, 72 per cent. Geographical interpretation

FIGURE 5—Correlation bonds (0.001 significance level)

of the component loadings (the elements of the eigenvector scaled by the square root of their associated eigenvalue) indicates zones with common patterns of trip assembly.

Because each zone, no matter how weakly it is linked with the overall system, exerts an influence on the principal pattern of variation that is extracted and because of the absolute dominance of the West End, this stage in the analysis also failed to isolate meaningful sub-systems. In fact every zone to the west of the City contributed strongly to the leading component, while subsequent components represented relatively minor patterns of variation.

Varimax Rotation of Leading Components

The final step in the analysis, therefore, involved the application of factor analysis rotation procedures which concentrate on the significant patterns of variation or common factors and ignore all other influences. Examination of the eigenvalues showed that beyond the sixth, each additional component added little to the overall level of explanation. These six components, representing 60 per cent of the initial variance, were subjected to a normal varimax rotation to satisfy the criterion of simple structure, whereby each variable, as far as possible, is highly associated with only one factor. This method of principal components analysis on the complete correlation matrix, with unities in the diagonal, followed by analytic rotation of a specified number of leading components, contrasts with the classic methods of principal factor analysis whereby communality estimates of the amount of variance explained by the number of common factors specified are inserted in the diagonal before extracting the latent roots of the matrix.[22] Table I compares components and rotated factors and indicates how the variance of the principal component has been redistributed among the factors.

The rotation succeeded in decomposing the dominant West End system, distinguishing between destinations and grouping them on the basis of their common origins. But what are the origins that are common to each set of destinations? The rotated factors are derived from the intercorrelations of the sixty-nine variables (destinations) which are measured over sixty-nine observations (origins). By computing each observation's score on each of the factors, it is possible to determine which zones are the principal origins for each group of destinations.[23]

The composition of the six factors into which the overall system has been resolved is described in Table II. Only factor loadings greater than ± 0.50 and factor scores greater than ± 1.00 are shown. An important characteristic of all the factors is that they have a single pole with only high positive or high negative loadings; this greatly facilitates interpretation.

TABLE I

Explained Variance: Unrotated and Rotated Solutions

Component	I	II	III	IV	V	VI	*Total*
Eigenvalue	17.96	7.23	4.83	4.44	3.86	2.92	
Percentage explanation	26.02	10.47	7.00	6.34	5.59	4.23	59.65
Rotated Factor	I	II	III	IV	V	VI	*Total*
Sum of squared factor loadings	10.84	8.29	6.85	6.02	5.43	3.83	
Percentage explanation	15.71	12.01	9.92	8.72	7.86	5.55	59.77

Note: The total percentage explanation for six components and six factors is approximately the same (59.7 per cent), although both totals do not exactly tally owing to different rounding levels in the computing.

TABLE II

Factor Loadings (Destinations) and Factor Scores (Origins)

Factor	I		II		III		IV		V		VI		Communality
Zone	F.L.	F.S.	F.L.	F.S.	F.L.	F.S.	F.L.	F.S.	F.L.	F.S.	F.L.	F.S.	per cent
1	—	—	—	—	—	—	0.58	—	—	—	—	—	40.1
2	—	—	—	—	—	—	0.64	2.52	—	—	—	—	47.9
3	—	—	—	—	—	—	—	—	—	—	—	—	17.7
4	—	—	—	—	—	—	—	—	—	—	—	—	28.5
5	—	—	—	—	—	—	—	—	—	—	−0.57	—	54.1
6	—	—	—	—	—	—	0.82	3.72	—	—	—	−1.41	75.9
7	—	—	—	—	—	—	0.76	5.23	—	—	—	1.85	79.5
8	—	—	—	—	—	—	0.66	—	—	—	—	—	55.3
9	—	—	—	—	—	—	—	—	—	—	—	—	42.1
10	−0.51	—	—	—	—	—	—	—	—	—	−0.51	—	63.3
11	—	—	—	—	—	—	—	—	—	—	—	—	44.1
12	—	—	—	—	—	—	—	—	—	—	−0.56	—	42.8
13	—	—	—	—	—	—	0.74	2.03	—	—	—	—	56.5
14	—	—	—	—	—	—	0.80	—	—	—	—	−1.01	68.1
15	—	—	—	—	—	—	—	—	—	—	—	—	22.2
16	—	—	—	—	−0.67	—	—	—	—	—	—	—	78.1
17	—	—	—	—	−0.52	−1.23	—	—	—	—	—	—	52.7
18	—	—	—	—	−0.68	−1.56	—	—	—	—	—	—	66.7
19	—	—	—	1.76	−0.74	−4.47	—	—	—	—	—	−1.90	88.5
20	—	—	—	—	−0.71	−2.40	—	—	—	—	—	—	77.4
21	—	—	—	—	—	−1.51	—	—	—	—	—	—	50.9
22	—	—	—	—	−0.55	—	—	—	—	—	—	—	38.7
23	—	—	—	—	—	−1.52	—	—	−0.75	−2.48	—	—	72.0
24	—	—	—	—	—	—	—	—	−0.80	−1.44	—	—	67.9
25	—	—	—	—	—	—	—	—	−0.80	−3.16	—	—	87.3
26	—	—	—	—	—	—	—	—	−0.70	−4.25	—	—	83.5
27	—	—	—	—	—	—	—	—	−0.78	−1.13	—	—	66.3
28	—	—	—	—	−0.53	—	—	—	—	—	—	—	58.7
29	−0.72	—	—	—	—	—	—	—	—	—	—	—	61.2
30	−0.79	—	—	—	—	—	—	—	—	—	—	—	69.8
31	−0.86	—	—	—	—	—	—	—	—	—	—	—	77.3
32	−0.84	−3.88	—	−1.23	—	—	—	—	—	—	—	—	76.4
33	−0.81	−1.88	—	—	—	1.10	—	—	—	—	—	—	79.3
34	−0.59	−1.22	—	−1.21	—	—	—	—	—	−2.09	—	—	62.2
35	−0.71	−2.61	—	—	—	—	—	—	—	—	—	—	77.6
36	—	—	—	—	−0.60	−2.21	—	−1.06	—	−1.38	—	2.50	66.2
37	−0.60	—	—	—	—	—	—	—	—	—	—	—	51.9

38	—	—	—	—	−0.69	−1.80	—	—	—	—	—	—	76.8
39	—	—	0.59	—	—	—	—	—	—	—	—	—	51.0
40	—	—	0.53	1.25	−0.59	−1.41	—	—	—	—	—	—	74.2
41	−0.73	−3.39	—	2.12	—	—	—	—	—	1.01	—	−1.68	79.5
42	−0.76	−3.22	—	—	—	−2.24	—	1.43	—	1.22	—	1.69	84.9
43	−0.64	—	0.51	—	—	—	—	—	—	1.00	—	—	81.7
44	—	—	0.73	2.23	—	—	—	—	—	—	—	—	57.2
45	—	—	0.67	—	—	—	—	—	—	—	—	—	61.6
46	—	1.11	0.84	1.92	—	—	—	—	—	—	—	—	75.3
47	—	—	—	—	—	—	—	—	—	—	—	—	33.5
48	—	—	0.86	4.56	—	1.13	—	—	—	−1.88	—	—	89.9
49	—	−1.01	0.76	3.29	—	1.11	—	—	—	—	—	—	71.6
50	−0.59	—	0.63	—	—	—	—	—	—	—	—	—	76.2
51	—	—	0.71	—	—	—	—	—	—	—	—	—	73.5
52	−0.66	—	0.57	—	—	—	—	—	—	—	—	—	76.8
53	−0.76	—	—	—	—	—	—	—	—	—	—	—	74.1
54	—	—	0.65	—	—	—	—	—	—	—	—	—	81.0
55	−0.60	—	—	—	—	—	—	—	—	—	—	—	62.9
56	−0.83	−1.13	—	—	—	—	—	—	—	—	—	—	73.9
57	—	—	—	—	—	—	—	—	—	—	—	—	8.8
58	—	—	—	—	—	—	—	—	—	—	—	—	23.7
59	—	—	—	—	—	—	—	—	—	—	−0.59	—	72.7
60	—	—	—	—	—	—	—	—	—	—	—	—	29.6
61	—	—	0.54	1.11	—	—	—	—	−0.50	−1.34	—	−5.45	63.6
62	—	—	—	—	—	—	—	—	—	—	−0.68	—	50.1
63	—	—	—	—	−0.58	—	—	—	—	—	—	—	46.4
64	—	—	—	—	—	—	—	—	—	—	—	—	38.9
65	—	—	—	—	—	—	—	—	—	—	−0·70	—	56.2
66	—	—	—	—	—	—	—	—	—	—	—	—	21.5
67	—	—	—	—	—	—	—	—	—	—	—	—	36.7
68	—	—	—	—	—	—	—	—	—	—	—	—	8.3
69	—	—	—	—	—	—	0.74	—	—	—	—	−2.75	66.0
Sum of squares	10.84		8.29		6.85		6.02		5.43		3.83		
Per cent common variance	15.71		12.01		9.92		8.72		7.86		5.55		
Cumulative	15.71		27.72		37.74		46.46		54.32		59.77		

Note: Only loadings greater than +0.50 and scores greater than ±1.00 are shown
F.L. = Factor loading *F.S.* = Factor score.

The percentage communality, indicating the amount of variance of each variable explained by the common factors, ranges considerably, from 8.3 per cent (zone 68) to 89.9 per cent (zone 48). The zones which are poorly accounted for by the six factors are those with low levels of trip generation that are hardly linked to the rest of the system. These contrast with the zones containing the main-line termini, which mostly have high communalities because they load high on several of the factors. Having isolated the major sub-systems, it is possible to return to cartographic analysis and delineate each of these on the map. This is done in a series of maps by linking each group of destinations (factor loadings) to its respective set of common origins (factor scores).

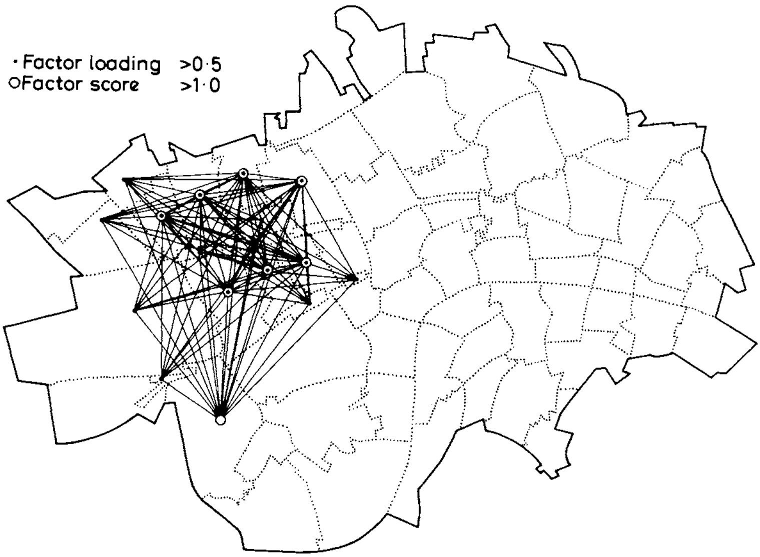

FIGURE 6—Factor one: the West End

Factor one (Fig. 6) accounting for 15.7 per cent of the common variance, represents the dominant nodal sub-system centred in the heart of the West End, including Mayfair, Oxford Street and the area to the north. This is a region characterized by a unique combination of offices, shopping facilities and high-class residences. The core zones of this system are both major generators and attractors of trips, this being shown by high factor scores and high factor loadings (zones 32, 35, 41 and 42). Peripheral to these are zones which are major destinations for trips generated in the core. These include other shopping areas along Piccadilly (37 and 43) and Knightsbridge (50).

Factor two, explaining a further 12 per cent of the common variance, distinguished a second sub-system in the western part of the Central Area which was far from apparent in the cartographic analysis. As shown in Figure 7, this is centred in the administrative area of Westminster. The focal zones, with high scores and loadings, include the distinctive localities of Belgravia (49), Victoria (44), St. James's (40) and Pimlico (46). Victoria station (48), Paddington (54), King's Cross/St. Pancras (26) and Waterloo stations (61) are all major destinations for trips generated in the core of this system, while Charing Cross (19) is a major origin. The importance of the stations in this pattern is in strong contrast to the West End system, which has no connections to the main line termini, only to the suburban services of Marylebone and Baker Street (zone 55).

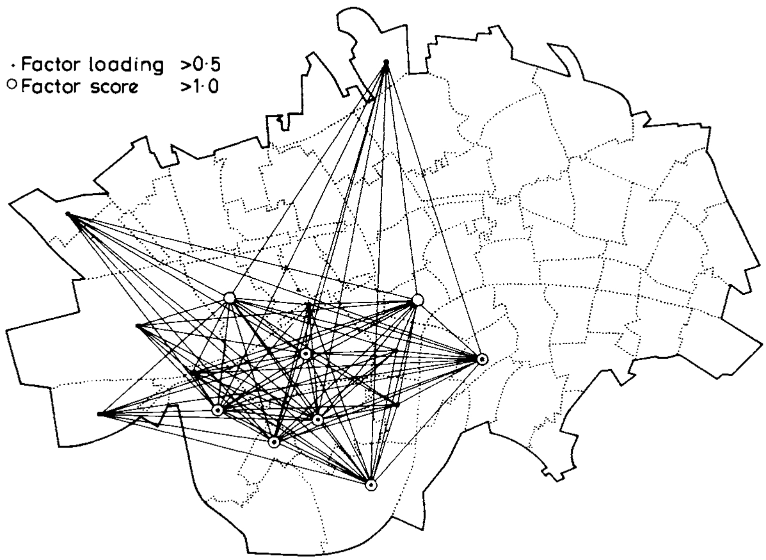

FIGURE 7—Factor two: Westminster

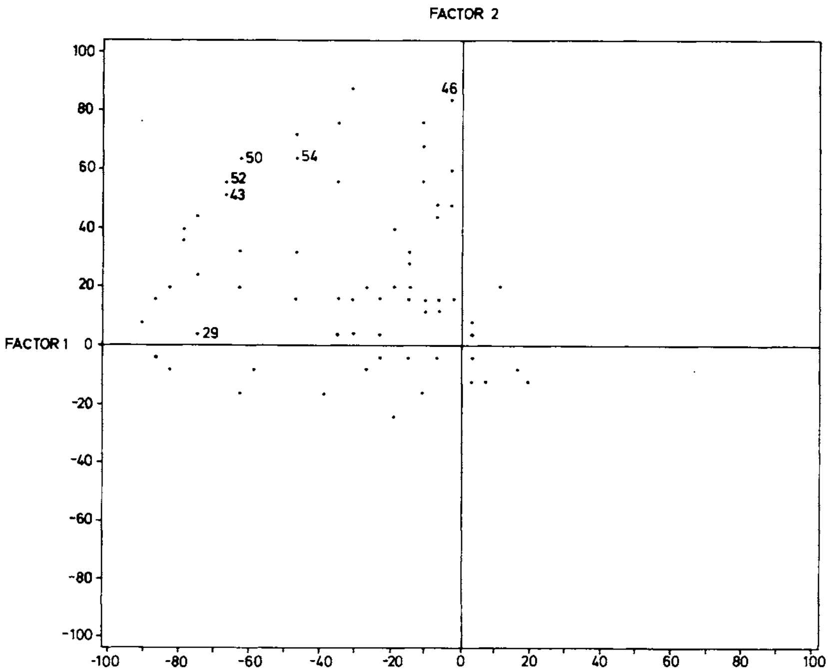

FIGURE 8—Plot of the loadings of the zones on factor one against the loadings on factor two

The failure of the initial unrotated solution to separate these first two systems was mainly owing to a certain amount of overlap between them. The maps show that certain zones, including Piccadilly and Knightsbridge, are major destinations for trips generated in both systems. Figure 8, a plot of factor one against factor two, emphasizes this overlap. It shows a number of zones, particularly numbers 43, 50, 52 and 54, loading high on both factors. The plot also shows the existence of independent sub-systems, indicated by zones that load highly on only one of the factors (for example, zone 29 on factor one and zone 46 on factor two). The varimax rotation of the leading components has therefore resolved a simple structure, picking out two distinctive hyperplanes or groupings of variables. With non-independent systems, however, an oblique rotation might have achieved a better factor resolution, but the requirements of the grouping techniques applied later in the paper preclude this.[24]

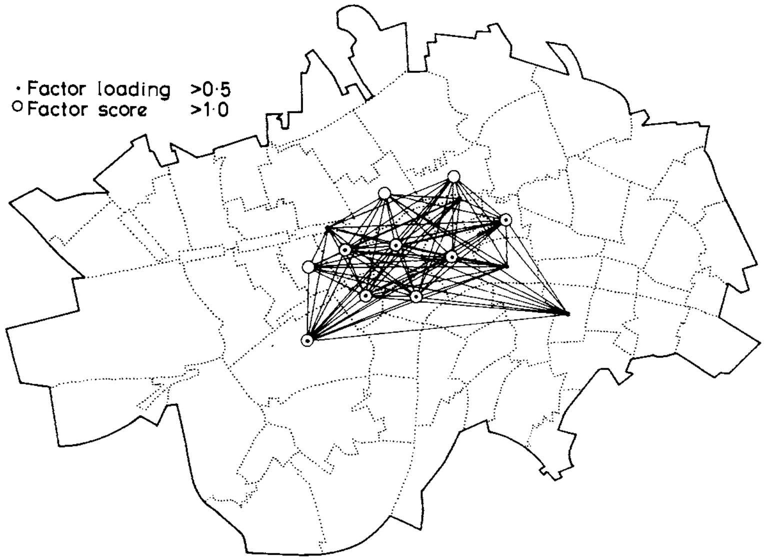

FIGURE 9—Factor three: Covent Garden, Soho and Fleet Street

The third sub-system focuses on the 'mid-town' part of the Central Area and includes Soho, Covent Garden and the Strand (Fig. 9). It accounts for a further 10 per cent of the variance. The core zones correspond to London's entertainment centre, including 'clubland', St. James's and Leicester Square. High Holborn (zone 22), the eastern end of Oxford Street (zone 28) and Fleet Street (zone 16) are peripheral destinations.

Factor four distinguishes a quite separate sub-system in the eastern part of the City that accounts for a further 9 per cent of variance (Fig. 10). It is the only sub-system that could be easily isolated from the map of flows. There is little overlap with the other systems, apart from a weak link between the City and Mayfair (zone 42). This is shown by a plot of factor two against factor four (Fig. 11). The plot reveals a clear-cut simple structure with only two zones (17 and 61) at any appreciable distance from the factor axes, representing Liverpool Street and Waterloo stations, which feed into both systems.

FIGURE 10—Factor four: the City

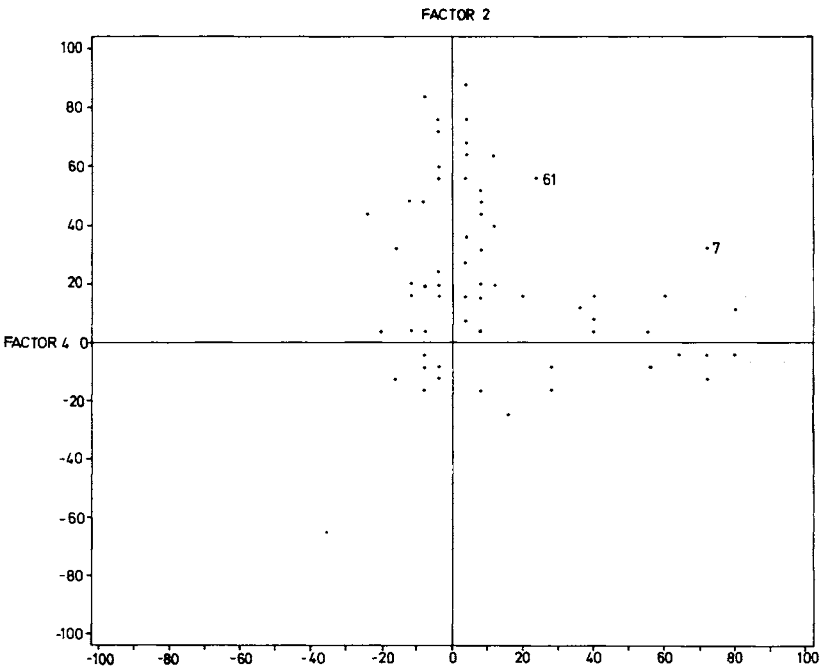

FIGURE 11—Plot of the loadings of the zones on factor four against the loadings on factor two

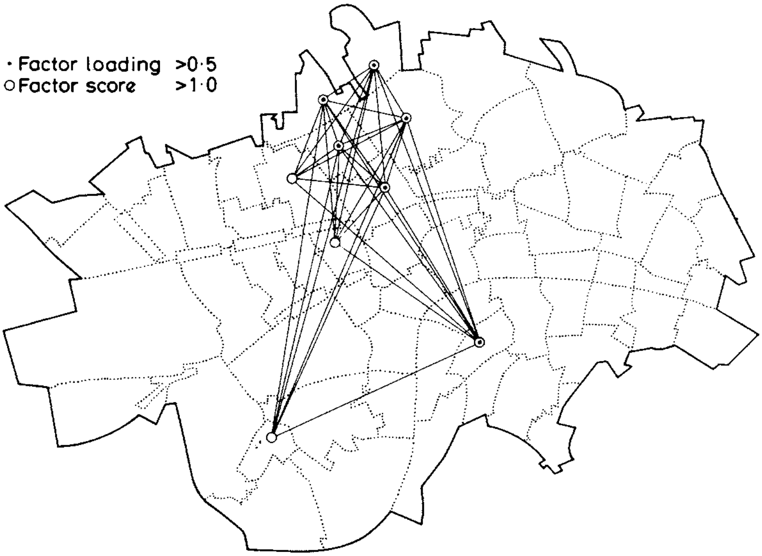

FIGURE 12—Factor five: Bloomsbury

Factor five (Fig. 12), with 8 per cent of the variance, has been termed a Bloomsbury system. The pattern is similar to that of the Westminster system (factor two) in that it includes a number of main-line termini as an integral part. These are Waterloo, Victoria, Euston, and King's Cross/St. Pancras. Again this area has a unique combination of activities. Besides the University, Bloomsbury has a large number of trade union offices whose London location is partly a reflection of their need for easy access from the rest of the country.

Factor six has a complex structure and has not been mapped. Experience has shown that it is necessary to rotate one more than the number of common factors identified by examination of the eigenvalues in order to obtain a meaningful composition in the last factor and this has not been done.

Grouping Analysis

When combined, the various sub-systems represent a functional regionalization of Central London on the basis of a single measure of internal linkage. However, amalgamating all the maps would fail to satisfy the basic criterion of any regionalization or classification scheme, namely that each observation should fall into a separate group. This is because several zones belong to more than one sub-system. Furthermore, a certain degree of subjectivity has been introduced into the maps by specifying arbitrary limits for factor scores and loadings.

To obtain a more objective regionalization in which zones fall into discrete classes, grouping techniques have been applied to the similarities between zones according to their scores on the six factors; zones have therefore been combined on the basis of similarities in the way they distribute their taxi trips to give a regionalization with a specifiable level of efficiency.

Multi-dimensional analysis permits exact measurement of the degree of similarity between zones with respect to all six patterns of variation simultaneously. The position of each zone is

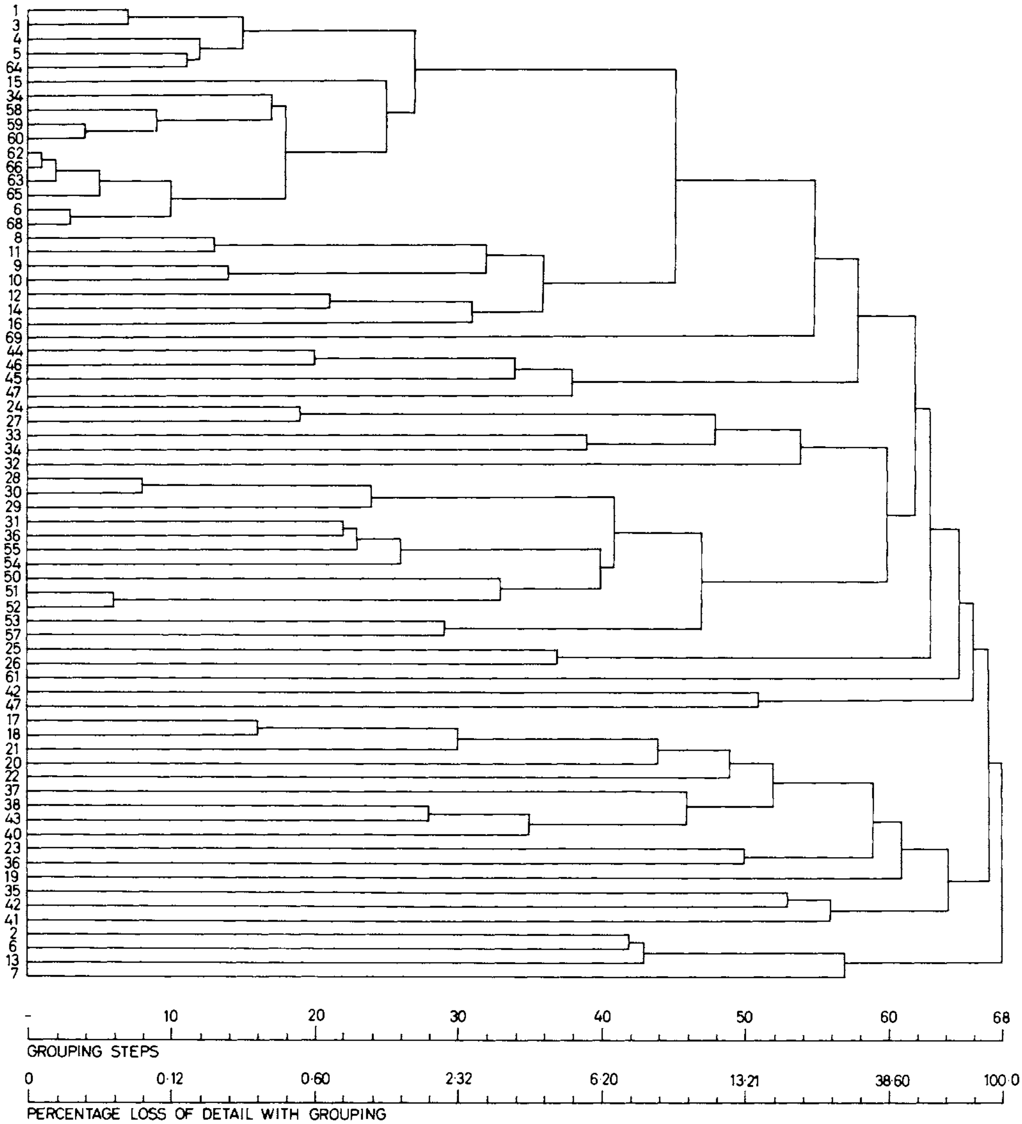

FIGURE 13—Linkage tree for contiguous grouping of traffic zones

defined in the orthogonal six-dimensional factor space by its score on all the factors. Similarity between each pair of zones is given by the euclidean distance separating them, which is calculated from the formula:

$$Dij = \sqrt{\sum_{m=1}^{n} (Smi - Smj)^2}$$

where s = factor score; m = factor; i = ith zone; j = jth zone; n = 6 (the number of factors).

This formula was used to compute a 69×69 inter-zone similarity matrix, in which the lowest distance indicates the two most similar zones. A stepwise grouping procedure was then applied to this matrix. Starting from sixty-nine single member groups, two groups (which might be either single or multi-membered) were combined, at each step, to form a new group, subject to the constraint that each new group made the least possible increment to the total within group distance.[25] In this way zones are combined into types of trip-generating area; there is, however, no guarantee that these will be uniform contiguous regions. To ensure this, a contiguity constraint has to be imposed on the grouping process whereby zones are combined only if they, or one member of the group, are contiguous.[26]

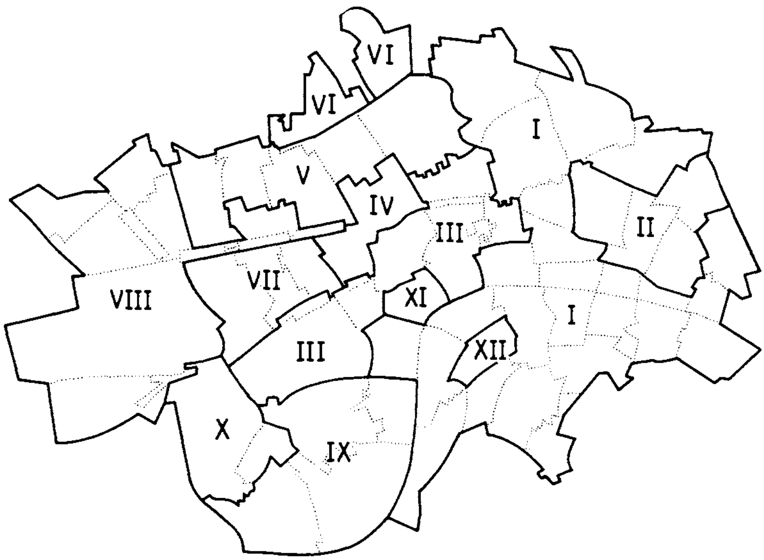

FIGURE 14—Contiguous grouping of origin zones on the basis of total flows (step 57)

The steps in the grouping process are described in a hierarchical linkage diagram (Fig. 13) which shows how the sixty-nine initial zones are combined ultimately to form one group. At each step the total amount of within-group variance can be specified. As the zones are combined into larger groups, an increasing amount of the initial detail is lost and the groups become more and more heterogeneous. The total within-group variance (in this case distance) therefore indicates the efficiency (or conversely the amount of information lost) for each stage in the grouping.

It is possible to break into the grouping at any level of efficiency or to specify a certain number of regions (steps) and then map the results. Inevitably an element of subjectivity enters in the choice of any particular level.[27] Bearing this fact in mind, Figure 14 represents a contiguous grouping of the originating zones into twelve regions. By so reducing the number of zones from sixty-nine to twelve, only 25 per cent of the initial detail has been lost; the regionalization is therefore 75 per cent efficient.

The largest region extends from Finsbury through the western part of the City to include all zones south of the river (except Waterloo) and also Whitehall (zone 39). The feature common to zones in this region is that they do not score high on any of the factors. Examination of the linkage tree suggests that this large region does contain clusters of zones or sub-regions that emerge at earlier stages in the grouping before being combined into one region. By step 25, zones 1, 3, 4, 5, and 15 form a separate group corresponding to the warehousing periphery of the City; at step 32, the zones to the north of the City (zones 8–11) and in the western half of the City, outside the financial nucleus (zones 12, 14 and 16) are separately distinguished from the larger group to the south of the river.

In contrast to this very large region with weak connections to the rest of the system, Waterloo (region 12), Charing Cross (region 11), Euston and King's Cross/St. Pancras (region 6),[28] do not join any groups. This is because these termini, unlike most other zones, are strongly associated with all the sub-systems that have been identified. Region 10 likewise indicates that Victoria has similarities only with near-by Belgravia.

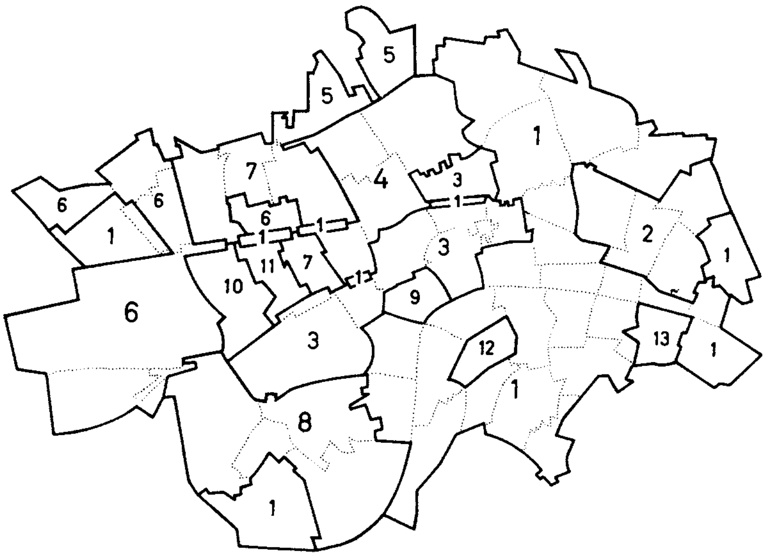

FIGURE 15—Non-contiguous grouping of origin zones on the basis of total flows (step 56)

The stage of the grouping mapped in Figure 14 picks out the distinctive nature of the eastern part of the City (region 2) and Victoria (region 9). Mayfair (region 7), in the context of the earlier sub-system maps, is surprisingly separate from the area to the north of Oxford Street (region 5), which would appear to have a greater similarity with Bloomsbury in terms of its trip distribution characteristics. The Oxford Street zones combine with a region in the far west of the Central Area, including Hyde Park and the areas to the north and south of it (region 8). The 'mid-town' sub-system (factor three) is isolated in origin from region four, which represents a combination of a St. James's group (zones 37, 40 and 43) in the west and a Fleet Street sub-group (zones 17, 18, 20, 21 and 22) in the east that occurs at step 52. The Soho area (region 4) does not join this general mid-town region until step 59.

In order to assess the degree of functional integration within central London, the grouping analysis was repeated without a contiguity constraint (Fig. 15). Zones were therefore combined solely in terms of similarities in trip distribution patterns, producing a typology of zones and not necessarily a uniform regionalization. A comparison of Figures 14 and 15 shows a considerable similarity between the alternative groupings. Area types one, two and three correspond closely to regions one, two and three. Again, the stations do not combine with any group, with the exception of Victoria which combines with its adjacent zone. The non-contiguous grouping is more realistic in that it distinguishes Bloomsbury (type 4) from the area to the north of Oxford Street (type 7). The Mayfair zones (types 10 and 11), that form part of region seven in the contiguous grouping, do not enter into any combination when no constraint is imposed; being the largest generators of trips, these zones exhibit the greatest variance in their trip distribution patterns. By adding an additional step to this grouping, types 12 and 13 combine, making this non-contiguous grouping 79 per cent efficient, 4 per cent more than the contiguous case at the same step. (The mapped typology, step 56, represents only an 18 per cent loss of information.) The fact that this unconstrained grouping yields a series of homogeneous and contiguous regions suggests a remarkable rationalization of space in terms of its functional organization.

Alternative Approaches

One of the characteristics of the taxi flow system is the great variation between zones in the total number of trips generated. In the study of linkage patterns it is useful to reduce this source of variation. This can be done by expressing movements from each origin to each destination as a proportion of that destination's total number of incoming trips. The raw data matrix is therefore standardized by columns. Correlations between columns reveal similarities in the relative rather than the absolute importance of each origin to each destination. Factoring isolates groups of destinations (indicated by the factor loadings) and the principal origins for each of these groups (high factor scores).

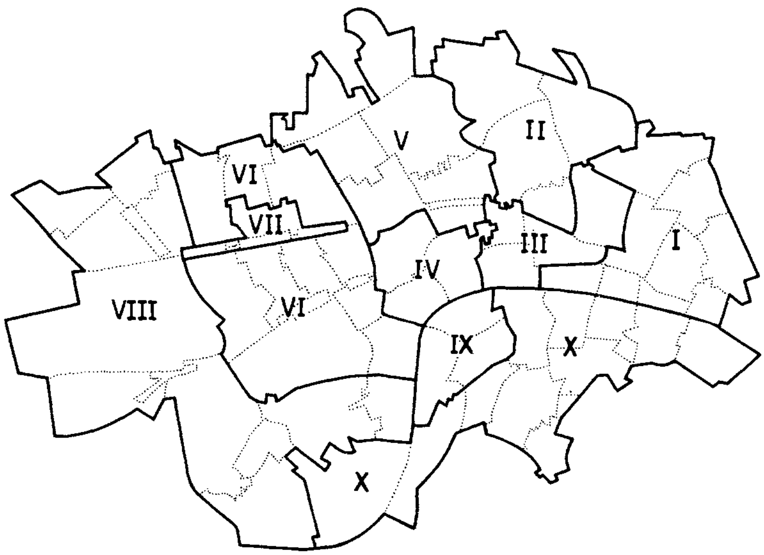

FIGURE 16—Contiguous grouping of origin zones on the basis of percentage flows (step 59)

Six factors were extracted from such a standardized matrix; together these accounted for 49 per cent of the variation, compared with 60 per cent with six factors from the untransformed matrix.[29] Grouping of origins with contiguity constraint gave the regional structure described in Figure 16, which at step 59 is 64 per cent efficient.

A good deal of correspondence exists between the regionalizations based upon total and percentage flows. However, the elimination of absolute differences in the number of trips generated means that zones in the West End unite into one region (region 6), though Oxford Street again forms a separate region (region 7). Euston and King's Cross/St. Pancras link into the Bloomsbury region (region 4) for similar reasons. Again an eastern City region appears, while the area of low trip generation is broken into three regions—in Finsbury (region 2), the western part of the City (region 3) and Southwark (region 10). A South Bank region around Waterloo is isolated from the rest of the zones south of the river (region 9). In fact, a general comparison of this functional regionalization with patterns of structural differentiation suggests that the standardization had been more successful in revealing the detailed linkages that the formal differences call forth.

Conclusion

Although this analysis is based on only one indicator of linkage, it nevertheless suggests that the Central Area of London contains a number of remarkably self-contained functional regions with strong internal bonds. Since data are not available on different purpose trips, it has not been possible to match this pattern statistically with a uniform regionalization and so test the basic isomorphism of formal and functional regions suggested by a general field theory. However, the broad correspondence between the pattern of functional regions produced in this analysis and familiar structural divisions of Central London, does lend support to the view that the Central Area contains a number of specialized sub-districts in which closely related activities congregate and that these districts are maintained by a clear pattern of internal circulation.

In view of the particular characteristics of taxi travel, it might be contended that the linkage patterns suggested by other modes of communication would indicate fundamentally different functional regions. Since it has not been the object of this paper to provide a detailed examination of taxi usage as such, it is only possible to postulate in general terms the likely effects of such factors as the frictional costs of distance, convenience and competing travel opportunities on the resultant flow patterns. In particular, the high costs of taxi travel might partly explain the relatively compact regions that emerge. Telephone and postal communications would not be influenced by distance in this way. Although no data are available on journey purpose, a large proportion of taxi trips are probably for the purpose of face-to-face business contacts and, in view of the fact that meetings are frequently preceded by telephone calls and followed by written memos, the patterns revealed by these latter flows should be broadly similar to those recorded here.

However, before any firm conclusions can be reached on the nature and pattern of functional linkages in the city centre, further studies are needed of the basic interdependencies that lie behind the movements of persons examined here. To this end, research is in progress which will attempt to ascertain the information flows connecting different types of office function within Central London. The particular value of the analysis of movement patterns presented in this paper is to indicate a method for investigating the structure of complex information flow systems which will reveal the principal groups of functions and the linkages they have in common.

NOTES

[1] B. J. L. BERRY, 'Interdependency of flows and spatial structure: a general field theory formulation' in 'Essays on commodity flows and the spatial structure of the Indian economy', *Univ. of Chicago, Dept. of Geography Res. Pap.* 111 (1966)

[2] B. J. L. BERRY, 'A synthesis of formal and functional regions using a general field theory of spatial behaviour' in *Spatial analysis* (Ed. B. J. L. BERRY and D. F. MARBLE)

[3] A. K. PHILBRICK, 'Principles of areal functional organization in regional human geography,' *Econ. Geogr.* 33 (1957), 299–336

[4] O. WÄRNERYD, *Interdependence in urban systems* (Gothenburg, 1968)

[5] J. RANNELS, *The core of the city* (New York, 1956)

[6] D. H. DAVIES, *Land use in central Capetown* (Capetown, 1965)

[7] J. B. GODDARD, 'Multivariate analysis of office location patterns in the city centre: a London example,' *Reg. Stud.* 2 (1968), 69–85

[8] J. B. GODDARD, 'The internal structure of London's Central Area' in *Urban core and inner city*, Proceedings of the International Studyweek, Amsterdam, September 1966 (Leiden, 1967)

[9] G. TÖRNQVIST, 'Flows of information and the location of economic activities', *Geogr. Annlr* 50 (1968), B, 99–107; B. THORNGREN, 'Regional economic interaction and flows of information', Paper presented to Regional Science Association seminar, Copenhagen (mimeo., 1967)

[10] *London County Council* (1964), *London traffic survey* 1, 154

[11] J. B. GODDARD, op. cit. (1967), 120

[12] E. M. HORWOOD and R. R. BOYCE, *Studies of the C.B.D. and urban freeway development* (Seattle, 1959)

[13] *London traffic survey* 1, 41. Data on other movement patterns were based on home interviews and are not reliable at the traffic-zone level.

[14] Because of the difficulties of illustrating all zone-to-zone desires, the major urban transport studies have made use of a method of accumulating these movements on a 'spider's web network' consisting only of links between adjacent zones. While this method helps in the cartographic display of flow patterns and is useful in traffic network analysis, it is of little value for a study of linkages which must be concerned with direct zone-to-zone connections.

[15] J. B. GODDARD, 'Changing office location patterns within Central London', *Urban Stud.* 4 (1968), 276–85

[16] E. M. HORWOOD and R. R. BOYCE, op. cit.

[17] J. NYSTUEN and M. F. DACEY, 'A graph theory interpretation of nodal regions', *Reg. Sci. Ass. Pap.* 7 (1961), 29–42

[18] W. L. GARRISON and D. F. MARBLE, 'A factor analytic study of the connectivity of a transport network', *Reg. Sci. Ass. Pap.* 12 (1964), 231–38; P. R. GOULD, 'On the geographical interpretation of eigenvalues', *Trans. Inst. Br. Geogr.* 42 (1967), 53–85; B. J. L. BERRY, op. cit. (1966); S. ILLERIS and O. PEDERSEN, 'Central places and functional regions in Denmark. Factor analysis of telephone traffic', *Lund Stud. Geogr.* (1968) Ser. B, 30

[19] A particular problem arises over the diagonal elements when correlating columns from a matrix containing flow information, since two unlike elements are compared, the flows within zone 1 being compared with the flows between zones 1 and 2. This source of error is very small, but cannot be eliminated.

[20] With 67 degrees of freedom an r value of greater than 0.38 is significant at the 0.001 probability level. It should be emphasized that the correlations only indicate similarities in the zonal profiles. Two zones might have similar profile variations but there might be considerable absolute differences in the total number of trips involved. The effects of this factor could be explored by repeating the analysis on a series of binary matrices for flows within specified ranges, for example, 501–750 trips per day, 751–1000 trips per day, etc.

[21] For a discussion of the methods, see P. R. GOULD op. cit.

[22] For a full statement of the methods of these two approaches see H. H. HARMAN, *Modern factor analysis* (Chicago, 1966), 136 ff; Varimax rotation of principal component axes is also discussed, p. 304 ff. See also J. B. GODDARD (1968) for an application and discussion of factor analysis in urban studies.

[23] The use of an initial principal components solution means that the rotated factor scores can be computed without recourse to estimation procedures using multiple regression (H. H. HARMAN, op. cit., 348)

[24] Oblique factors in geographical problems are discussed in J. IMBRIE, 'Factor and vector analysis programs for analysing geological data', *Office of Naval Research, Tech. Rep.* 6 (1963), Task No. 389–135, Contract 1228(26) (Northwestern University, Ill.) The orthogonality requirement is discussed in E. CASETTI, 'Classificatory and regional analysis by discriminant iterations', *Office of Naval Research, Tech. Rep.* 12 (1963)

[25] J. H. WARD, 'Hierarchical grouping to optimise an objective function', *J. Am. statist. Ass.* 58 (1963), 236–44. This grouping procedure does not guarantee an optimal partitioning of the set of zones. A closer approximation to an optimal grouping can be obtained by applying discriminant analysis to the classification at the required step. Cf. E.

CASETTI, op. cit., and Q. AHMAD, 'Indian cities: characteristics and correlates', *Univ. of Chicago, Dept. of Geography Res. Pap.* 102 (1965)

[26] For a full statement of the methods see N. SPENCE, 'Multi-factor uniform regionalisation of British counties on the basis of employment data for 1961', *Reg. Stud.* 2 (1968), 87–104.

[27] R. J. JOHNSTON, 'Choice in classification: the subjectivity of objective methods', *Ann. Ass. Am. Geogr.* 58 (1968), 575–89

[28] Euston and King's Cross/St. Pancras were operationally defined as contiguous.

[29] The skewed nature of the percentage distributions necessitated a log transformation.

RÉSUMÉ—*Régions fonctionnelles au sein de la cité: étude analytique factiorielle de l'écoulement des taxis dans le centre de Londres.* Cette communication examine le problème de la mesure des rapports entre la structure des mouvements et l'emplacement des activités au sein du centre de la cité. Elle insiste particulièrement sur le problème de la définition de l'analyse du système complexe de liens. Les données sur les écoulements de taxi sont présentées comme un indicateur unique de liens fonctionnels polyvalents au sein de Londres. L'analyse cartographique révèle une structure très complexe des écoulements. Au moyen de l'analyse des facteurs et de la corrélation, il s'agit donc d'établir sur la base des données la structure sous-jacente du système d'écoulement des taxis. Les zones de circulation sont comparées en fonction des zones d'où les courses proviennent; les facteurs de charge élevés révèlent des groupes de zones qui reçoivent des courses d'origines communes. L'indice de facteurs apparentes indique les origines. En reliant sur une carte des groupes d'origine et de destination, l'on voit se dégager cinq régions fonctionnelles. Un groupement hiérarchique des zones produit une classification objective des zones en fonction des analogies quant aux caractéristiques de liens. La comparaison entre la structure des régions fonctionnelles et celle des régions structurelles familières laisse entendre un isomorphisme de base entre de telles régionalisations. Ce qui est encore plus important, l'analyse révèle l'utilité des méthodes d'enquête sur la structure des systèmes complexes d'écoulement, afin de spécifier les principaux groupes d'observations et les liens qu'ils ont en communs.

ZUSAMMENFASSUNG—*Funktionelle Regionen innerhalb des Stadtzentrums: eine Faktorenanalyse der Taxibewegungen in Zentrallondon.* Diese Arbeit untersucht das Problem einer Messung des Verhältnisses zwischen Gesetzmässigkeiten der Bewegungen und der örtlichen Lage von Aktivitäten innerhalb des Stadtzentrums. Die Betonung liegt auf dem Problem der Definition und der Analyse komplexer Verbindungssysteme. Angaben über Taxibewegungen werden als Einzelanzeiger der vielseitigen funktionellen Verbindungen in Zentrallondon vorgelegt. Eine kartographische Analyse zeigt ein höchst komplexes Bewegungsmuster; die Angaben werden daher einer Korrelation und Faktorenanalyse unterworfen, um die zugrunde liegende Struktur des Systems der Taxibewegungen zu bestimmen. Verkehrszonen werden nach denjenigen Zonen miteinander verglichen, in denen die Taxis ihre Fahrten antreten; hohe Ladungsfaktoren zeigen Zonengruppen, in die Fahrten aus gemeinsamen Ausgangspunkten führen; die entsprechenden Faktorenergebnisse bestimmen diese Ausgangspunkte. Die Verbindung von Gruppen von Ausgangsorten und Zielpunkten auf einer Landkarte zeigt fünf funktionelle Regionen. Eine hierarchische Gruppierung von Zonen ergibt eine objektive Klassifizierung der Zonen nach Ähnlichkeiten in den Verbindungs-Charakteristiken. Ein Vergleich des Musters von funktionellen Regionen mit den vertrauten Strukturregionen lässt einen grundlegenden Isomorphismus zwischen solchen Regionalisierungen erkennen. Wichtiger ist, dass die Analyse die Brauchbarkeit der Methoden zur Untersuchung von komplexen Bewegungssystemen zeigt, wodurch die Hauptbeobachtungsgruppen und die ihnen gemeinsamen Verbindungen spezifiziert werden können.