Chapter 1
Geodemographics and Urban Planning Analysis: An Historical Review

Peter Batey

Abstract The focus in this chapter is on geodemographics—essentially the analysis of people by where they live—from the perspective of those involved in planning and policy-making. Like other analysts, they seek to distil the main sources of social and economic variation in cities. However, what distinguishes them is usually some practical purpose directly related to policy formulation, analysis, and evaluation. The chapter takes the form of an historical review in which four main themes are examined: (1) early efforts to apply rudimentary geodemographic classifications, in order to inform and influence policy; (2) harnessing advances in computing and multivariate statistics that make it possible to handle the large datasets needed to explore urban spatial structure; (3) pioneering applications of geodemographics to enable local authorities to identify multi-dimensional needs and to indicate priorities for spatial targeting of resources, and (4) geodemographics in action as an evaluation tool to measure the success of spatial targeting of area-based policy initiatives and identify potential improvements. In a final section, some aspects of the present status of geodemographics are considered and related to the findings of the historical review.

Keywords Geodemographics · History of planning methods · Spatial targeting · Surveys of London life and labour · Charles Booth

1.1 Introduction

Geodemographics is concerned with the classification of neighbourhoods into categories or clusters based on their socio-economic characteristics. In lay-person's terms, it can be said to be 'the analysis of people by where they live' (Harris et al. 2005, pp. 16–17). It uses a qualitative description—a 'pen portrait'—to summarise

P. Batey (✉)
University of Liverpool, Liverpool, UK
e-mail: pwjbatey@liverpool.ac.uk

the distinctive attributes of each category or cluster. It works on the principle that 'birds of a feather flock together': people who live close by (i.e. in the same neighbourhood) are assumed to have more in common than a random group of people. This is a well-established notion in human geography, commonly known as Tobler's first law of geography: 'everything is related to everything else, but near things are more related than distant things' (Tobler 1970, p. 236; Harris et al.[1] 2005, pp. 16–17). Furthermore, geodemographics works on the principle that people tend to align themselves with the behaviour and aspirations of the local communities in which they live (Alexiou et al. 2016, p. 382).

Urban planners and policy-makers have long had a practical interest in geodemographics, usually directly related to policy formulation, analysis, and evaluation. Typically, the aim is to develop a consistent and systematic approach to spatial resource allocation, involving the definition of priority areas to receive favoured treatment. Such areas may be defined in relation to particular policy sectors, such as education, housing, crime or health or, in a more general sense, as in the case of designating 'inner city areas'. The geodemographic classification here serves as a composite measure of need and is usually constructed using census data, where feasible supplemented by other sources of small area data. It generally takes the form of a map displaying the spatial distribution of neighbourhood types, together with a set of pen portraits.

Without doubt, geodemographic classifications have made a major contribution to the regional scientist's toolkit since their introduction more than 50 years ago. This applies especially to those whose background training is in the quantitative branches of geography, planning, and sociology but not, it must be said, economics. Over the years, geodemographics has retained its strong empirical focus and has largely operated as a separate sub-field of applied regional science. Unlike many of their regional scientist colleagues, the proponents of geodemographics seem not to have chosen to engage with recent and current theoretical debates in regional economics and the burgeoning field of the New Economic Geography, preferring to concentrate on more technical matters at a fine level of spatial detail. However, as will be demonstrated later in this chapter, there are encouraging signs of new developments in *integrated* analysis, bringing together geodemographics with other forms of regional analysis, such as spatial interaction modelling.

In what follows, the emphasis will be on the public sector: how can and does geodemographics support the development of public policy in general and urban planning in particular? The chapter will take the form of an historical review, examining key developments in geodemographics through the medium of a series of themes, beginning with Charles Booth's pioneering street-by-street surveys and poverty maps of the late nineteenth century. As Singleton and Spielman (2013)

---

[1]Harris et al. (2005, p. 17) suggest that Tobler's Law should be modified when viewed in a geodemographic context: not only are nearby populations related but so too are populations living in the same type of neighbourhood. In other words, near and far things are related—by neighbourhood type.

observe, in the US most applications of geodemographics take place in a commercial environment and involve proprietary geodemographic classifications. The situation in the UK is different, largely thanks to a long tradition of making classifications accessible to public sector and academic users at very little cost.

Private sector applications, including retail planning, locational analysis and market segmentation, are a vast topic in themselves, and to cover them adequately would be well beyond the scope of the present chapter. They are, however, documented in the key geodemographics texts: see, for example, Webber and Burrows (2018); Harris et al. (2005); and Leventhal (2016). These texts also provide accounts of applications elsewhere in the world.

The chapter is organised into four main themes:

- *Precursors of geodemographics*: early efforts to apply rudimentary geodemographic classifications, in order to inform and influence policy.
- *Exploring urban spatial structure*: harnessing advances in computing and multivariate statistics that make it possible to handle the large datasets needed to explore urban spatial structure.
- *Pioneering geodemographic classifications* enabling local authorities to identify multi-dimensional needs and to indicate priorities for spatial targeting of resources
- *Geodemographics and the evaluation of spatial targeting:* geodemographics in action as an evaluation tool to measure the success of spatial targeting of area-based policy initiatives.

Interspersed in the text is a series of discussions that draw together some of the main points that emerge from the review, while in the final section there is an attempt to relate the historical advances covered in the chapter to a number of recent and current developments in geodemographics.

1.2 Precursors of Geodemographics

1.2.1 Charles Booth's Descriptive Map of London Poverty 1889

Charles Booth's *Descriptive Map of London 1889*, first published in 1891, is generally regarded as the earliest antecedent of geodemographic classifications. Booth,[2] a wealthy ship-owner and businessman from Liverpool, was also an energetic social reformer deeply committed to finding out the full extent of poverty in

---

[2] See Pfautz (1967, pp. 43–46) for a detailed account of Booth's life, character, and the motivation that lay behind his prodigious efforts to understand and campaign on poverty among London's population; see also Morgan (2019, pp. 40–41).

London. This was just one part of his broader lifetime mission to foster a deeper understanding of the origins of poverty in urban Britain.

His privately funded city-wide enquiry aimed to discover how many of London's residents were living in poverty, what kept them in that state and what might be done to alleviate it (Vaughan 2018, p. 69). In addition to his study of poverty, Booth made a series of detailed studies on the working conditions in the principal London industries. This ended with an inquiry into religious influences, including interviews with clergy to ascertain church attendance (Vaughan 2018, p. 70). The results of his research, including his poverty map, were published in a series of 17 volumes under the general heading of the *Life and Labour of the People of London* over the period 1886–1903.[3]

Booth was an empiricist who believed in collecting evidence to gain political support for a more systematic approach to the elimination of poverty than was being provided by the 'sporadic and untargeted efforts of the charitable classes' (Webber and Burrows 2018, p. 32). He was a member of the official committee in charge of the 1891 Census which suggests that he clearly understood the value of systematic data collection, even though he had no formal education in statistics (Webber and Burrows 2018, p. 32).

Booth began his study of poverty at Tower Hamlets in London's East End, in 1887, extending the inquiry a year later to include the people of East London and Hackney. Reaction from the press and the public was favourable and this gave Booth the confidence to turn his attention to gathering data on the rest of the city. His plans for the London-wide survey were ambitious. The outline of each street in London was carefully shaded on a 6″ to the mile base map to indicate the general socio-economic condition of the residents. The basis of the classification was the reports of the school board visitors (SBVs) to households in each street. These reports contained detailed records compiled from continuous home visits, of every family with children of school age, (Vaughan 2018, p. 70). The SBVs had been established as a result of the Compulsory Education Act of 1877 as a means of tracking the children of the poor in order to ensure they were receiving an adequate education. Each SBV kept a 'detailed record of every poor family in his district, noting such details as the occupation, his income, the number, ages, and sexes of the children, the parents' habits of sobriety, the cleanliness of the household, and so on' (Selvin and Bernert 1985, p. 73).

Booth preferred to interview SBVs rather than household members on the grounds that a direct interview would have been considered an 'invasion of privacy'. Each SBV was interviewed for 20–30 h based upon the contents of their notes and record books (Bales 1991). After carefully checking their returns, Booth personally inspected each neighbourhood covered by the SBV, and checked his findings with data from the census whose collection he also oversaw. It is significant that people

---

[3]The work is divided into three parts: the Poverty series (four volumes); the Industry series (five volumes); and the Religion series (seven volumes). The final volume is a conclusion to the whole work.

Table 1.1 Booth's classification of streets in London by general condition of inhabitants

Colour code	Description (or 'pen portrait')	Class	Description (or 'pen portrait')
Black	The lowest grade (corresponding to Class A), inhabited principally by occasional labourers, loafers and semi-criminals—the elements of disorder	A	The lowest class—occasional labourers, loafers and semi-criminals
Dark blue	Very poor (corresponding to Class B), inhabited principally by casual labourers and others living from hand to mouth	B	The very poor—casual labour, hand-to-mouth existence, chronic want
Light blue	Standard poverty (corresponding to Classes C and D) inhabited principally by those whose earnings are small (say 18 s (shillings—a unit a currency) to 21 s a week for moderate family), whether they are so because of irregularity of work (C) or because of a low rate of pay (D)	C and D	The poor—including alike those whose earnings are small, because of irregularity of employment, and those whose work, though regular, is ill-paid
Purple	Street mixed with poverty (usually C and D with E and F, but including Class B in many cases)	E and F	The regularly employed and fairly paid working class of all grades
Pink	Working class comfort (corresponding to Classes E and F, but containing also a large proportion of the lower middle class of small tradesman and Class G). These people usually keep no servant	G and H	Lower and upper middle class and all above this level
Red	Well-to-do; inhabited by middle-class families who keep one or two servants		
Yellow	Wealthy; hardly found in East London and little found in South London; inhabited by families who keep three or more servants and whose houses are rated at £100 or more		

Source: Harris et al. (2005)

living and working in an area were asked about the *neighbourhood* they lived in, not their own *personal* circumstances (Bales 1991).

These data were then used to place each household into one of eight mutually exclusive and exhaustive categories of 'class', rank ordered from A to H in ascending order of status. The eight were then grouped into five higher order categories (Webber and Burrows 2018, p. 32; Pfautz 1967, p. 91) again ordered by social status, as shown in Table 1.1.

Table 1.1 describes each of the classes by means of a pen portrait, a vivid description of the conditions experienced by London's residents. The idea of using pen portraits was novel at the time but proved very effective when used in

conjunction with the maps emerging from the street survey. It has stood the test of time and still today forms an important component of a geodemographic classification.

The streets covered by the survey were mapped using the colour scheme described in Table 1.1. Booth recognised that the spatial distribution of these different classes of household was far from random. Although the pattern was not entirely uniform, for the most part households in similar classes tended to be 'clustered' in close spatial proximity to each other (Webber and Burrows 2018, p. 33). Booth was 'the first, and by no means the last, to use colour to indicate the locations where distinct categories of household lived' (Vaughan 2018, p. 71). Moreover, in mapping streets he permitted some streets to be assigned to more than one class, representing streets of a mixed socio-economic character, an idea which caught the attention of analysts much later when commercial geodemographic classifications[4] were being designed in the 1980s.

The map itself was published in 1891, in 12 sheets, as the *Descriptive Map of London Poverty 1889*. Figure 1.1 shows two extracts from the poverty map, the first centred on Bloomsbury, an affluent area of central London, and the other focusing on Lincolns Inn Fields not far away, a more mixed area. What is striking is that both areas contain such a variety of socio-economic conditions with many instances of the poorest living cheek by jowl with some of those most well off.

The map of poverty and the survey results was widely disseminated and proved very effective in drawing attention to the scale of poverty experienced by London's residents.

Discussion 1: The Wider Influence of Booth's London Survey and the Prospect for Follow-ups

Charles Booth's work in compiling the *Life and Labour of the People of London* proved to be very influential and led to similar studies being carried out elsewhere, especially in the UK and the USA. Notable examples, referred to by Vaughan (2018, pp. 92–128), are Rowntree's studies of York; Hull-House in Chicago; and Du Bois' map of the seventh ward in Philadelphia. Of particular interest here is the *New Survey of London Life and Labour*, 1928–35, conducted under the leadership of Hubert Llewellyn Smith, one of Booth's former assistants. Like the earlier survey, the New Survey focused on poverty and was intended to make comparisons with the earlier *Life and Labour* survey, 40 years on. It aimed to replicate the methods used in Booth's survey.

The *New Survey* was well received, particularly in its examination of urban change. It showed a general rise in income, with a shorter working day and improved literacy and more money to spend in increased leisure time. However, there were still substantial numbers of people continuing to live in poverty (Alexander 2007).

Maps were just as important in the *New Survey* as they had been in the first survey. The study area was covered by six sheets at $4''$ to the mile and Booth's colour

---

[4] Such classifications became known as fuzzy geodemographics, in view of the overlapping cluster boundaries. See Flowerdew and Leventhal (1998).
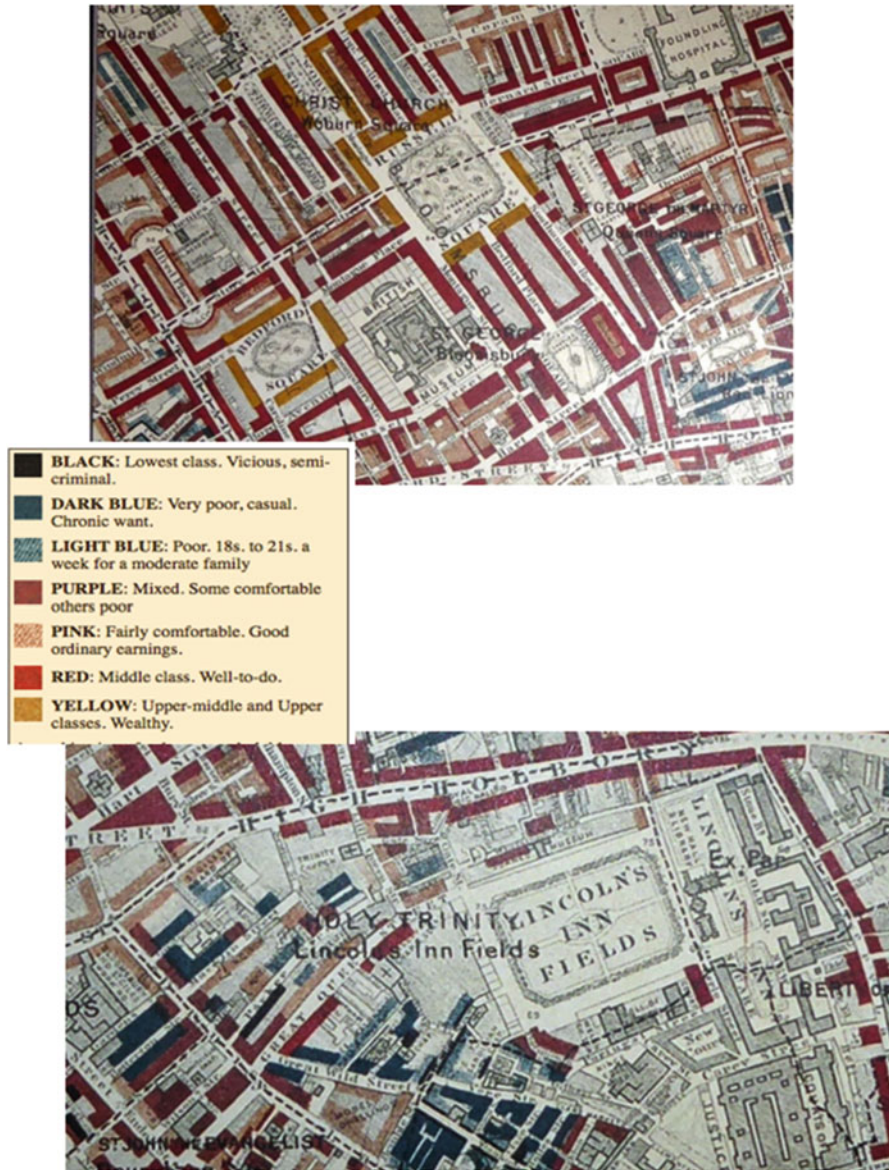
BLACK: Lowest class. Vicious, semi-criminal.

DARK BLUE: Very poor, casual. Chronic want.

LIGHT BLUE: Poor. 18s. to 21s. a week for a moderate family

PURPLE: Mixed. Some comfortable others poor

PINK: Fairly comfortable. Good ordinary earnings.

RED: Middle class. Well-to-do.

YELLOW: Upper-middle and Upper classes. Wealthy.

Fig. 1.1 Two extracts from Charles Booth's Descriptive Map of London Poverty 1889: Lincoln's Inn Fields and Bloomsbury

scheme was largely repeated. In the nine volumes that accompanied the maps, there is a huge amount of detailed analysis of the spatial structure of the city pinpointing areas of continuity and change. While poverty was now more dispersed, it remained entrenched in some areas (Vaughan 2018, pp. 115–125).

By the late 1950s there was a proposal to undertake a third survey. Emanating from the newly established Centre for Urban Studies based at University College London, the *Third Survey of London Life and Labour*, would again make a detailed analysis of residential areas, but this time making extensive use of small area data from the 1961 Census of Population. As will be seen later in this chapter, the Centre was in the forefront of developments in census analysis at that time, work that led on to a range of applications of geodemographics in the public sector, from the late 1960s onwards.

1.2.2   Carter Goodrich and the Plane of Living

The second example has been chosen partly because it illustrates geodemographic analysis at a different spatial scale. Whereas Booth was concerned with the fine detail of poverty in London's streets, in this American example the spatial unit of analysis is the county and the map in question refers to the 3000 plus counties in the whole country.

The subject of interest here is the so-called *Plane of Living*, a concept first developed during the 1930s as a means of characterising levels of living across the entire USA by Carter Goodrich et al. based at the University of Pennsylvania. This work, largely forgotten over the years, has been rediscovered quite recently by Carruthers and Mulligan (2008). The aim was to devise and apply a rough measure, by small geographical units, that would enable comparisons of the level of prosperity in various parts of the country immediately before the onset of the Great Depression. Which were the areas where the standard of living was low before 1929? Did people succeed in moving from the worse to the better areas? Did those who left the country for the city gain by moving? (Goodrich et al. 1935, p. 14). To be able to judge this, it was important to measure all parts of the country in a consistent manner, recognising the difficulties of finding indices that were equally applicable to both urban and rural areas. Hence the need for a careful comparison of possible measures before a final selection could be made.

The original *Plane of Living* map was prepared by Warren Thornthwaite[5] on behalf of Goodrich's team and is reproduced in Fig. 1.2. The map displays a composite index of three variables that reflects, as a percentage of the national average: (1) household income; (2) the proportion of homes having radios; and (3) the proportion of homes having telephones, equally weighted.

The *Plane of Living* data used here, expressed at county level, refer to the period 1928–1929, immediately before the Great Depression. Overlaid on the map are mining and manufacturing areas, a particular concern of Goodrich's research, as

---

[5] At this time, Thornthwaite was working as an urban geographer. Later in his career, he was to become a renowned climatologist, known for a climate classification still in use more than 70 years after it was first conceived.
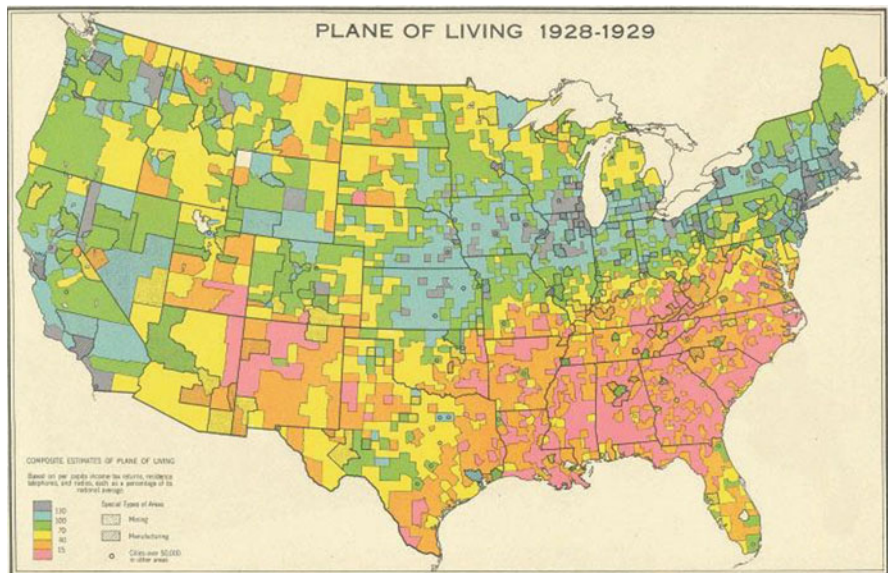
Fig. 1.2 Planes of Living in the United States 1928–1929 (Source: Goodrich et al. (1935))

well as towns with more than 50,000 population. A noteworthy feature is the concentration of low *Plane of Living* scores in the rural South, a reflection of the deficiency in 'those attributes of a modern standard of life' (Hoover 1948, p. 204). Only in a limited number of *urban areas* of the South are these attributes to be found. Here and elsewhere in the USA, the research team carried out a remarkably thorough analysis of the *Plane of Living* results and their implications for public policy.

At the time of Goodrich's research project, policy-makers wanted to understand how the distribution of the population had evolved in the period leading up to the Great Depression—and, going forward, how to influence migration flows in a way that enhanced economic opportunity and personal well-being (Goodrich 1936). It was vital that the analysis extended to the entire country rather than particular localities.

Discussion 2: **Connecting Place-to-Place Variations in the Quality of Life to the Greater Economic Landscape**
Carruthers and Mulligan viewed the *Plane of Living* map as one of the earliest examples of what would now be regarded as geodemographics. They could see that it had a valuable role in explicitly connecting place-to-place variation in the quality of life to the greater economic landscape (Greenwood and Hunt 2003). Considered by Carruthers and Mulligan as exceptionally innovative for its time, the work of Goodrich and his team 'helped establish an enduring framework wherein living conditions are viewed as fundamental to a wide array of socio-economic processes and outcomes' (Carruthers and Mulligan 2008, p. 2). Carruthers and Mulligan proceeded to apply the *Plane of Living* concept in their own research examining

the wider impacts of the Financial Crisis of 2008 and the possible role of public policy interventions.

1.3 Exploring Urban Spatial Structure[6]

1.3.1 United States

The task of identifying urban spatial structure in the USA has generally focused on specific cities. Among the best-known work is that of the Chicago urban sociologists Park and Burgess (Park et al. 1925). They used empirical urban research to develop and test concepts about the form, structure, and processes of development operating within cities. Park's work on defining 'natural areas' in cities (Light 2009, pp. 12–15)—'geographical units distinguished both by physical individuality and by the social, economic and cultural characteristics of the population' (Gittus 1964, p. 6)—typified work in a field which subsequently became known as human ecology (Theodorsen 1961; Light 2009, p. 7).

Early attempts at 'within city' classification, particularly those which involved the definition of natural areas, generally lacked methodological rigour. It is not clear how the various classification criteria (social, housing, ethnicity, etc.) were combined, nor was it evident as to which classification method was used. Despite these shortcomings, natural areas, once defined, remained in use as a summary device for reporting census and local statistics. Rees (1972) quotes the example of the *Local Community Factbook* for the Chicago Metropolitan Area which in 1960 was still using a city-wide application of the natural area concept in which 75 community areas defined 30 years earlier were employed as basic statistical units. Such areas had been classified according to a vaguely specified combination of historical, social, physical, commercial, and transportation criteria (Kitagawa and Taeuber 1963).

An undoubted stimulus to research in human ecology was the availability, for an increasing number of cities in the USA, of tabulations of data for census tracts, each with a population of about 4000. Census tracts had been introduced in 1910 when the US Bureau of the Census agreed to prepare tabulations of such areas as New York, Baltimore, Boston, Chicago, Cleveland, Philadelphia, Pittsburgh, and St Louis. Over the years, the number of tracted areas grew rapidly so that by the time of the 1960 Census, there were as many as 180 tracted areas, of which 136 were entire Standard Metropolitan Statistical Areas (Robson 1969, p. 42). Local advisory committees helped in the definition of tracts and where possible boundaries were drawn to follow permanent recognisable lines and to contain people of similar racial and economic status and areas of similar housing.

Notable among the studies that made extensive use of census tract data were those of Shevky and Williams (1949) and Shevky and Bell (1955) for Los Angeles and

---

[6]This section is a revised and expanded version of a section in Batey and Brown (1995).

San Francisco. They classified the census tracts of those cities into a number of classes that, because of their geographical proximity, were called social areas. The form of analysis was referred to as 'social area analysis'[7] and centred around three theoretical constructs: economic status, family status, and ethnic status. Shevky and his co-workers proposed three indices, one per construct, made up from one to three census variables, to measure the status of census tract population on scales of economic, family, and ethnic status, and to enable tracts to be classified on the basis of their scores on the indices (Berry and Horton 1970, p. 314). Social area analysis thus used classification criteria unique to each particular case study, which meant that the original analysis was incapable of being replicated by research workers in other cities. Rees (1972) includes a comprehensive bibliography of studies carried out in the USA and elsewhere following the principles set out by Shevky, Williams, and Bell.

Social area analysis was used to perform a variety of functions: to delineate socially homogeneous sub-areas within the city; to compare the distribution of such areas at two or more points in time; and compare the social areas in two or more places; and to provide a sampling framework; to enable other types of research to be undertaken, particularly for the design and execution of behavioural field studies (Rees 1972, p. 275).

In its original form social area analysis was severely criticised on two counts: first in terms of its theoretical basis (the theory underlying the constructs); and secondly for empirical reasons (the method of measuring the constructs).

Efforts were made subsequently to test the correctness of the census variables used to measure the constructs by employing factor analysis (Bell 1955). This work had some initial success, but extension to a wider range of cities revealed the shortcomings of the original choice of census variables (van Arsdol et al. 1958). It led to the inclusion of a wider range of socio-economic census variables and to the adoption of factor analysis (or the related technique of principal component analysis) as a standard method for identifying the underlying dimensions of urban social and spatial structure. This development of social area analysis became known as factorial ecology and was widely used by quantitative geographers in the 1960s and 1970s, not only in the USA but also in a range of cities throughout the world (Rees 1972; Berry and Horton 1970). Factorial ecology generally led to the production of maps and cross-sections using factor scores for each of the main factors. In this way it was possible to summarise the main features of spatial variation in socio-economic and demographic characteristics.

In some instances, the scores from the two factors were used to cross-classify census tracts. Rees's study of Chicago (Berry and Horton 1970), for example, employs a simple graphical technique to categorise areas according to the economic status of their residents. It was uncommon at this time to proceed one step further and

---

[7]Over time the term 'social area analysis' has been used in a number of different ways. For example, later in this chapter, the term will be used again to refer to work carried out in Liverpool in the late 1960s and 1970s.

use cluster analysis to create a multivariate classification of social areas. One exception, Tryon's (1955) study of the San Francisco Bay Area, was of limited value because of the imprecise way in which cluster analysis was used. Other researchers found it difficult to reproduce the results that Tryon had obtained (Robson 1969, p. 51).

1.3.2 United Kingdom

In Britain, early studies of urban spatial structure were hampered by the almost complete absence of small area census data. For many years the smallest units for which census data was published were the ward and civil parish and even here the range of information was small. Gittus (1964) describes attempts made in 1951 to define zones within the major conurbations that were relatively uniform with respect to the siting of industry and commerce, the rate of population change, and the age and type of housing. Within zones, distinctive areas, both natural and planned, were recognised and their boundaries determined on the basis of 'purely local considerations' (Gittus 1964, p. 9). These divisions and sub-divisions were intended to provide a more rational basis for presenting social data than that offered by administrative boundaries. However, it proved difficult to achieve consistency from one conurbation to another and in practice little use was made of the zones in comparative studies.

A more promising initiative was the establishment of the Inter-University Census Tract Committee. This committee, formed in Oxford in 1955, was originally intended to consider the definition of census tracts similar to those used in the USA. The city of Oxford served as the prototype for British census tracts and some 48 tracts were delineated with an average population of 2645 (Robson 1969, p. 44). Although these census tracts were similar to their American equivalents, they were nevertheless fairly large aggregates, likely to exhibit a high degree of internal heterogeneity. One possible advantage compared with other geographical units was that they were more certain of retaining their boundaries over time, allowing comparisons to be made.

However, the British Registrar General's Department had a different idea. Instead of adopting census tracts, it would make data available by enumeration district.[8] Such units were considerably smaller, containing on average less than 1000 people. Data on this scale were purchased for most of the conurbations included in the 1951 scheme and for a smaller number of smaller administrative areas. Members of the Inter-University Committee continued to meet and began to develop a series of

---

[8]An enumeration district was the area covered by a single census enumerator, distributing and collecting census forms.

comparative studies of urban structure. Gittus (1964) used 1951 Census data for sub-divisions[9] of the Merseyside and South East Lancashire conurbations in an experimental project, applying correlation analysis and principal component analysis to a set of 27 census variables. This preliminary work paved the way for further studies using 1961 Census enumeration district data, including Gittus's study of South Hampshire and Merseyside (Gittus 1963–1964) and Robson's study of Sunderland (Robson 1969; Robson 1984, pp. 110–112).

The work of the Centre for Urban Studies at University College London, under the direction of urban sociologist Ruth Glass, was probably the most important in terms of the development of method, scale of study and influence upon other research. Founded in 1958, one of its early projects was a pioneering inter-urban study of British towns. This study, carried out by two researchers at the Centre, Claus Moser and Wolf Scott, drew upon 60 variables for 157 towns in England and Wales with a population of more than 50,000 and classified them into 14 groups, using principal component analysis and a graphical plot of scores from the first two components (Moser and Scott 1961). The study was based on a combination of 1951 Census data and other sources of social and health data, while other variables were intended to measure change over time. Undoubtedly, the British Towns study was a remarkable achievement given the size of the data set and the limits of computing power available at the time. It provided the stimulus for much of the work in the UK on what was to become known as 'geodemographics'.

The Centre's original research programme included plans to undertake a *Third Survey of London Life and Labour*, intended to carry on some aspects of Charles Booth's *Life and Labour of the People in London* (1886–1903) and its follow-up the *New Survey of London Life and Labour* (1930–35), both of which were referred to earlier. The *Third London Survey* was seen as an opportunity to study how London had changed over time. And, rather than focus largely on poverty, as the earlier surveys had done, the scope would be wider, including London's economy, society, and culture. Like its two predecessors, the *Third Survey* would not be a purely academic exercise but where possible would use the survey results to influence social policy (Glass 1963, p. 181).

The Centre's research plan envisaged that four volumes would be published, covering: (1) Thirty Years of Change (essays on the main features and trends of change); (2) The Socio-Geographical Pattern (mainly the report on the analysis of special census tabulations; (3) The Diverse London (studies of particular areas, groups, problems, and aspects); (4) Maps and Sources (Descriptive material, as well as detailed tables relevant to the other three volumes). The second of the proposed volumes is of particular relevance here and reflected the Centre's:

---

[9]Sub-divisions consisted of a mix of smaller local authority areas and sub-divided larger local authority areas. They were used extensively by local planning authorities at that time to analyse population and employment change.

> . . . interest in developing large scale comparative analyses of urban patterns, both of intra-
> urban and inter-urban classifications—which make it necessary to identify types of towns
> and of urban components (Glass 1963 p. 182).

Here the intention was to make extensive use of the 1961 Census in the expectation that enumeration district data would become available early on in the Survey. Some 7000 enumeration districts were covered, 5000 of which referred to the County of London and the remainder in an out-county ring. A classification of enumeration districts in terms of their socio-economic characteristics would be produced.

The Centre research team experimented with 1961 Census data at both the ward and enumeration district levels. For Inner London[10] they produced a six-fold classification of enumeration districts using principal component analysis and a least-squares cluster analysis (Norman 1969). Several features of this work stand out, as may be seen in Table 1.2 which presents the results of the Inner London classification. First, and no doubt influenced by Charles Booth's poverty mapping, the naming of clusters ('Upper Class', 'Bed Sitter', 'Poor', 'Stable Working Class', 'Almost Suburban', 'Local Authority Housing'); secondly, the use of location quotients[11] to produce a statistical profile of each cluster); thirdly, the use of these statistical profiles to produce a verbal description of the main census characteristics of each cluster; and fourthly, the considerable variation that was found in the size of clusters: in this example, the Stable Working Class cluster accounted for almost a third of all enumeration districts, a fair reflection of the spatial structure of Inner London at that time. From that point onwards, these four features would become standard elements of geodemographic classifications.

The *Third Survey* was initiated in 1961 at a time when proposals were being considered for a comprehensive re-structuring of London's administration: by 1965, the Greater London Council (GLC) had been created. On the one hand, this new structure was bound to generate considerable interest in the results of the *Third Survey* and what it said about London's changing characteristics; while on the other hand some of the *Survey*'s findings would, by this time, be looking rather dated based as they were on the 1961 Census. Added to this, the newly established GLC would have its own technical capability in its Research and Intelligence Unit to carry out the same kind of census analyses.

---

[10]Inner London represented the area covered by the old (pre-1965) London County Council. There is no evidence that the Centre for Urban Studies team produced an enumeration district classification for the entire *Third Survey* study area.

[11]Table 1.2 shows location quotients that compare the cluster mean value of a variable with the Inner London value of the same variable. A quotient of less than 100 indicates the variable is under-represented in a cluster while a value of more than 100 indicates over-representation. See Batey and Brown (1995, pp. 95–102).

Table 1.2 Selected characteristics of six types of enumeration district in Inner London, 1961 Census

Number and Name of Variable		Enumeration District Type						All Inner London	
		Type 1 Upper Class	Type 2 Bed Sitter	Type 3 Poor	Type 4 Stable Working Class	Type 5 Local Authority Housing	Type 6 Almost Suburban	Mean	Coefficient of Variation
1.	% Population under 5	57	60	140	111	99	86	7.0	37
2.	% Population under 15	58	52	110	105	133	92	19.2	31
3.	% Population 65 and over	120	94	83	103	77	123	11.8	34
6.	Female/male ratio, all ages	130	105	92	96	94	102	1.123	21
7.	Female/male ratio, 25-44	126	91	115	95	106	105	1.005	22
8.	% women 20-24 never married	45	51	119	125	85	104	50.9	40
9.	% adults single	139	167	101	88	113	87	28.7	30
11.	% one-person households	149	205	112	94	58	77	22.5	51
12.	% one- and two-person households	125	144	105	99	141	97	52.6	25
13.	% households of five or more persons	68	44	97	94	159	92	12.4	50
15.	% all households overcrowded	49	159	228	91	86	32	6.9	83
16.	No. of persons per room	83	104	118	99	114	81	0.78	18
18.	% households sharing dwelling	38	66	218	122	18	62	29.1	105
19.	% households sharing dwelling stove sink	48	206	239	104	28	48	6.9	119
20.	% households sharing WC	57	159	195	120	390	63	30.4	78
21.	% households sharing dwelling stove sink WC	45	206	235	97	26	43	6.5	18
22.	% owner-occupiers	92	56	128	79	25	244	15.3	112
23.	% council tenants	23	16	31	63	352	42	21.2	146
24.	% private tenants, furnished accom.	138	355	182	68	13	68	13.0	110
25.	% private tenants, unfurnished accom.	118	83	120	135	35	94	46.0	56
28.	% born in India, Pakistan and Ceylon	190	350	120	50	30	80	1.0	150
30.	% born in British Caribbean	19	76	324	105	14	48	2.1	171
32.	% born outside British Isles and New Commonwealth	281	290	129	58	33	51	6.9	112
36.	% born in Ireland (both parts)	92	59	200	90	57	59	4.9	82
37.	% moved within year into local authority area	206	242	120	68	59	80	8.3	102
38.	% early school leavers	46	55	116	114	118	95	73.5	29
40.	% males in prof./manag. occupations	304	191	200	55	39	128	13.2	111
41.	% males in manual occupations	44	63	119	115	118	85	65.3	32
	%of ed's of each type	9.1	7.3	14.9	31.1	18.2	19.4	100	

Source: modified from Norman (1969)

This is largely what happened in practice. The recent availability of the 1966 10% Sample Census[12] opened the way for a classification of the 32 newly created London Boroughs (Kelly 1971); a report, also by Kelly, on the methodology used to construct this classification, including helpful advice on the choice of input variables and clustering methods (Kelly 1969); and a classification of wards in Greater London by Daly (1971) that comes closest to what had been intended in the *Third Survey*. The notion of a classification based on 1966 Census enumeration districts was postponed because of concerns about sampling and enumeration errors (Kelly 1969, p. 18). In due course the 1971 Census, containing a substantial amount of 100% data, would prove more suitable for this kind of analysis.

The GLC's classification work as described here was strongly influenced by that of the Centre for Urban Studies and fully acknowledged the methodological contribution of Ruth Glass. Its practical value lay in enabling systematic comparisons to be made at different spatial scales across the whole of Greater London. It is certainly the case that its didactic reports did much to encourage other local authorities in the UK to carry out their own census classification.[13]

In the meantime, some of the research findings from the *Third Survey* were published but never the full range that had been promised when the idea was first contemplated.[14]

1.4 Pioneering Geodemographic Classifications

1.4.1 *The City of Los Angeles and its Urban Information System*

A pertinent example of an early application of the public sector use of geodemographics in the USA concerns the city of Los Angeles. In the late 1960s and 1970s, the city's administration began to develop a comprehensive urban information system that integrated a wide range of spatial data relevant to the

---

[12]The 1966 10% Sample Census was intended to plug the gap left between decennial censuses which was felt to be too long. However, sample censuses have never been repeated, largely because of difficulties in measuring change at a local level caused by incompatible enumeration district boundaries, confidentiality problems, and sampling error.

[13]For example, in 1972 the present author carried out a multivariate analysis, at ward level, of the 1966 Census for the Greater Manchester area. The approach taken was strongly influenced by the Greater London studies.

[14]The Third Survey reports were announced as 'forthcoming' as late as 1973 (Glass 1973). A Fourth Survey, *Working Capital: Life and Labour in Contemporary London* (Buck et al. 2002) was carried in the late 1990s and published to coincide with the centenary of Charles Booth's survey in 2002. Based on an extensive programme of interviews across London and on detailed analysis of socio-economic data, the book's main message was that despite London's outstanding economic success as a global city, the city's prosperity was unequally distributed and was increasingly marred by social exclusion and deprivation.

activities of the city council, particularly in housing and planning. A recent retrospective review by Mark Vallianatos considered this ambitious venture into computer-assisted data and policy analysis to be well ahead of its time, comparing favourably with current initiatives to create 'smart cities' (Vallianatos 2015).

In establishing its Community Analysis Bureau (CAB), Los Angeles sought new tools to address the old challenges of deteriorating housing by providing detailed local data to identify neighbourhoods showing early signs of obsolescence. The bureau's data would, it was felt, help identify blighted areas across the city and inform measures aimed at alleviating the poverty that led to blight in the first place.

The US Census Bureau had gathered and reported statistics on housing quality between 1940 and 1960 but had abandoned this approach when it became clear that it was seriously over-estimating the amount of dilapidated housing. After 1960, the Census Bureau recommended looking at other characteristics such as building age, lack of plumbing, and overcrowding to infer housing quality. The CAB adapted and developed a range of analytic approaches to assess housing (and related social) conditions to fill this void left by the Census Bureau, and provide detailed local data to identify neighbourhoods showing early signs of obsolescence. First, however, the bureau had to digitise and centralise relevant information from the US Census, the Los Angeles Police Department, the LA County Assessor, and other private and public sources. In an effort to create a comprehensive Los Angeles Urban Information System, the bureau assembled a database containing 550 categories available to analyse individual census tracts. As Vallianatos (2015) points out, given the computing power then available, this would certainly have been regarded as 'big data'.

The CAB used cluster analysis in order to allow "the data to suggest its own 'natural' grouping." Clustering could identify parts of the city that might be geographically far apart but shared important social and physical characteristics. Sixty-six key items were chosen from the database, including population, ethnicity, education, housing, and crime data, and an environmental quality rating and LA's 750 census tracts were sorted into 30 clusters. It emerged that nowhere near 66 data variables were needed to identify which parts of the city had the worst blight and poverty. Three sets of data considered together—birth weight of infants, sixth-grade reading scores, and age of housing—were found to be an accurate indicator of housing decline and socio-economic deprivation. The bureau's data and analyses were intended to spur interventions in the city. They helped the city to move away from the traditional approach to urban renewal, with its focus on the treatment only of physical problems, to a more broadly-based approach that dealt with the social, economic, and physical nature of urban decay. Ultimately, however, the CAB was a victim of its own success. The data it collected proved so useful in securing federal grants that the city focused the CAB's activities on grant development and administration, with continued data analysis to justify these funds. Instead of using research to guide the city's actions, the bureau found itself reacting to the city's predetermined goals as set out in funding applications. By 1980, it had stopped

producing research reports and had been absorbed into the city's community development department (Fig. 1.3).[15]

1.4.2   Liverpool and its Social Area Analyses

In the late 1960s, the city of Liverpool in North West England was in serious social and economic decline. The City Council was engaged in a massive programme of slum clearance and new house building as part of efforts to regenerate the city. To help achieve these goals, the planning function in the City Council was being strengthened by creating a strong social research orientation. Planning was now seen as much broader in scope than physical planning. The City Council's role in social and community development was under active discussion in the aftermath of the UK Government's Seebohm Committee report (1968) which recommended fundamental changes in how social services were to be organised and delivered.

In 1967, the City Council hosted a conference to discuss community development in Liverpool. Attending the conference were city councillors, council officials, and representatives of community organisations in the city. The conference agreed that the city planning department should undertake a study to identify areas with large numbers of social problems, to help guide the allocation of social services resources and the establishment of a community development programme.

The so-called *Social Malaise Study,*[16] commenced in 1968, would concentrate on three elements: (1) an examination of 'social malfunction' throughout the city, to guide the allocation of extra physical and social resources; (2) an exploration of the degree of association between malaise and census indices, and of the important of better coordination of services to those in need; (3) a consideration of the impact of slum clearance and housing redevelopment, as well as various economic factors on the distribution of problem areas within the city. The information collected would guide policy but would also serve as an as educational exercise for city officials, helping to improve their understanding of the complex processes at work in the city. Unusually for a British local authority, the study team was advised in its early stages by an expert in community development, Professor Arthur Dunham, from the University of Michigan (Batey and Brown 1995, pp. 83–84). The City Council clearly wanted to be seen as a leading local authority in this field and the appointment of an international adviser would, no doubt, add to the prestige of the study.

Like the Los Angeles study described earlier, the *Social Malaise Study* assembled data from a range of sources:

- 1966 Census data, at enumeration district level (58 variables).
- Operational data assembled for 36 social malaise indicators from six City Council departments and from a wide range of other agencies on, e.g. job instability,

---

[15]For a full account of the Los Angeles Community Analysis Bureau, see Light (2003, pp. 78–83).

[16]See Flynn et al. (1972) for a comprehensive account of the Social Malaise Study.
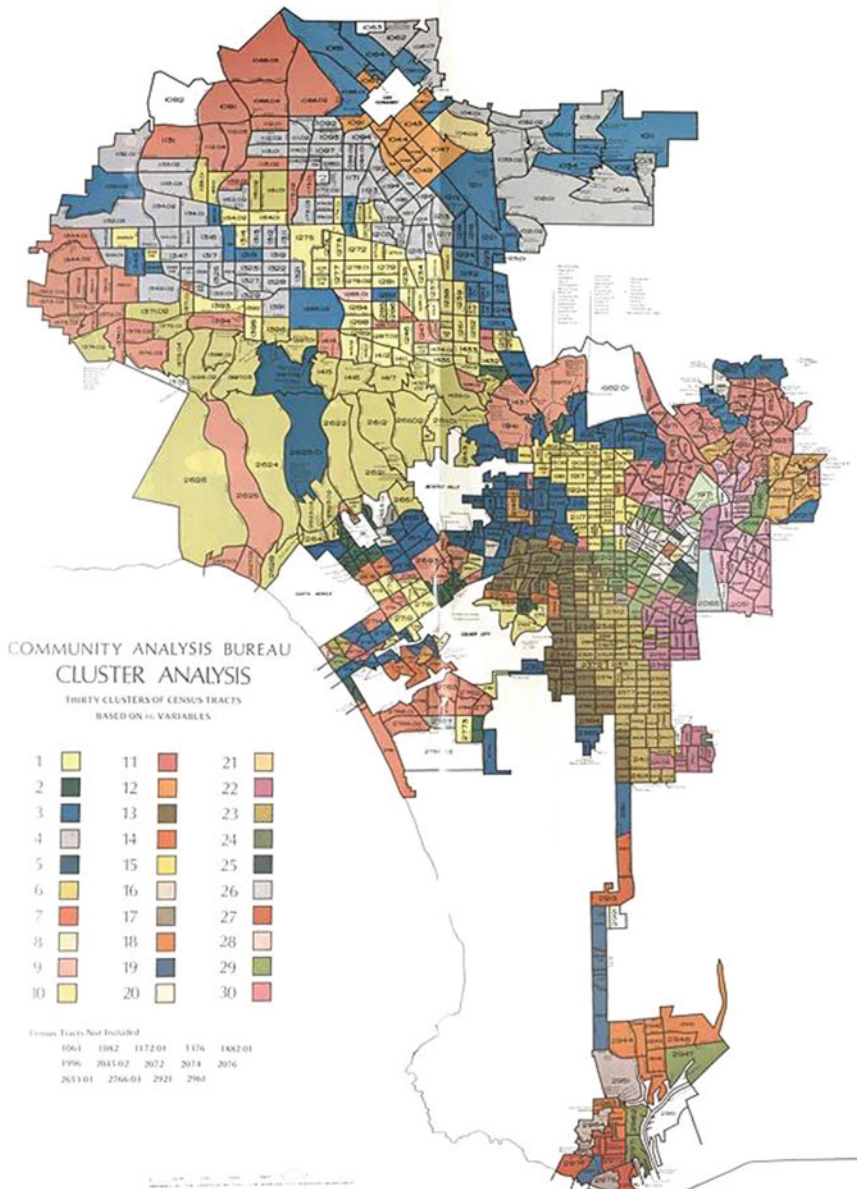
Fig. 1.3 The state of the city: a cluster analysis of Los Angeles, 1974 (Source: Vallianatos (2015))

crime, debtors, and possession orders, together with six housing variables (Amos 1969).

The task of collecting data was formidable and met with widespread resistance from those unconvinced as to the value of the study, as well as technical problems in coding operational data to the census enumeration districts. A correlation exercise was carried out for census and social malaise data at ward level, leading to a principal component analysis. The loadings on the first principal component were used as weights in creating a single index of social malaise which could then be used to define priority areas in the city.

The City Council was prepared to learn from the experience of this first *Social Malaise Study*. Four practical lessons were identified[17]:

- By concentrating on a single aggregate measure (score on principal component 1), the study failed to recognise the different kinds of urban stress experienced in cities.
- The proportion of the city's population shown to live in areas of severe need was out of all proportion to the funds available for allocation to priority areas.
- The analysis was carried out in the Planning Department, with other agencies playing little part in the project design and so not having much 'buy in' to the study, despite its high profile nationally.
- The technical capability was not retained with the result that further analysis that could have been done never materialised.

As the first of its kind, the *Social Malaise Study* proved controversial[18] attracting a lot of attention both locally and nationally: it was widely emulated by other local authorities in the early 1970s. This led to the holding of a conference in 1970 at which representatives from a range of disciplines were encouraged to criticise the study and suggest how it might be improved. The critics, who were largely academics, unsurprisingly pointed out the lack of underlying theory, limitations of the statistical analysis and the fact that, in the 3 years since the initial 1967 conference, the institutional and policy context had changed, as the Central Government implemented the Seebohm Committee's recommendations on the management and delivery of social services. In fairness, however, the scope of these changes could not have been fully anticipated by those commissioning the study.

The story of Liverpool's engagement with social area analysis does not end there. By 1974, there were calls for a new study. These came from a variety of sources: Liverpool City Council which was keen to extend its initial *Social Malaise Study* with the benefit of 1971 Census data; consultants Hugh Wilson and Lewis

---

[17]Webber (1975, p. 12).

[18]Some of the controversy arose because of the sensitive nature of the data being collected and the probability that some areas would be stigmatised as a result of high social malaise scores. The social malaise report was duly marked as 'highly confidential'.

Womersley who wanted an objective basis for defining boundaries of the Government-commissioned Liverpool Inner Area Study[19]; the newly established Merseyside (Metropolitan County Council), embarking on a strategic spatial plan and interested in developing cross-county information systems; the Centre for Environmental Studies and its Planning Research Applications Group (PRAG) which was seeking a test-bed for its area classifications and their application; and the Office of Population Censuses and Surveys (OPCS) which was interested in partnering PRAG in their census classification work.

This led to PRAG being commissioned to carry out a *Liverpool Social Area Study*. This was to be a well-resourced four-year study demonstrating advanced practice in assembling data sets, using computer-based analytical techniques and a range of actual and potential applications of the methodology. At the heart of the Study was a two-level geodemographic classification with 25 clusters and 5 so-called families (groupings of clusters).

The results of the study were widely disseminated in a number of reports each with a different target audience. The City Council produced a report combining the Social Area Analysis with more traditional methods for studying urban structure and long-term change in Liverpool and in comparator cities (Evans 1977); PRAG's report, written by leading census analyst Richard Webber as a demonstration project aimed at a wider audience of practitioners and academics (Webber 1975); PRAG's report on behalf of the Inner Area Study consultants making extensive use of visual and graphical techniques to illustrate clusters to show how they differed (Wilson and Womersley 1977); and a report showing how the classification had been extended to the wider area of Merseyside County and focusing on a range of applications (Webber 1978; Webber and Burrows (2018, pp. 54–61).

The collaboration with the OPCS proved to be very fruitful, leading to PRAG creating a series of classifications of parliamentary constituencies, the system of post-1974[20] local authorities in Britain and a number of individual local authorities. It culminated in the creation of two 1971 Census-based classifications, of wards and parishes and of enumeration districts, for Great Britain as a whole. These national classifications were to prove an important stepping stone for Webber as they started to generate serious interest from the private sector. Market analysts were quick to see the potential of area classifications. CACI, a leading US marketing company, made the first move and recruited Webber in order to help them launch ACORN[21] the first UK commercial product of its kind. ACORN was essentially a re-branding of the national classification that Webber had developed while working for PRAG (Batey and Brown 1995; Webber and Burrows 2018, pp. 62–63). For Webber, who had

---

[19]Inner Area Studies were commissioned by the UK Government in the early 1970s and were concerned with developing policy responses to a growing problem of multiple deprivation in the inner cities of England. Three studies were carried out: Liverpool, Birmingham, and Lambeth (London).

[20]In 1974 the system of local government was re-organised and a new, two-tier system of counties and districts was introduced. Local authority boundaries were comprehensively revised, hence the need to understand the socio-economic spatial structure of the new authorities.

[21]The term ACORN was an acronym for 'A Classification Of Residential Neighbourhoods'.

been the leading figure in developing public sector census classifications, this now opened up a highly successful career spent largely working in a commercial marketing environment.

From 1980 onwards there was a rapid growth in the development of proprietary area classifications, with several companies competing in the market place with products based on the 1981 Census. These products benefited greatly from increases in computing power and from new clustering methods that enabled much larger datasets to be handled. Although there were some variants, generally the statistical methodology adopted was very similar. Henceforth the field was to be known as 'geodemographics'.

Discussion 3: Los Angeles and Liverpool Experience Compared

It is interesting to note the parallels between the experience in Liverpool and that in Los Angeles. Both cities were engaged in a major programme of urban renewal/slum clearance and had started to question the wisdom of pursuing an entirely physical approach. There was general agreement on the need to consider the social dimension of housing renewal and to collect data that would enable this to be done. Data collection would involve a multi-agency approach and considerable effort would be needed to ensure consistency in this data. Careful planning was essential if all participant organisations were to be persuaded to 'buy in' to the project and, importantly, stay with it beyond the early stages. Both cities were conscious of the fact that what they were attempting to do was new and untried and at the limits of technical and computing capability of that time. The two cities were also in the throes of reorganising their community development provision making it difficult to maintain the information function that had been built up during the project.

Whereas Liverpool was fortunate in very soon after being part of a multi-faceted project leading to some important developments nationally in geodemographic analysis, in Los Angeles the City Council re-structuring meant an end to the urban information system by the late 1970s. However, this was not the end of the story in Los Angeles. An executive order from the city mayor in December 2013 instructed each city department to gather all the data it collects and share it on a publicly accessible website. By later the following year, Los Angeles had appointed its first Chief Innovation Officer and launched DataLA, the city's online data portal. Forty years on, the era of 'big data' and Smart City had finally arrived (Vallianatos 2015).

1.5 Geodemographics in Action as an Evaluation Tool

1.5.1 Area-Based Urban Policy Initiatives (ABIs)

Much of the discussion so far has viewed geodemographic classifications as a tool to guide the spatial targeting of resources.[22] In this section, the process is reversed and

---

[22]For a comprehensive guide to the uses and limitations of geodemographic classifications in a public policy context, see Lupton et al. 2011).

geodemographics is used to evaluate targeting that has already been done by other means, which may or may not be rational. The focus is on urban policy operating at a neighbourhood scale in the UK.

For at least 40 years, area-based initiatives (ABIs) have been an important feature of urban policy in the UK and have been seen as an effective means of targeting the poor. Successive governments have pursued a spatial targeting approach and introduced a range of policies and programmes identified through the use of area deprivation indices (for example, the Index of Deprivation (1980s); the Index of Local Conditions (1990s); and the Index of Multiple Deprivation (2000s) (Harris et al. 2005, pp. 42–45)). This prompts the question: how effective is spatial targeting in reaching the people for whom an urban policy initiative is intended and where areas are targeted as a proxy for individuals? With this question in mind a geodemographic evaluation tool will now be introduced and tested.[23]

What is the purpose of targeting associated with ABIs? Potentially, there is a wide range of possibilities. At one extreme, there could be a situation where the sole purpose of targeting is to identify a group of individuals who share a common set of characteristics that are relevant to the initiative (individual-, or people-oriented targeting). In this case, the attraction of an area-based approach is that it gives ready access to a concentration of such individuals and may help in the delivery of the initiative. At the other extreme, there could be an initiative that is entirely geographically-based, to the extent that the characteristics of the local population are completely irrelevant (area- or place-oriented targeting).

In practice, ABIs invariably lie somewhere between these two extremes. Even initiatives that appear at first sight to be either place-oriented or people-oriented turn out to be a combination of the two. What distinguishes them is the relative importance attached to targeting the individual and the area.

Following a study of government urban policy initiatives, Tunstall and Lupton (2003) put forward two simple concepts that help in considering the effectiveness of targeting: the notions of *efficiency* and *completeness*. Because the population of any given area is never perfectly differentiated by income, every area is, to some extent, mixed. This means that a degree of inefficiency is built into targeting by area, because people who are not the intended beneficiaries will be included. At the same time, the targeting will be incomplete, because deserving cases living outside the targeted area will be excluded.

The Tunstall and Lupton concepts can be put into practice by developing a method to measure the degree to which spatial targeting is successful. The proposed method draws upon a geodemographic classification system. The utility of the method is demonstrated by employing the P^2 People and Places geodemographic system[24] to assess the targeting of the Sure Start initiative in eight large provincial cities in England.

---

[23] A fuller account of the geodemographic evaluation tool and its application is provided in Batey and Brown (2007).

[24] P^2 People and Places was a proprietary geodemographic classification based on the 2001 Census. It was developed jointly by Beacon Dodsworth Ltd. and researchers at the University of Liverpool.

1.5.2 Characterising ABIs

Geodemographic classification systems may be used to establish the main types of residential neighbourhood associated with particular area-based initiatives. Moreover, they provide a means of judging how well the boundaries of regeneration initiatives reflect the spatial distribution of socio-economic need.

Any targeted area may be described in terms of a series of census Output Areas.[25] Local examples of the areas defined for ABIs are generally larger than a single Output Area and, although the match will not be perfect, it should be relatively easy to list the relevant Output Areas that constitute a targeted area. Describing targeted areas in this way enables them to be linked to the geodemographic classification which itself is based on Output Area level data. In the geodemographic system, each Output Area is assigned to a specific residential neighbourhood type (cluster), along with other Output Areas sharing similar characteristics.

The neighbourhood types conveniently summarise the main features of the population that is being targeted by an initiative. In practice several different neighbourhood types will be needed, rather than a single dominant type. Two closely related technical issues are important here: the mechanism by which these neighbourhood types are identified; and the task of measuring the closeness of fit between these neighbourhood types and the population targeted by the initiative. The objective here is to obtain the best possible approximation.

Two complementary approaches have been adopted in identifying the list of relevant neighbourhood types. The first of these is referred to as a 'penetration ranking' or concentration approach and identifies the neighbourhood types that have the greatest over-representation of the ABI population. The second approach employs a method of ranking based on the overall similarity between particular neighbourhood types and the general socio-economic profile of the ABI. This is described as a programme profile distance approach. In drawing up a final list of neighbourhood types, elements of the two approaches are combined.

By studying the composition of neighbourhood types that make up local instances of targeted areas across the complete set of local authorities, it is possible to establish whether there are particular types that occur more frequently than others. Taken together, such neighbourhood types are likely to account for the bulk of the total population resident in the targeted areas. These may be regarded as *Category 1 Neighbourhoods*. These neighbourhoods are likely to play an important part in characterising the areas targeted by a particular initiative.

The latest (2011) version of P^2 People and Places is described on the Beacon Dodsworth website: https://beacon-dodsworth.co.uk

[25] In the 2001 UK Census, Output Areas replaced enumeration districts as the smallest spatial units used to publish census data. They were designed to align with postcode geography and were created by aggregating unit postcodes. This made it possible to link post-coded data with small area census data much more accurately than had been possible previously, with obvious benefits for spatial targeting. By the time of the 2011 Census there were approximately 190,000 Output Areas in the UK, containing an average of 125 households.

However, it is also important to recognise that certain types of neighbourhood are concentrated in particular parts of the country, and may not emerge near the top of a national ranking of prevalent neighbourhoods. The method used here must be sufficiently flexible to reflect local and regional distinctiveness of this kind. To do this, it is necessary to define a second group of neighbourhood types, namely *Category 2 Neighbourhoods.* Such neighbourhoods have to satisfy the criterion that they are locally important (local here could mean a particular local authority area) in that they are over-represented in that area.

Inevitably, some neighbourhood types will lie outside Categories 1 and 2. These are defined as Category 3. Successful spatial targeting implies that most of the targeted areas are either Category 1 or 2 Neighbourhoods (reflecting the *efficient* targeting of the initiative), and that the incidence of Category 1 and 2 neighbourhoods outside the targeted areas is kept to a minimum (reflecting more *complete* targeting of those whose needs are greatest).

1.5.3   Measures of Targeting Performance

A number of simple measures can be calculated to describe how successful targeting has been, based on the cell values contained in a 2 × 2 table. To illustrate these, an example has been selected for the city of Nottingham. Here the characterisation of an unspecified area-based initiative has been done using the P^2 People and Places geodemographic system referred to earlier. The different categories of residential neighbourhood were identified using the Branch (40 cluster) level in P^2 People and Places by adopting the procedure outlined earlier (see Batey and Brown (2007) for a more detailed account). The generalised socio-economic profile of the initiative is based on the combined evidence of targeting across the entire set of eight cities.

In Table 1.3 the two rows represent the combination of Category 1 and Category 2 neighbourhoods (i.e. those whose needs are greatest) and Category 3 neighbourhoods (i.e. those whose needs are least), and the two columns represent, respectively, output Areas within, and outside, the areas on which the ABI programme is targeted.

In this table, the two main diagonal entries represent *correct targeting*—comprising, respectively, the *deserving* Categories (1 and 2) that fall within the defined initiative area boundaries and the *undeserving* Category (3) that fall outside the defined area. This "correctness" can be translated into a rate by adding the two figures together and dividing by the total population of the city and expressing the result as a percentage.

In this table, the two main diagonal entries represent *correct targeting*—comprising, respectively, the *deserving* Categories (1 and 2) that fall within the defined initiative area boundaries and the *undeserving* Category (3) that fall outside the defined area. This "correctness" can be translated into a rate by adding the two figures together and dividing by the total population of the city and expressing the result as a percentage.

Table 1.3 The match between targeted areas and neighbourhood categories: a population analysis for Nottingham

	Targeted areas	Non-targeted areas	Total
Category 1 or 2	42,648	40,294	82,942
Category 3	31,090	141,956	173,046
Total	73,738	182,250	255,988

Source: Batey and Brown (2007)

The two off-diagonal entries each represent different types of error, as follows:

Type 1 Error refers to *inefficiency*, or the capturing, within the initiative area, of people who are in the less deserving Category 3, and *Type 2 Error* refers to *incompleteness*, or the omission, from the defined area, of people who are in the more deserving Categories 1 and 2.

Table 1.3 shows the relevant counts relating to Nottingham for the selected area-based initiative. These counts are then used to derive the corresponding measures of inefficiency and incompleteness, as follows:

1. *Correct Targeting*: $(42,648 + 141,956) \times 100/255988 = 72.1\%$
2. *Targeting Error*: $100 -$ Correct Targeting $= 27.9\%$
3. *Type 1 Error (Inefficiency)*: $31090 \times 100/255988 = 12.1\%$, or 43.6% of total error
4. *Type 2 Error (Incompleteness)*: $74719 \times 100/255988 = 15.7\%$, or 56.4% of total error

In this example, approximately three-quarters of Nottingham's population is found to be correctly targeted, implying that the remaining quarter is not. For this quarter, it is possible to apportion the error between Types 1 and 2. Here, Type 2 (incompleteness) turns out to be appreciably more important than Type 1 (inefficiency). The implication is that in Nottingham, the boundary of the area-based initiative needs to be drawn more extensively, to include a greater number of people living in Category 1 and 2 neighbourhood types.

1.5.4 Application to a Specific ABI

The Sure Start programme is used here to demonstrate the practical application of the geodemographic assessment tool. By concentrating on a comparison between eight large provincial cities, the application also provides an opportunity to demonstrate how the assessment tool can be used to identify variations in targeting performance between areas with markedly different social and economic conditions.

Sure Start was a £3bn. 10-year national programme, launched in 1998, in which the intention was to work with parents, parents-to-be, and children to promote the physical, intellectual, and social development of babies and young children, particularly those that are disadvantaged. The programme was focussed on combating

Table 1.4 Comparison of inefficiency and incompleteness in the definition of Sure Start areas by city

	Sure start rate	Correct targeting (%)	Targeting error (%)	Type 1 error (inefficiency) (%)	Type 2 error (incompleteness) (%)
Bristol	138	86.6	13.4	61.7	38.3
Sheffield	194	82.2	17.8	58.7	41.3
Birmingham	119	79.5	20.5	43.3	56.7
Leeds	146	77.7	22.3	50.1	49.9
Newcastle	335	73.8	26.2	70.3	29.7
Liverpool	285	73.4	26.6	35.2	64.8
Nottingham	288	72.1	27.9	43.6	56.4
Manchester	312	67.1	32.9	41.8	58.2
All 8 Cities	199	77.2	22.8	51.5	48.5

Source: Batey and Brown (2007)

child poverty in neighbourhoods with concentrations of children aged 0–4 by reshaping existing support services (see Sure Start 2005).

Districts in receipt of Sure Start funding were selected according to levels of deprivation, but detailed decisions about the definition of individual Sure Start programme area boundaries were made locally. The starting point in each case was the national list of the 20% most deprived wards, as measured by the IMD (Index of Multiple Deprivation) 2000 (Noble et al. 2000). Draft Sure Start area boundaries were then modified using local knowledge (Frost 2005).

Table 1.4 presents the results for the eight cities. The first column shows the Sure Start rate: the number of local residents targeted by Sure Start per 1000 total population in each city. It indicates that there is substantial variation among the cities in the penetration of the Sure Start initiative. The next four columns show the measures of correct targeting and targeting error introduced in the Nottingham example. The resulting values provide a basis for ranking the eight cities. This ranking places Bristol at the top, with a correct targeting measure of 86.6 or 12% above the average for the eight cities as a group. The complement, targeting error, ranges from 13.4% for Bristol to 32.9% for Manchester, the latter 44% higher than the eight-city average of 22.8%.

The same table also records the two components of targeting error: Type 1 (Inefficiency) and Type 2 (Incompleteness). The table reveals that, in those cities with a higher rate of correct targeting, there is a tendency for Inefficiency to exceed Incompleteness, i.e. for a larger number of less deserving people to be included in Sure Start areas than should be. Similarly, towards the bottom of the table, notably in Liverpool (with the highest value of 64.8%), Nottingham and Manchester, incompleteness is more marked, implying that, in these cities, the Sure Start area boundary has been drawn too tightly, causing a greater proportion of potentially deserving recipients to be excluded.

A clear indication of the success of spatial targeting can be obtained by mapping the Category 1 and 2 neighbourhoods and the boundaries of the ABI. Figure 1.4
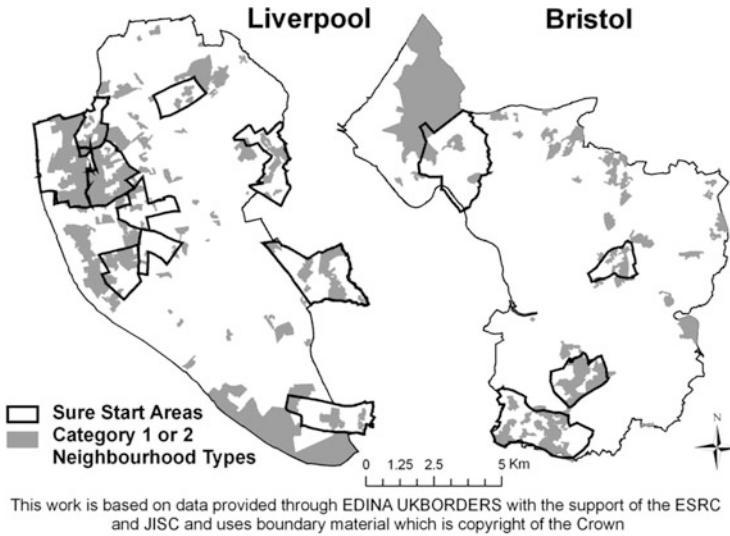
This work is based on data provided through EDINA UKBORDERS with the support of the ESRC
and JISC and uses boundary material which is copyright of the Crown

Fig. 1.4 The relationship between Sure Start areas and Category 1 and 2 neighbourhood types in Bristol and Liverpool (Source: Batey and Brown (2007))

presents maps of Bristol (where the targeting is relatively successful) and Liverpool (where it is less successful). The maps reveal that in both cities there are substantial areas that could equally well have been targeted and that there are some neighbourhoods where the targeting is hard to justify.

Discussion 4: **Benefits of a Geodemographic Evaluation Tool**

This case study has shown how a geodemographic approach can be employed to measure the success of spatial targeting of area-based urban policy initiatives. Even though the original basis for targeting may be obscure, and reflect political as much as technical factors, the approach presented here allows targeted areas to be analysed consistently and systematically. The geodemographic approach works by characterising the main types of residential neighbourhood that account for the bulk of the population in targeted areas. Some neighbourhood types are widely represented while others are distinctive to particular localities. Neighbourhoods that have been wrongly targeted can be easily identified, as can those that have been missed in targeting.

The geodemographic approach is flexible. The Sure Start case study used here has shown that it is possible to compare targeting performance in one city with that in other cities and thus to draw conclusions about the consistency with which particular nationally initiated area-based initiatives are implemented. In some instances, poor targeting is found to be a product of incomplete targeting, where the definition of targeted areas has stopped short of including the full complement of deserving areas. In other cases, the poor targeting outcome reflects an inefficient definition in which areas are targeted wrongly, resulting in a targeted population that includes a mixture

of neighbourhood types, only some of which are closely related to the socio-economic profile for Sure Start.

Some degree of spatial mis-targeting is inevitable and, indeed, it may be argued that this is no bad thing since it implies that, in any given targeted area, there will be some less deprived households that can serve as positive role models for those households intended to benefit from the policy initiative. However, the empirical results presented here in relation to Sure Start indicate that the quality of targeting is highly variable among cities and reveal that, even in the best cases, there is a substantial amount of mis-targeting. Taken as a whole, these results do give cause for concern and suggest that there is a plenty of scope for achieving better spatial targeting of urban policy initiatives. The geodemographic assessment tool described here provides clear guidance about where the emphasis should be placed in making these improvements.

1.6   Geodemographics Now

In this final section attention is drawn to two important developments of the last 10 years: *open geodemographics* in which data and computer software is made available free of charge to potential users in the public sector and in academia; and *geodemographics and spatial interaction data combined*, making it possible to join together residential and workplace classifications.

1.6.1   Open Geodemographics

Open geodemographics is intended to be highly flexible, in terms of geography, spatial scale, and choice of classification variables. The UK Office of National Statistics (ONS) first collaborated with Leeds University on constructing an Output Area Classification (OAC) based on the 2001 Census, and later worked in conjunction with University College London on a new classification based on the 2011 Census. The end product in each case was an Output Area Classification (OAC), with a hierarchical structure containing three levels, 8 Supergroups, 26 Groups, and 76 Subgroups in the 2011 version. The classification was based entirely on census data and the use of Output Areas—more than 190,000 covering the UK—meant that it was possible to tap the full spatial detail of the Census. There were 60 census variables in all, covering five domains: demographic; household composition; housing; socio-economic; and employment, broadly similar in scope to the variables used by the Centre for Urban Studies in the Third Survey enumeration district classification (Norman 1969).

The 2011 OAC went much further in terms of flexibility, allowing users to create their own classification based on a different geography and a different set of classification variables if desired.[26] A good case in point was London. In earlier, national classifications, in the 1990s and 2000s, London had proved problematical because, in many respects, it differed markedly from the UK as a whole and with each successive census these differences were becoming more pronounced. For the many users requiring a London classification, it was felt preferable to create a separate classification—L(ondon) OAC—to be made of the Greater London area with a set of more appropriate classification variables.[27] The successful creation of the LOAC led on to local authorities throughout the country being offered geo-data packs containing data specific to their area and access to software that enabled them to produce a tailor-made classification, without the need for specialised expertise and without cost, a major attraction to the many local authorities running tight budgets.

Alongside the 2011 OAC, the ONS created further classifications for different geographies and spatial scales. Like the OAC, these were hierarchical classifications with three levels, in this case 8 Supergroups, 16 Groups, and 24 Subgroups. Notable among these was a classification of UK local authorities, reminiscent of Moser and Scott's British Towns Study of the early 1960s. As shown in Fig. 1.5, there is a clear spatial structure to the classification, particularly in south east England where a series of concentric rings of area types radiate from the centre of London.

The second example of open geodemographics, *Patchwork Nation*, is quite different from the classifications described so far. Begun at the time of the 2008 US elections, the project was intended to create a usable, easily understandable tool for the media that would help combat simplistic views about America's socio-economic and political divides. It brought together academic social scientists and journalists working for a range of news media outlets, including the *Christian Science Monitor*, *PBS*, *Politico*, and the *Wall Street Journal*. Funding was provided by the Knight Foundation, a not-for-profit philanthropic organization that supports innovative projects in journalism, communities, and the arts, and the project hosted by the Jefferson Institute in Washington DC.

Like the *Planes of Living* example described earlier, *Patchwork Nation* used the 3144 counties as the building blocks for its area classification. However, unlike *Planes of Living*, which relied on just three classification criteria the *Patchwork Nation* project assembled a huge database, consisting of 150 variables, the vast majority drawn from the US Census. Chosen by the researchers for their relevance to present-day American politics, the scope of the variables was very wide indeed. It included data on population, local economic activity, and occupational mix, categories of consumer expenditure, racial and ethnic composition, religious adherence,

---

[26]The design of the 2011 OAC benefited from feedback gathered from users of the earlier OAC. See Gale et al. (2016).

[27]The London Datastore (https://data.london.gov.uk) is a free and open data sharing portal where anyone can access data relating to London. The LOAC may be accessed at https://data.london.gov.uk
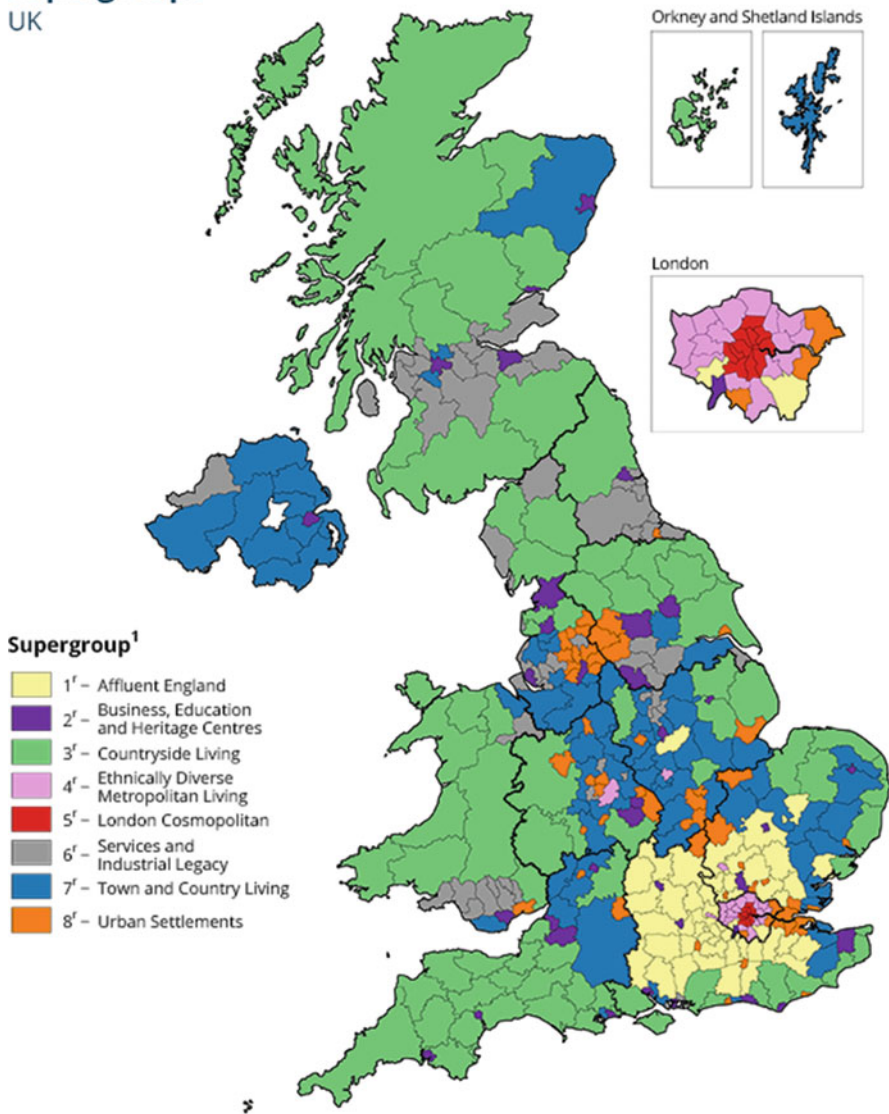
Fig. 1.5 2011 Area classification for UK local authorities: Groups (Source: https://www.ons.gov.uk/methodology/geography/geographicalproducts/areaclassifications/2011areaclassifications/maps, website accessed 17 Feb 2021)

immigration, education level, population density, housing stock, as well as several measures of income. There were also a number of variables measuring change, primarily relating to population. Where appropriate variable counts were converted into rates or percentages.

At the heart of *Patchwork Nation* was a geodemographic classification of the USA into different types of community. The methodology used to construct the classification was relatively simple and relied upon a standard principal component analysis in which all variables were included. Unlike many of the classifications described earlier in this chapter, there was no cluster analysis. Instead, the study focused on the leading principal components, in terms of overall variance explained, and the component score for each county on each of the principal components was used to decide to which of 12 community types a county should be assigned. The choice of 12 types was fairly arbitrary and reflected perceived ease of use by potential users as much as any particular statistical consideration.

Having created the 12 community types, considerable effort went into creating a profile of each type, along with specific examples of particular types. Journalists had an important role here, producing popular articles and features for both local and national media. Dante Chinni and James Gimpel, whose idea the project was, published a book that provided a more systematic account of each of the community types (Chinni and Gimpel 2010).

Table 1.5 shows the 12 community types, their labels and brief pen portraits in much the same manner as Charles Booth's poverty maps,[28] albeit with far less granularity. Figure 1.6, reminiscent of Carter Goodrich's *Plane of Living*, presents the community typology in interactive map form, enabling the reader to see how individual counties relate to the whole scheme of things.

As well as making *Patchwork Nation* comprehensible and accessible to the educated lay-person, the project team went out of its way to encourage readers to carry out their own analyses. This could involve producing maps for different geographies (state, region, nation), re-classifying counties according to different criteria, and maps for single variables. For this purpose the *Patchwork Nation* website supplies spreadsheets containing both the database as a whole and the database relating specifically to the 12 community types. Indeed the whole purpose of *Patchwork Nation* can be seen to be educational, aimed at achieving a better informed electorate able to see beyond traditional stereotypes.

1.6.2 Geodemographics and Spatial Interaction Data

Thus far in this chapter, geodemographics has been presented as a separate research tradition without any suggestion as to how it might link to other research fields of regional science. There are signs, however, that the picture is changing, with more
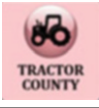
---

[28] See Fig. 1.1

Table 1.5 *Patchwork Nation* community types

Cluster name	Pen portrait
Boom towns	Fast growing communities with rapidly diversifying population
Immigration nation	Communities with large Latino populations and lower-than-average incomes, typically clustered in the south and south west
Monied 'burbs	Wealthier, highly educated communities with a median income of $15,000 above the national country average
Campus and careers	Cities and towns with young, educated populations; more secular and Democratic than other American communities
Industrial metropolis	Densely populated, highly diverse; incomes trend higher than the national average and voters lean Democratic
Mormon outposts	Homes to a large share of members of the Mormon Church and slightly higher median household incomes
Emptying nests	Home to many retirees and aging baby boomer populations: less diverse than the nation at large
Military bastions	Areas with high employment or related to the presence of the military and large veteran populations
Service worker centers	Midsize and small towns with economies fuelled by hotels, stores and restaurants and lower than average median household income by country
Evangelical epicenters	Communities with a high proportion of Evangelical Christians found mostly in small towns and suburbs; slightly older than the US average; loyal Republican voters
Minority central	Home to large pockets of black residents but a below average percentage of Hispanics and Asians

(continued)

Table 1.5 (continued)

Cluster name		Pen portrait
	Tractor country	Mostly rural and remote smaller towns with older populations and large agricultural sectors

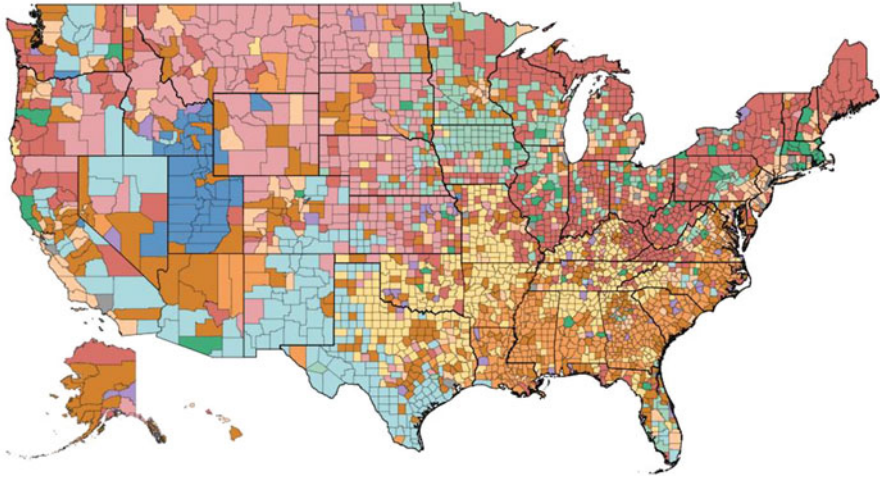Source: https://patchworknation.org/regions-page, website accessed 15 Feb 2021

Fig. 1.6 Map of *Patchwork Nation* community types (Source: https://patchworknation.org/regions-page, website accessed 15 Feb 2021)

attention being paid to various forms of integrated analysis that includes geodemographics.

A good opportunity is the link between geodemographics and spatial interaction data. In a study examining the effectiveness of area-based urban regeneration policy, Buck and Batey (2021) showed how, by combining small area census data on migration with a geodemographic classification of residential neighbourhoods, much could be learnt about the structure of migration patterns. Using 13 geodemographic area types, they were able to show the key elements of these patterns: an underlying pattern of migration to more affluent area types; three migration sub-systems showing strong interaction within groupings of affluent area types, deprived area types, and metropolitan area types; and an outlier representing new starters in the housing market. Moreover, the study was able to draw firm conclusions about the impact of spatially targeted urban policy initiatives upon migration between these geodemographic area types.

Without exception, the geodemographic classifications examined up till this point have been residential classifications. Thanks to a number of refinements in the 2011 Census, it became possible for the first time to construct a workplace geodemographic classification, COWZ-EW. The refinements consisted of a wider

spread of workplace variables and a system of purpose-designed Workplace Zones that replaced the Output Areas that up till then had had to suffice in representing workplace data.

In an important and ambitious paper, Martin et al. (2018) not only built a workplace classification, to stand alongside the 2011 OAC residential classification, but went a step further in constructing a classification of travel-to-work flows.[29] In doing so, they were able to analyse the 26 million travel-to-work flows in England and Wales, and to understand more clearly the different types of flow. This is likely to prove an important innovation that would find applications in a number of policy fields that had so far remained largely untouched by geodemographics, such as transport planning, labour market analysis, economic development, population mobility, gender studies, and energy consumption and pricing.

References

Alexander S (2007) A new civilisation? London surveyed 1928-1940s. Hist Workshop J 64:296–320

Alexiou A, Singleton AD, Longley P (2016) A classification of multi-dimensional open data for urban morphology. Built Environ 42:382–395

Amos FJC (1969) Social malaise in Liverpool: report on social problems and their distribution. Liverpool City Council, Liverpool

Bales K (1991) Charles Booth's survey of Life and Labour of the People in London. In: Bulmer M, Bales K, Sklar KK (eds) The social survey in historical perspective. Cambridge University Press, Cambridge, pp 1880–1940

Batey PWJ, Brown PJB (1995) From human ecology to customer targeting. In: Longley P, Clarke G (eds) GIS for business and service planning. GeoInformation International, Cambridge, pp 77–103

Batey PWJ, Brown PJB (2007) The spatial targeting of urban policy initiatives: a geodemographic assessment tool. Environ Plan A 39:2774–2793

Bell W (1955) Economic, family and ethnic status: an empirical test. Am Sociol Rev 20:45–52

Berry BJL, Horton FE (1970) Geographic perspectives on urban systems. Prentice-Hall, Englewood Cliffs

Buck M, Batey PWJ (2021) Migration, neighbourhood chance and the impact of area-based urban policy initiatives. In: Cochrane W et al (eds) Labor markets, migration and mobility. Singapore, Springer Nature

Buck N, Gordon I, Hall P, Harloe M, Kleinman M (2002) Working capital: life and labour in contemporary. Routledge, London

Carruthers J, Mulligan GF (2008) The plane of living and the pre-crisis evolution of housing values in the US. J Econ Geogr 12(1):739–773

Chinni D, Gimpel J (2010) Our patchwork nation: the surprising truth about the 'real' America. Gotham Books, New York

Daly M (1971) Greater London research: characteristics of 12 clusters of wards in Greater London, Research Report 13. GLC Department of Planning and Transportation, London

Evans ESP (1977) City in transition. Liverpool City Council, Liverpool

Flowerdew R, Leventhal B (1998) Under the microscope. New Perspect 18:36–38

---

[29] See also Hincks et al. (2018) for a similar geodemographic classification of commuting flows.

Flynn M, Flynn P, Mellor N (1972) Social malaise research: a study in Liverpool. Social Trends 2:
    42–52
Frost M (2005) Principles of GIS: a review of applications. http://bbk.ac.uk/geog/study/courses/
    pgis.pdf
Gale CG, Singleton AD, Bates AG, Longley PA (2016) Creating the 2011 area classification for
    output areas (2011 OAC). J Spat Inf Sci 12:1–27
Gittus E (1964) The structure of urban areas: a new approach. Town Plan Rev 35:5–20
Glass R (1963) The Centre for Urban Studies: extracts from the Quinquennial Report. Town Plan
    Rev 34:169–184
Glass R (1973) The mood of London. In: Donnison D, Eversley D (eds) London, urban patterns,
    problems and policies. Heinemann, London
Goodrich CB (1936) Migration and economic opportunity: the report of the study of population
    redistribution. University of Pennsylvania Press, Philadelphia
Goodrich CB, Allin W, Hayes M (1935) Migration and planes of living: 1920–34. University of
    Pennsylvania Press, Philadelphia
Greenwood MJ, Hunt GL (2003) The early history of migration research. Int Reg Sci Rev 26:3–37
Harris R, Sleight P, Webber R (eds) (2005) GIS, geodemographics, and neighbourhood targeting.
    Wiley, Chichester
Hincks S, Kingston R, Webb B, Wong C (2018) A new geodemographic classification of commuter
    flows. Int J Geogr Inf Sci:663–684
Hoover EM (1948) The location of economic activity. McGraw Hill, New York
Jefferson Institute. https://patchworknation.org/regions-page, website accessed 15 Feb 2021
Kelly F (1969) Classification of urban areas, Quarterly Bulletin of the Research and Intelligence
    Unit, 9. Greater London Council, London, pp 13–19
Kelly F (1971) Classification of the London Boroughs, Greater London Research and Intelligence
    Unit, Research Report 9. Greater London Council, London
Kitagawa EM, Taeuber KE (1963) Local community fact book Chicago metropolitan area 1960.
    Chicago Community Inventory, University of Chicago, Chicago
Leventhal B (2016) Geodemographics for marketers: using location analysis for research and
    marketing. Kogan Page, London
Light JS (2003) From warfare to welfare; Defense intellectuals and urban problems in Cold War
    America. The Johns Hopkins University Press, Baltimore
Light JS (2009) The nature of cities. Johns Hopkins University Press, Baltimore
Lupton R, Fenton A, Tunstall R, Harris R (2011) Place typologies and their policy applications,
    CASE Report 65. Centre for the Analysis of Social Exclusion, London School of Economics,
    London
Martin D, Gale C, Cockings S, Harfoot A (2018) Origin-destination geodemographics for analysis
    of travel to work flows. Comput Environ Urban Syst 67:68–79
Morgan MS (2019) Charles Booth's poverty maps. Thames & Hudson, London
Moser CA, Scott W (1961) British towns: a statistical study of their social and economic differ-
    ences. Oliver and Boyd, Edinburgh
Noble M, Smith G, Penhale B, Dibben C, Owen T, Lloyd M (2000) Measuring multiple deprivation
    at the local level: indices of deprivation 2000. Department of Environment, Transport and the
    Regions, London
Norman P (1969) Third survey of London life and labour: a new typology of London districts. In:
    Dogan M, Rokkan S (eds) Quantitative ecological analysis in the social sciences. MIT,
    Cambridge, pp 371–396
Park RE, Burgess EW, Mackenzie RD (eds) (1925) The city. University of Chicago Press, Chicago
Pfautz HW (1967) The times, the man and the inquiry. In: Pfautz HW (ed) Charles Booth on
    the city: physical pattern and social structure. University of Chicago Press, Chicago, pp 3–46
Rees PH (1972) Problems of classifying sub-areas within cities. In: Berry BJL, Smith KB (eds) City
    classification handbook: methods and applications. Wiley, London, pp 265–330

Robson BT (1969) Urban analysis: a study of city structure. Cambridge University Press, Cambridge

Robson BT (1984) A pleasant pain. In: Billinge M, Gregory D, Martin R (eds) Recollections of a revolution: geography as spatial science, vol 104-16. Macmillan, London

Selvin H, Bernert C (1985) Durkheim, Booth and Yule: the non-diffusion of an intellectual innovation. In: Bulmer M (ed) Essays on the history of British sociological research. Cambridge University Press, Cambridge, pp 70–82

Shevky E, Bell W (1955) Social area analysis: theory, illustrative application and computational procedures. Stanford University Press, Stanford

Shevky E, Williams M (1949) The social areas of Los Angeles: analysis and typology. University of California Press, Berkeley

Singleton AD, Spielman SE (2013) The past, present and future of geodemographic research in the United States and the United Kingdom. Prof Geogr 66:558–567

Theodorsen GA (ed) (1961) Studies in human ecology. Harper & Row, New York

Tobler W (1970) A computer movie. Econ Geogr 46:234–240

Tryon RC (1955) Identification of social areas by cluster analysis. University of California Press, Berkeley

Tunstall R, Lupton R (2003) Is targeting deprived areas an effective means to reach poor people? An assessment of the rationale, CASE Paper 70. Centre for Analysis of Social Exclusion, London School of Economics, London

Vallianatos M (2015) Uncovering the early history of 'big data' and 'smart cities': a critical appreciation http://www.boomcalifornia.com/2015/06/uncovering-the-early-history-of-big-data-and-the-smart-city-in-la/ Accessed on 15 Jan 2021

Van Arsdol MD, Camilleri SC, Schmid CF (1958) The generality of urban social area indexes. Am Sociol Rev 23:277–284

Vaughan L (2018) Mapping society. UCL Press, London

Webber R (1975) Liverpool social area study, 1971 data, PRAG Technical Paper 14. Centre for Environmental Studies, London

Webber R (1978) The social characteristics of residential neighbourhoods: Merseyside, PRAG Technical Paper 24. Centre for Environmental Studies, London

Webber R, Burrows R (2018) The predictive postcode: the geodemographic classification of British society. Sage, London

Wilson H, Womersley L (1977) Inner area study Liverpool social area analysis. Department of the Environment, London