

**Exploiting national survey and census data:
the role of locality and spatial effects**

Edited by
Angela Dale

CCSR Occasional Paper 12

April 1997

£5.00
ISBN 1 8990 514 5

Further copies of this paper may be obtained from:

CCSR
Faculty of Economic & Social Studies
The University of Manchester
Oxford Road
MANCHESTER M13 9PL

Tel: 0161 275 4721; Fax: 0161 275 4722
Email: ccsr@man.ac.uk
<http://les.man.ac.uk/ccsr>

Contents

	Page
Authors	ii
Introduction	iii
What are Spatial Effects and how can we measure them? Sally Macintyre	1
Some Thoughts on Space and Locality in Sociological Theory and Research Mike Savage	19
Multilevel Approaches to Modelling Contextuality: from Nuisance to Substance in the Analysis of Voting Behaviour Kelvyn Jones	29
Context and Culture in the Health of Ethnic Groups Chris Smaje	53

Authors

Professor Sally Macintyre
MRC Medical Sociology Unit
6 Lilybank Gardens
Glasgow
G12 8RZ

Professor Mike Savage
Department of Sociology
University of Manchester
Oxford Road
Manchester M13 9PL

Professor Kelvyn Jones
Department of Geography
University of Portsmouth
Lion Terrace
Portsmouth
Hants PO1 3HE

Dr Chris Smaje
Department of Sociology
University of Surrey
Guildford
Surrey GU2 5XH

Preface

This collection of papers arises from a meeting of the ESRC seminar series Exploiting the temporal and spatial dimension of census and survey data, jointly organised by Angela Dale, Sara Arber, Peter Elias and Denise Lievesley. The topic of the seminar, held at Manchester University in October 1995, was 'The role of locality and spatial effects', and brought a cross-disciplinary perspective to the meaning of 'place', its value in understanding social processes and how it can be incorporated into analysis.

Sally Macintyre's opening paper distinguishes between composition (the characteristics of the people who occupy the place), collective (the way in which those people may generate a collective identity or may function as a collectivity) and contextual (characteristics which may be ascribed to the place as distinct from its inhabitants). This very valuable distinction runs through the other papers in the volume. Sally's paper opens up a discussion on how the effects of context might operate and how one can make direct measurement of context. She provides an excellent example, drawn from a Glasgow study of two contrasting localities - one in the better north-western area of the city and another worse-off area in the south-west. The study obtained 'objective' measures of the environment and then checked this against the perceptions of the respondents who lived in the two localities.

Mike Savage provides a sociological context for debates over the role of space and place and traces the development of interest from the Chicago school through to the 1980s when a number of ESRC-funded studies of localities were funded. He suggests that the notion of discrete spatial localities with given effects has become hard to sustain and the more interesting issue relates to the spatial dimensions of social processes of interest - thus analysis of processes of movement through space rather than spaces as fixed places.

How one can establish the role of context is provided in a paper by Kelvin Jones which explains, with excellent clarity, the use of multilevel modelling. In this respect, statistical theory and software development is keeping pace with theoretical and conceptual understandings. A number of different examples are given, relating to voting behaviour and health, where the complexities of the relationship between place and individual can be captured. Finally, Chris Smaje discusses collective effects in relation to ethnicity. He asks whether geographical density - the 'collective' of Sally Macintyre - may have an effect and sets out an agenda for measuring this.

All the authors make reference to the problems of measuring place: these include locating appropriate boundaries - for example, if there is to be an adequate theory of place then one must be able to posit the process by which it operates and this requires different boundaries for different processes. The boundary of the education authority may be appropriate for educational research but for employment one may need the local labour market whilst for health research the health authority boundary may be more appropriate. This poses an immediate problem and data availability. Whilst the Samples of Anonymised Records provide a finer geography than available in any other nationally representative datasets, it is often inadequate for contextual analysis and provides a single set of boundaries only. The extent to which GIS-based innovations may be able to help overcome some of these problems forms the topic of a further seminar.

We are very grateful to the ESRC for funding the seminar series: Grant number R45 12643394

Angela Dale, Manchester, 1997

What are spatial effects and how can we measure them?

Sally Macintyre

INTRODUCTION

In a paper published in 1993 I argued that;

‘It may not be possible to make everyone middle class, but it might be possible to try to upgrade the social and physical environments of poorer people in ways which might be health promoting... Regional and District Councils should be encouraged to conduct health impact assessments on a wide range of policy options in different areas (for example in relation to housing, land use, transport, industrial development, policing, recreation, retail food provision). Health Authorities or Health Boards should be encouraged, as part of their health needs assessments, to focus on features of the local social environment as well as on characteristics of the local population’ (Macintyre, Maciver & Sooman 1993, p233)

The following year Sloggett and Joshi argued from their study of mortality in the OPCS Longitudinal Study that;

‘The evidence does not confirm any social miasma whereby the shorter life expectancy of disadvantaged people is further reduced if they live in close proximity to other disadvantaged people.Deprivation appears to be adequately assessed by personal and household circumstances, which are themselves associated with income. Area based measures are not efficient substitutes. For maximum effectiveness, health policy needs to target people as well as places’ (Sloggett & Joshi, 1994, pps 1473&1474)

These two quotations seem to be arguing the opposite; the first that more emphasis in health and social policy should be given to places rather than people, and the second that more emphasis should be given to people rather than places. In fact for a variety of reasons I believe that they are not as contradictory as they might initially appear.

One reason for the apparent contradiction might lie in the way the authors are conceptualising spatial effects. I and my co-authors were thinking of features of the local social and physical environment; Sloggett and Joshi seem to have been thinking of the effects of proximity to certain sorts of people. In this context it might be helpful to draw a distinction between three types of spatial effects commonly implied in the literature.

Firstly there is the idea that spatial effects, in this case strong associations between deprived areas and low life expectancy, are purely **compositional**. Poor people die early, so areas with lots of poor people have low average life expectancy. Poor people would die early wherever they live and rich people live longer wherever they live; the spatial effect is purely due to the spatial concentration of poor or rich people in different sorts of areas, and life expectancy is therefore a property of the individual, not of areas. This is the argument used in a recent comparison between the Paisley/Renfrew study and the Whitehall study, in which the authors argued that the observed differences in mortality between the two studies were due to differences in the distribution of types of people in the two study populations, not to differences

between the West of Scotland and London (Davey Smith, Shipley, Hole et al 1995). In the educational field this would suggest that bright pupils do well wherever they go to school, and schools that do well on average do so because they have a high proportion of bright pupils.

Secondly there is the idea that there might be, as Sloggett and Joshi put it, some type of 'social miasma'; that is, that the relations between different types of people, or the **collective** properties of members of a group, exert some sort of effect over and above the properties of the individual. A bright pupil put in a class with less bright pupils might do less well educationally than if she were in a class with pupils brighter than her, because the group culture, average level of aspirations, and interactions with teachers might be different. In this model living near lots of deprived people might decrease your life expectancy, and living near socially advantaged people might increase your life expectancy.

Thirdly, there is the idea of a broader social or physical **context**, an environment over and above either individuals or social groups, which might predate the individual or group, and over some aspects of which the individual or group has little control. Whatever your individual level of brightness, and whatever the social dynamics and average brightness of your fellow pupils, you might do better educationally if your school is well resourced in terms both of material plant and quality of teachers. Similarly, people of whatever levels of personal social advantage or disadvantage might live longer if they lived in non polluted areas with a pleasant climate and an excellent range of services and amenities.

I think its important to keep in mind the distinction between these three levels, which rather arbitrarily I am here calling **compositional**, **collective**, and **contextual**, because in the field of area variations in health they are often confused or conflated. What are often referred to as 'area effects' are often compositional effects defined in terms of aggregated individual data. For example, the work which relates mortality and morbidity rates to area level deprivation indices essentially uses a compositional approach, using aggregated mortality or morbidity and aggregated deprivation data as surrogates for individual level data in order to test the hypothesis of a link between deprivation and ill health. (Carstairs & Morris 1991, Townsend, Philimore and Beattie 1987)

As Jones and Moon have noted;

'Seldom however does location itself play a real part in the analysis; it is the canvass on which events happen but the nature of the locality and its role in structuring health status and health related behaviour is neglected' (Jones & Moon, 1993, p 515)

This suggests, correctly I think, that much work on spatial effects on health focus on the compositional level, or implicitly on the collective level, rather than on the contextual level. I'd like to illustrate this with reference to the literature on social class, area, and health, and then move on to discuss some of the ways in which we've tried to conceptualise and measure contextual locality based properties which might influence health and well being. In the course of doing this I will also mention some issues of measurement.

SOCIAL CLASS, AREA, AND HEALTH

Socio-economic status - whether measured by income, education or occupation - correlates closely with health and life expectancy. The lower you are down the socio-economic scale, the worse you do healthwise at every age. Not only are you more likely to die prematurely, but while you are alive you're more likely to be disabled or unhealthy. This observation, which is as

old as we have records - in the UK and US well known in the mid nineteenth century - has been attracting renewed attention recently. It has been shown that socio-economic inequalities in health and life expectancy seem to have been increasing. The gap between the richest and the poorest increased in the 1970s and 1980s in the US, the UK, New Zealand and Finland . (Macintyre, in press)

There are two main responses to these observations. The first is that the gap is due to poor health behaviours on the part of the poor - smoking, drinking, diet etc.- and that if poor people could behave more like the rich, the gap would disappear. The second response is that the only thing to do is to redistribute income and wealth so that they are more evenly spread, and then health would also be more evenly spread.

What both these approaches do is focus on the individual. What I want to suggest is that we should also focus on the social and physical environments which people inhabit and which may influence their health, either directly, or through the opportunities they provide for people to live healthy lives.

In Britain there has been a long tradition of research into associations between areas of residence and health. Since the 1850s statistics on mortality rates in the healthiest districts were held up as a standard against which other areas, with higher death rates, could be compared. However, most studies of areas and health have not actually looked directly at features of areas - for example, 'healthy districts' - which might promote health. Rather, there have been two sorts of studies of area.

The first sort of area study has been trying to discover the causes of specific diseases, and has focused on specific properties of the physical environment which might cause the disease. The characteristics of the environment typically examined include air pollution, water quality, toxins or other threats from industrial processes, climate etc. Usually this work is done at an ecological level i.e. rates of disease among the population in the most exposed areas are compared with rates of disease among the population in less exposed areas. However, the real interest is in individual exposure, and the area based analysis is usually undertaken because individual level data are not available. Although these studies look at areas with high levels of, for example, air pollution, what they are really interested in is individual exposure to air pollution.

The second type of study also uses area level data, but this time to look at the relationship between poverty and ill health. Data on mortality rates are linked to area level information, derived from censuses, such as median income, education, or unemployment rates. In Britain for example death rates have been examined in relation to scales of local deprivation which include male unemployment, proportion with no car, proportion of overcrowded housing, and proportion in the lowest two occupation social classes (the semi-skilled and unskilled manual workers). These have shown that areas with high rates of poverty measured at this area level tend to have high death rates. However, the real interest of these studies is the link between poverty and ill health at an individual level; the area level of analysis is only used because individual level information is not available or cannot be linked (Carstairs & Morris 1991, Townsend, Philimore & Beattie 1987).

So both these sorts of studies - those looking at specific environmental agents of disease, and those looking at poverty and mortality rates - tend to use areas as a proxy for individual data, rather than being interested in areas themselves (Macintyre, Maciver & Sooman 1993).

Several recent British studies have examined the relative importance of individual level and area level deprivation. Some find that area level deprivation is unimportant compared to individual level deprivation (Sloggett & Joshi 1994), but others have found that area of residence has an

influence on some health outcomes over and above individual level deprivation (Duncan, Jones & Moon 1995, Gould & Jones, in press). These latter studies echoed work by Mary Haan and colleagues on the Alameda County study in California, which showed that residents in a federally designated poverty area experienced higher mortality rates over a follow up period compared to residents of non poverty area. This increased risk of death persisted when account was taken of all the individual level characteristics which predict mortality - baseline health status, race, income, employment status, education, access to medical care, health insurance coverage, and a whole range of behavioural factors often assumed to be the link between socio-economic status and poor health. The authors concluded that "These results support the hypothesis that properties of the socio-physical environment may be important contributors to the association between low socio-economic status and excess mortality, and that this contribution is independent of individual behaviours" (Haan, Kaplan & Camacho 1987). In other words this suggests that over and above individual attributes of poverty, people of low socio-economic status may have poorer health because they had to live in areas which are health damaging.

But apart from the studies of particular environmental toxins, there is actually very little research on what environmental or cultural features of local areas or regions in contemporary societies might promote or damage health. This lack of information may stem from several causes. One is that people accidentally slip into thinking that characteristics of the people living in the area - the socio-demographic composition of the area - actually describes the area. Another is the assumption that any differences between areas are due to the character of the people living there. In Glasgow death rates in some neighbourhoods are two and a half times higher than those in other neighbourhoods. One response to this observation on the part of many people is to say "well that's because the people living there are all in poorer social classes - what can you expect?" Another is the assumption that we all know what different sorts of areas are like and therefore do not need to study them.

The problem with these sorts of reactions is firstly that they treat 'social class' and 'area' as if they are explanations in themselves, rather than attributes whose links to health need further clarification; and secondly they do not give any suggestions for policy other than trying to make people of low social class more like higher social class people (either by changing their behaviours or by changing their socio-economic circumstances - actions which may be either ineffective or not politically feasible).

HEALTH PROMOTING OR HEALTH DAMAGING FEATURES OF LOCAL AREA

I would therefore like to advocate looking at features of local areas which might be health promoting or health damaging. The literature does not give much guidance on what these might be, mainly because of the repeated use of aggregate individual measures means that much of the discussion tends to focus on the merits and demerits of various census based indices.

We have however been using as an organising framework the following five types of features of local areas which might influence health (Macintyre, Maciver & Sooman 1993)

1. Physical features of the environment shared by all residents in a locality. These include the quality of air and water, latitude, climate etc. and are likely to be shared by neighbourhoods across a wide area. In Glasgow, for example, all the drinking water for a city of nearly a million comes from the same loch, so the two and a half fold differences in death rates between neighbourhoods cannot be explained by variations in drinking water.
2. Availability of healthy environment at home, work and play. Areas varying their provision of decent housing, secure and non-hazardous employment, safe play areas for

children etc. These environments may not affect everyone living in an area in the same way as air and water quality do; they may affect the employed more than the unemployed, families with children more than elderly people, and so on.

3. Services provided, publicly or privately to support people in their daily lives. These include education, transport, street cleaning and lighting, policing, health and welfare services. Again, how these affect people may depend on personal circumstances. Public transport may matter more if you do not have a car.

4. Socio-cultural features of a neighbourhood. These include the political, economic, ethnic and religious history of a community: norms and values, the degree of community integration, levels of crime, incivilities and other threats to personal safety, and networks of community support.

5. The reputation of an area. How areas are perceived, by their residents, by service or amenity planners and providers, by banks and investors, may influence the infrastructure of the area, the self-esteem and morale of the resident, and who moves in and out of the area.

We all know about areas that either go downhill or become gentrified.

The effect of these environmental features may interact with personal circumstances. People with better personal resources may be better placed than their poorer neighbours to make use of certain opportunities, for example, by driving to out of town superstores to buy their food, using recreational facilities outside the area, or sending their kids to better schools outside the area. There may also be interactions between these different characteristics of the neighbourhood: areas with a poor reputation may find it hard to attract decent private or public services or secure inward investment for small businesses. Areas with a strong history of community integration may, however, be able to fight to retain or gain decent services such as schools.

Aspects of the physical and social environment may influence both mental and physical health, and may do so directly or in interaction with individual attributes.

Health services may not be a major determinant of variations in mortality between areas. However, inaccessible or poorly resourced local health services may be an additional stress for people already stressed by other personal and local circumstances - for example, poor, disabled or unemployed people or those caring for children or disabled relatives. The provision of facilities for outdoor recreation - tennis courts, football pitches, bowling greens - may not only enhance opportunities to develop or maintain cardiovascular fitness but may also promote mental health via self-esteem and social participation. A general sense of a threatening environment - one with high noise levels, high traffic density, dirty and poorly lit streets, high levels of crime, vandalism, litter and graffiti - may demoralise people and harm their mental health as well as directly compromising their physical health.

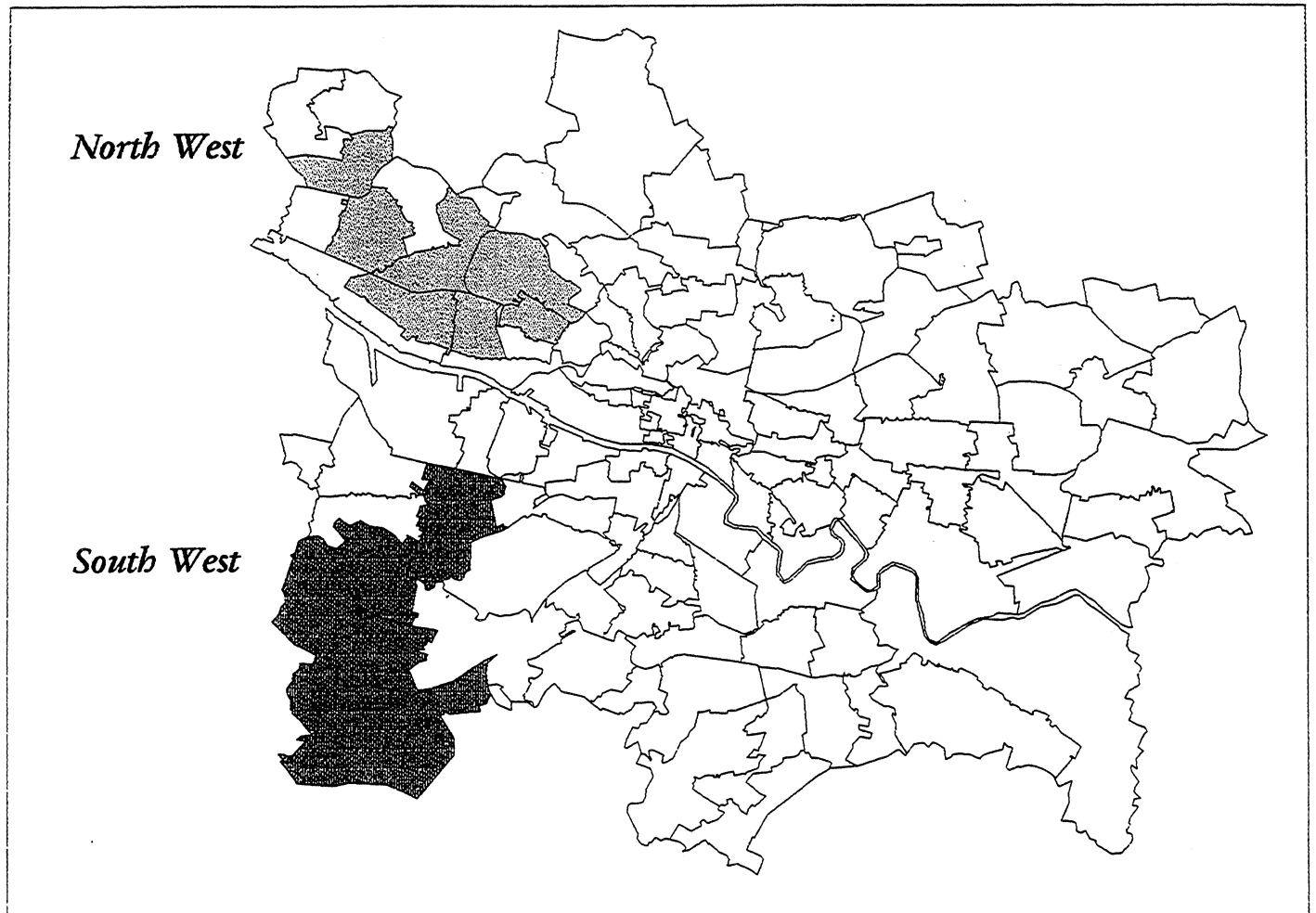
Is it possible to study these features of the social and physical environment directly?

RECENT FINDINGS FROM GLASGOW

I'd like to suggest that it is, and to illustrate this with some data we are collecting in the West of Scotland. The West of Scotland has relatively high death rates compared to other industrialised countries but within the West of Scotland some neighbourhoods have death rates as low as Japan or Sweden, that is, among the lowest in the world, while others have death rates two and a half times as high.

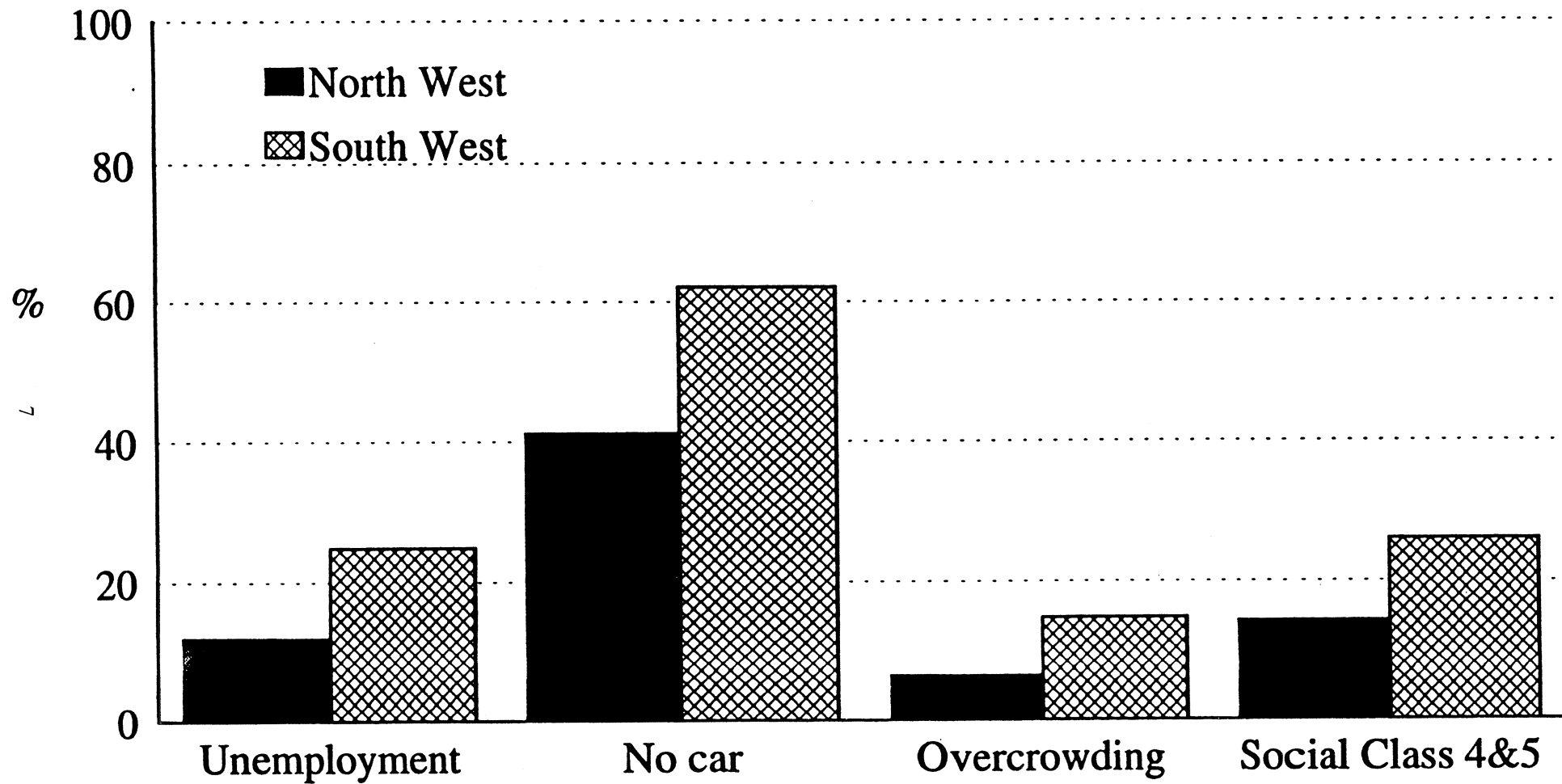
We have selected two localities in Glasgow City with contrasting socio-residential characteristics. The two areas are not at the extremes of a social continuum - not the richest and poorest - because we wanted to get some mixing in the individuals living in them.

Figure 1: Two Study Localities in Glasgow City



The NW is the better area, containing mainly private housing. Its standardised mortality rate is 17% below the average for Glasgow. The SW is the worse off area, containing mainly public housing. Its standardised mortality rate is 14% above the average for Glasgow. On a number of measures of aggregate social advantage/disadvantage the NW was better off at the 1991 census - less unemployment, fewer households without cars, less overcrowding, and fewer people in semi or unskilled blue collar jobs.

Figure 2 CHARACTERISTICS OF LOCAL AREA
1991 Census data



I will present some of the data we have collected about these two areas. We have been following up residents in them but the material shown here is mainly that directly collected by researchers. What is the service provision context of these localities?

Figure 3: Selected Shopping Statistics

1988

	North West	South West
Shops per 1,000 population	11	4
Number of:		
Strategic centres	6	3
Major local centres	9	5
Minor local centres	45	45
Single outlets	10	13

The NW has much better retail provision than the SW. In 1989 the NW had 11 shops per 1,000 population compared with 4 in the SW. The NW had 6 strategic shopping centres and 9 major local centres, compared with 3 and 5 respectively in the SW. There was the same number of minor local centres in both localities, and the SW had more single outlets (13 compared with 10 in the NW).

**Figure 4: Comparison of ‘Healthy’ and ‘Less healthy’
Food Baskets**

	‘Healthy’ basket (A)	‘Less healthy’ basket (B)	Extra cost of basket (A)
South West	£10.48	£9.02	£1.46 (16%)
North West	£9.94	£8.99	£0.95 (10%)

In 1992 we costed both a ‘less healthy’ basket of foodstuffs (white bread, whole milk, sausages etc.) and a ‘healthy’ basket (wholemeal bread, semi-skimmed milk, low-fat sausages). This involved researchers going round checking in a stratified sample of shops in both places. Both baskets cost more in the SW than in the NW; £9.03 in the SW compared with £8.99 in the NW for the ‘less healthy’ basket, and £10.48 in the SW compared with £9.94 for the ‘healthy’ basket. The difference in price between the ‘healthy’ and ‘less healthy’ versions was proportionately greater in the SW than in the NW (£1.46 [16%] more expensive in the SW, compared with £0.96 [10.7%] more expensive in the NW), thus creating a greater price disincentive to eating healthily in the poorer area (Sooman, Macintyre & Anderson 1993).

Figure 5: Recreation Facilities, North West and South West Localities

1990

	North West	South West
Number of:		
Athletic Tracks	2	1
Boating Ponds	3	0
Bowling Greens	12	8
Cycle Tracks	1	0
Golf Courses	1	2
Indoor Bowling Rink	1	0
Pitch and Putt	0	1
Playing Fields	15	1
Putting Greens	2	0
Recreation Centres	0	3
Sports Centres	0	1
Sports Hall	3	3
Swimming Pools	4	4
Tennis Courts	17	4
TOTAL	61	28

Source: Parks and Recreation Department, Glasgow City Council

There was strikingly less local access in the SW to healthy recreation. These data show a higher availability of both publicly and privately provided sporting recreation facilities in the NW compared with the SW; in total, people in the NW have access to 61 facilities within their local area, while those in the SW have access to only 28. (Since we collected these data a large public

sports hall has been opened in the NW which tips the balance even more in favour of that area.)

Figure 6: Selected Transport Statistics

1988

	North West	South West
<i>Bus Services: number of routes</i>		
shoppers bus service	3	0
hospital bus service	2	0
<i>Taxi stances (official)</i>	8	5
<i>Railway Stations</i>	8	4
Peak trains per hour	65	8
Off peak trains per hour	44	5
Sunday trains per hour	20	<1
<i>% households with no car</i>	42	62

Transport services show a marked contrast between the two areas, to the disadvantage of the SW. In 1991 62% of households in the SW had no car, compared with 42% in the NW. Yet there is no compensatory better provision of public transport services in the SW. In 1988 the SW was worse off in terms of special bus services and taxi stances and markedly worse off in terms of train services. This quantitative analysis does not take into account additional, more qualitative, contrasts such as the inconvenient siting of stations in the SW, or their less attractive physical appearance.

**Figure 7: Selected Primary Care Statistics, North West
and South West Localities**

1989

	North West	South West
Number of*:		
Health Centres	1	2
Clinics`	4	2
GP Practices	36	12
GPs	92	37
Dental Practices	24	10
Dentists	38	13
Community Dental Clinics	4	2
Opticians	18	5
Pharmacies	29	17

* including those on the locality peripheries

Source: Primary and Community Care Unit, Greater Glasgow Health Board

The quantity of primary health care services is also less in the SW than in the NW. There were three times as many general practices, over twice as many primary care physicians, three times as many dentists, nearly four times as many opticians, and one and a half times as many pharmacies in the NW than in the SW. This lower provision of primary care in the two localities also has to be seen against the backdrop of higher death rates for all major causes of death in the SW compared to the NW; the lower car ownership in the SW; and the higher mean number of GP consultations in the SW compared to the NW (around three and a half compared to two and a half for adults in the previous year). In relation to these admittedly crude indicators of "need", even identical primary care provision in the two areas could be interpreted as relatively disadvantageous to the SW (Wyke, Campbell & Mciver 1992).

**Figure 8: Monthly Average of Reported Crime Rates, North West
and South West Localities**

1990

	North West (Partick)	South West (Pollok)
Serious Assault	4	16
Assault and Robbery	12	16
Vandalism	70	96
Breaking into houses	120	138
Breaking into other premises	100	28
Theft of car	100	57
Theft from car	100	61
Attempted theft of/from cars	100	9
Breach of the Peace	110	61

Source: Strathclyde Police

Finally, data were obtained from Strathclyde Police on reported crime rates in the sub-divisions or divisions covering our two study areas. This shows a greater incidence of crimes against the person, and against personal space, in the SW; and a greater incidence of crimes against commercial premises and cars in the NW, where there are more cars to steal or steal from, and more shops and other premises from whom theft might be profitable.

I will move on now to some of the data collected from residents in these localities. As described above, we had found that the availability of healthy food was higher and prices lower in the NW than in the SW. We did an analysis of reported food consumption and found significant differences between neighbourhoods after controlling for individual socio-demographic characteristics. Some foods (for example, low fat milk, white fish, confectionery, cakes and pastries, and savoury snacks) seemed to be part of a wider Scottish or Glaswegian diet and not to vary between neighbourhoods. Others showed marked differences between neighbourhoods after controlling for sex, age, income and social class: these include fruit, vegetables, meat (particularly processed meat products), bread, spreading fats, sugar, natural fruit juice and alcohol. This suggests that such intraurban variations in food consumption cannot be explained simply by socio-demographic or socio-economic factors in individuals and that cultural and supply side factors need to be taken into account (Forsyth, Macintyre & Anderson 1994).

We asked about residents' perceptions of their local areas within 6 main domains - local amenities, local problems, reputation, neighbourliness, fear of crime, and general satisfaction with the area. Here I present some selected data on problems and reputation.

Figure 9: Perceptions of Problems

	North West	South West
	%	%
vandalism	56	40
litter & rubbish	56	69
smells & fumes	25	16
assaults & muggings	30	62
burglaries	88	81
disturbances from youths	25	41
traffic	53	56
discarded needles	6	16
uneven pavements	43	51
dogs	54	56
poor public transport	56	65

There is a generally consistent pattern of more problems being reported in the poorer area, although the magnitude of differences between them varies with the particular problem. An interesting exception is for smells and fumes, which were more likely to be reported in the better areas. We were puzzled by this but when we checked we found there was a rape seed processing factory and a sewage plant to the SW of these neighbourhoods, and the prevailing winds blow the fumes into the NW. This suggests that respondents in the poorer areas are not simply responding more negatively because of generally more pessimistic response set but are responding to real environmental differences.

In fact some of our data suggest that the reported perceptions may minimise rather than exaggerate socio-physical differences. We had found marked differences in public transport and sports provision when we looked at these directly: but there was less difference between the localities in the proportion reporting poor transport; and more people in the SW said they had sports facilities within half a mile, suggesting that those in the poorer areas, although more in need of public transport and with poor public transport available to them, had lower standards and expectations than those in the better areas.

Another example of the way in which people's perceptions may minimise problems compared to an external observer's perceptions was apparent from some work we have done on respondents' views about crowding. When interviewed in 1992, 11% of our respondents reported households densities which would be defined as overcrowded according to the census definition. However, of these only around 40% reported that they were in overcrowded accommodation. This raises the measurement issue of whether you describe features of places in insiders' or outsiders' terms. Our data show bigger differences in outsiders' terms.

We asked questions about whether people had ever been refused services because of their address. This had only happened to very few people but in the case of the police and credit, was significantly more often reported in the poorest area. This indicates something about how agencies view the areas, and confirms the picture obtained from external sources.

Figure 10: Area Reputation

	North West	South West
<i>In the last three years:</i>	<i>%</i>	<i>%</i>
refused taxi because of address	2.1	3.8
refused ambulance because of address	1.7	4.3
refused police because of address	2.2	7.0
refused credit because of address	2.3	7.1

The detail of these findings about our study localities is less important than the fact that we have been trying to develop ways of conceptualising and measuring people's local social and physical environments, in order to see whether these have any influence on people's health and their opportunities to lead healthy lives. In terms of the compositional, collective and contextual distinction I made earlier, it should be clear that we have been trying to think about the contextual rather than the compositional level. We have also not yet tried to look at the collective level, which we conceptualise in terms of social relations, social cohesion etc. We do not accept that the only way in which place and locality can affect peoples lives and health is through some social miasma process. Rather we have been thinking in quite concrete ways about services, facilities, amenities, incivilities in the environment, and barriers to healthy living.

We also conceptualise spatial effects as being dynamic rather than static. In analysing of local variations in diet we felt that one had to see demand and supply for healthy foodstuffs dynamically. There is no point blaming suppliers for not selling healthy foods at low prices in deprived areas if there is no demand for these foods. Areas with lots of poor people in them

differ in a range of ways from other areas; not only in the built environment, in services and amenities, but also in terms of traditions and the expectations of their residents.

CONCLUSIONS

What are the implications of our work for the question 'what are spatial effects and how can we measure them?'

This work represents an attempt to move beyond census level data to try to characterise features of areas and their possible effects on health. Rather than thinking in terms of percentage of male unemployment, or percentage of residents in social classes IV and V, it means looking at local employment opportunities for men and women (since these will, in interaction with personal characteristics of the residents, determine the percentage of unemployed or in particular social classes). When shipyards close, skilled manual jobs for males disappear, so the proportion in social class III manual will reduce; when new service or high tech industries open then there will be more people employed and a hike in the proportion in white collar employment.

What about measurement issues? I am particularly interested here in problems of measuring locality characteristics, rather than problems associated with measuring health outcomes.

Firstly, there is the issue I've already mentioned; direct measurement by outsiders, for example of sports provision or transport facilities, may give both a more accurate and a more contrasting picture of services available than asking residents for their perceptions. This is probably because of lower expectations among residents of poorer areas.

Secondly, there remains the tendency to think of individual behaviours rather than structural or systems properties. As well as looking at crime rates, which we have done, we should be looking at policing policies and levels. However, as we discovered when looking at social work provision, it is much harder to look at aspects of policies and provision than it is to look at records of number of clients dealt with, or records of activities (similarly it is easier to obtain unemployment rates than job vacancy rates).

Thirdly, if one is interested in national or large scale studies it is difficult both to obtain much area level data consistently over the whole country, and to obtain it in spatial units which can be linked to the spatial units in which mortality or morbidity data can be analysed.

However, in conclusion I would suggest that both for analytic reasons-- trying to understand the social production of health-- and for policy reasons-- trying to find strategies for improving the health of the population-- it is important to focus on places as well as people. This means trying to supplement census or other survey data on individuals with structural or area level data. This is difficult and poses a number of challenges for access to and manipulation of data, but is worth trying. In doing this it is important to maintain the conceptual distinction between the three types of spatial processes--compositional, collective, and contextual--described earlier.

ACKNOWLEDGEMENTS

I am grateful for the financial support of the Medical Research Council who funded the research reported here. I am also grateful to Anne Ellaway, Sheila Maciver, Jenny Taggart, and Alastair Forsyth for data collection and analysis.

REFERENCES

- Carstairs V, Morris R (1991) *Deprivation and Health in Scotland*. Aberdeen University Press
- Davey Smith G, Shipley M, Hole D, Hart C, Watt G, Gillis C, Marmot MG, Hawthorne V. (1995) Explaining male mortality differentials between the west of Scotland and the south of England. Paper presented at annual meeting of Society for Social Medicine 1995. *Journal of Epidemiology and Community Health*, 49,5, 541
- Duncan C, Jones K, Moon G (1995) Psychiatric morbidity; a multi level approach to regional variations in the UK. *Journal of Epidemiology and Community Health* 49,, pp. 290-295
- Forsyth A, Macintyre S, Anderson A (1994) Diets for disease? Intraurban variation in reported food consumption in Glasgow. *Appetite* 22,, pp. 259-274
- Gould M, Jones K (in press) Analysing perceived limiting long term illness using UK census microdata. *Social Science and Medicine*
- Haan M, Kaplan G, Camacho T (1987) Poverty and health; prospective evidence from the Alameda county study. *American Journal of Epidemiology* 125,, pp. 989-998
- Jones K, Moon G (1993) Medical geography; taking space seriously. *Progress in human geography* 17, 4, pp. 515-524
- Macintyre S (in press) The Black report and beyond; what are the issues? *Social Science and Medicine*
- Macintyre S, Maciver S, Sooman A (1993) Area, class and health; should we be focusing on places or people? *Journal of Social Policy* 22,, pp. 213-234
- Sloggett A, Joshi H (1994) Higher mortality in deprived areas: community or personal disadvantage? *British Medical Journal* 309,, pp. 1470-1474
- Sooman A, Macintyre S, Anderson A (1993) Scotland's health; a more difficult challenge for some? The price and availability of healthy foods in socially contrasting localities in the West of Scotland. *Health Bulletin* 51,, pp. 276-284
- Townsend P, Philimore P, Beattie A (1987) *Health and deprivation: Inequality and the North*. London, Croom Helm
- Wyke S, Campbell G, Maciver S (1992) The provision of primary care services in a working class and middle class locality in Glasgow. *British Journal of General Practice* 42,, pp. 271-275

Some Thoughts on Space and Locality in Sociological Theory and Research

Mike Savage

For some considerable time now the question of space has loomed large in sociological inquiry. As with the study of time (for which see Adam 1990), the question has increasingly been seen as central to social theory. The key problem is how to move from analysing how things happen in space (or time), towards seeing how things happen through space (or time). Whereas in the positivist tradition explanation involved abstracting from space and time, and then introducing spatial and temporal parameters to abstract law-like propositions (so that, for instance, in industrial societies, class was more important than status), it is now increasingly recognised that explanation involves seeing **how** things occur in and through time and space. Explanation is related to issues of process¹. These issues are now well rehearsed in theoretical debate² and I do not intend to take them up directly here. This paper will rather examine some of the practical difficulties involved in taking space seriously in social research and will offer a survey of recent developments. My argument will be that in the past decade there has been a shift from seeing space as 'locality', characterised by fixed boundaries, to more mutable notions of space which are more attuned to change. I will consider some of the reasons for this shift and discuss some of the problems in seeing places as 'localities' drawing partly on earlier critiques (Duncan 1989; Duncan and Savage 1989; Savage 1990). The main point will be that the promise of 'space' remains remarkably difficult to redeem in rigorous empirical research, and this paper is written in the hope that a broad awareness of this point is the best way to develop inquiries further.

The paper begins by examining the rise of 'national' level research in the years after 1945 and then emphasising the remarkable recovery of locality in the 1980s. The paper will argue that the reasons for the 'recovery of locality' were diverse and, ultimately, difficult to reconcile. During the 1990s it has become increasingly clear that the theorisation of space has led interest away from locality and that something of an impasse has now been reached. The paper concludes with a few speculative thoughts concerning the way forward.

1 SPACE IN BRITISH SOCIOLOGICAL INQUIRY: A BRIEF SURVEY

A concern with space lies at the heart of the sociological tradition, but that this tended to be forgotten in the 1960s and 1970s when the growth of survey research tended to valorise the institution of the national survey. Until the 1960s, nearly all sociological research was closely involved with community studies in definite local contexts. This can be traced through to the precursors to social research such as Mayhew's and Booth's studies of poverty in London (which in the case of Booth produced an extremely detailed social geography of London). The development of urban sociology, both in Britain and America led to a preoccupation with spatial processes, often thought of in ecological terms (see Savage and Warde 1993; Saunders 1981), and even stratification based research was frequently based on local case studies. This was true of studies such as the Affluent Worker series based in Luton (Goldthorpe and

¹ I do not wish to defend this argument here, but it should be noted that I see this argument as resting upon a critical realist approach to social inquiry, for which see Collier (1983), Layder (1990), Sayer (1984).

² A good starting point for this debate is Gregory and Urry's collection *Social Relations and Spatial Structures* (Basingstoke 1985). Giddens's work, notably *Central Problems in Social Theory* (1979), and *The Constitution of Society* (1984), offers one well known excavation of this problem. A recent account of issues of space in social theory is Gregory's *Geographical Imaginations* (1984).

Lockwood 1968/69), or the study of deferential agricultural workers in East Anglia (eg Newby 1977) as late as the 1960s and 1970s.

Nonetheless, after 1945 two developments allowed a radical break with this localist tradition. Methodologically, the development of sample surveys permitted research to take place across entire nations, and indeed it became a central claim that such national surveys provided more reliable findings than the more idiosyncratic pictures provided by local studies. Indeed it became something of a hallmark of modern nationhood to provide national survey instruments to complement the Censuses which had historically been in existence for longer periods of time. In the British case annual surveys such as the Family Expenditure Survey, or the General Household Survey became major research instruments from the 1960s. Theoretically, this development was related to the fact that the nation state appeared to be the key spatial unit of the modern world with the result that it was possible to talk about 'national' societies, with their own social structures, and so forth. Interestingly, the focus on national social structures, was developed largely by researchers within the 'conflict' tradition of sociology who were opposed to claims about the convergence of industrial societies around the industrial (or American) model. One of the clearest formulations of this argument was that of C Wright Mills, who in 1959 provided a rationale for national studies which was to have a long lasting impact:

'In our period, social structures are usually organized under a political state. In terms of power, and in many other interesting terms as well, the most inclusive unit of social structure is the nation state.... I think that when most social scientists come seriously to examine a significant problem, they find it most difficult to formulate in terms any smaller than the nation-state' (Mills 1959: 136-137).

These currents in social research therefore led to a focus on the nation. National borders became the boundaries in which distinctive social relations were constituted, the sorts of 'empty containers' for various social process. Such a formulation allowed a variety of research projects examining national class structures (eg Goldthorpe 1980), gender relations (eg Dex 1985), welfare regimes (Esping-Andersen 1990), and so on to be instituted. Insofar as space was taken seriously it was largely through international comparative research, for instance in research involving comparison between (rarely within) nations (eg Erikson and Goldthorpe 1992; Esping-Andersen 1990), and which in turn were linked to macro-historical studies of the development of nation states (eg Moore 1967; Anderson 1974; Skocpol 1979).

Although it would be foolish to exaggerate the trend, since other more localised forms of research continued during the post war years, it is nonetheless clear that the years after 1945 saw a distinct shift to a national frame of inquiry using national surveys as their prime instrument. In Britain during the 1980s however, a marked shift in emphasis occurred. The mid 1980s can be seen as marking a distinct 'return to locality' which appeared to see a major challenge to the established focus on national level analysis. One striking example of this quite dramatic shift of focus was the ESRC funding of a number of major research projects which were all based on comparisons of different 'localities' within Britain. In the mid to late 1980s there were no less than four research initiatives of this type. The Social Change and Economic Life Initiative, directed by Duncan Gallie, obtained work histories, information on local labour markets and employer strategies in six 'localities' (Aberdeen, Swindon, Rochdale, Northampton, Kircaldy, and Coventry) (see eg Gallie 1994). The Changing Urban and Regional Systems initiative compared local economic restructuring and social change in Lancaster, the Isle of Thanet, Cheltenham, Teesside, Swindon, South West Birmingham and outer Liverpool (see Cooke

1990). The 16-19 Initiative looked at the transition from school to work in four places (Liverpool, Sheffield, Kircaldy and Swindon), and finally the Economic Restructuring, Social Change and the Locality programme at Sussex University compared Slough, Derby and Burnley (see Saunders 1990 for one example).

What, then was this recovery of locality all about? It appeared to indicate a much greater appreciation of space and locality than existed in earlier work. But what explains this dramatic change of emphasis? It is my contention that the 1980s 'recovery of locality' contained within itself a series of incompatible ideas and emphases which have more recently proved to pull apart the promise this research suggested. Let me develop this argument by examining four reasons why the shift to locality took place in the 1980s in order to explain the excitement occasioned by this research.

2 RECOVERING THE LOCALITY

My main point is that there were no less than four analytically separate, reasons why researchers looked to the locality, but that because researchers did not always recognise the contingent nature of the connections between these four aspects, they were unable to systematically relate them together.

The first reason was a purely descriptive recognition of the growth of local variation within Britain. In a number of ways it was clear that local variation appeared to be markedly increasing during the 1980s, with the result that social division appeared to have a more distinct geography and concern about the rise of a 'North-South divide' (Lewis and Townsend 1986) was marked. Three sorts of spatial division were especially evident, both to academics and to laypeople. The economic recession of the early 1980s hit the old industrial areas much harder than the Home Counties and London, and the result was a marked opening up of regional differences in unemployment and wage rates (Massey 1984; Cooke ed 1990; Savage 1989). Secondly, and relatedly, there were marked divergences in regional house prices, with price rises in the South East of England rapidly outstripping those elsewhere. Finally, spatial variation in voting patterns became more marked, with the Conservative Party becoming more dominant in the Home Counties and the South East, whilst the Labour Party retained much of its strength in its traditional industrial heartlands (eg Johnson and Pattie 1990). These three interlinked changes made researchers directly aware, at a purely descriptive level, that Britain appeared to be a more divided and spatially fragmented country and suggested that such variation could best be examined by local case studies.

This purely descriptive account rested alongside an altogether different trend, related to a significant theoretical shift in social science research. A variety of sources suggested that the 'local' might help resolve a series of deep seated theoretical problems which dogged the social sciences. Most important of these was the 'structure-agency' question which was emphasised by Giddens (1979) as one of the most intractable issues, yet also one of the most significant. The key question was to explore how human action could be related to the structural properties of society in ways which were neither entirely voluntarist (people's actions are ultimately a matter of choice), nor reductionist (people are forced to act in particular ways because of their circumstances). There had always been a current of sociological research, associated with the interactionist and latterly ethnomethodological tradition, which emphasised the importance of the small scale, face-to-face setting as a key arena for exploring this interface, since it was in small scale settings that people 'acted'. This stress, perhaps especially clear in the dramaturgical sociology of Goffman, was taken up by Giddens, who coined the phrase 'locale' to define 'a physical region involved as part of a setting of interaction, having definite boundaries which

help to concentrate interaction in one way or another' (Giddens 1984: 375). Giddens emphasised therefore that sociology needed to explore the locales which made up 'society'. He drew also on the 'time-geography' of Hagerstrand who explored the movement of individuals through space in the course of their daily lives. The details of Giddens's argument are not important now (see Gregory 1994: 109ff for a detailed discussion and critique). I will only point to a few crucial points. Firstly, following Goffman's lead, Giddens places great emphasis in his work on the idea of co-presence - the distinctive social relations possible in face-to-face settings. Indeed much of Giddens's later work offers a view of modernity whose defining feature is the decline of 'co-present' social relations with the development of what he terms 'time-space distanciation'. The forms of trust and security possible in small settings become more difficult to accomplish as the social world becomes globalised (see Giddens 1990).

For Giddens, then, it is essential to root the agency-structure interface in the routines of daily life, based in particular locales whilst recognising the creation of social mechanisms which 'stretch' across space. Giddens (1985) saw the nation state not as a pre-given, 'natural' entity which defined national societies, but as particular historical creations which, in the era of late modernity, were being supplemented by other globalising forces. This formulation points to a focus on social relations at the local, as well as the national and global levels. It is not the world of bureaucratic national institutions, but of routine daily life, that is at the heart of Giddens's analysis.

Giddens's work, and his formulation of the 'structure-agency' problem, sensitised sociologists to the significance of 'locale'. But his work was perhaps less immediately influential than that within the political economy tradition which also pointed to the distinctiveness of local social relations. Doreen Massey's geological metaphor (see Warde 1985) was perhaps the best known approach (though Harvey's work also commanded attention - see Harvey 1985). Massey's work was influential in guiding the research programme of Cooke, in particular. Massey focused on processes of capital accumulation, and on the way that large firms used space for competitive advantage, locating their different productive activities in specific places to draw upon appropriate labour forces. Localities were not simply passive in the face of these global processes, however, but were able to draw upon local resources to contest economic change, and this again tended to lead to a view of the locality as a battleground where structural capitalist forces met with diverse forms of popular agency. Although Massey's work was in a different theoretical tradition to that of Giddens, it nonetheless ended up endorsing the local for similar reasons - as the crucible in which economic and social relations were contested and possibly changed. Admittedly, Massey's focus was less on the face to face locale but rather on the local economy, frequently understood as local labour market (see eg Cooke 1983). Nonetheless, the same sort of critique of 'national' level analysis followed.

Sub-national research became 'trendy' for theoretical reasons, therefore. But this was not all. The third factor leading to a new interest in locality was innovation in methodological techniques which allowed alternatives to reliance of the national sample survey, or ways of exploiting national surveys which permitted closer scrutiny of spatial variation. A variety of methodological developments were important here. There were new techniques such as entropy maximisation which allowed Johnson and others to analyse aggregate, spatial data in sophisticated ways which allowed them to correlate changes in local voting outcomes with local social change (Johnson, Allsop and Pattie 1989). Multi-level modelling techniques allowed researchers to bring out the significance of contextual effects, as they were able to establish whether particular variables had different effects in varying places or contexts. Ragin's (1987) work offered a view of comparative research which considered the problem of 'small N'

research. Traditionally, most quantitative techniques had relied upon having large sample sizes in order to produce the degree of statistical significance which was thought to be essential for reliable research to be obtained. This tended to sanction the role of the large random sample survey. Ragin's adoption of Boolean logic stressed that rigorous comparisons could be made on only a few cases using algebraic methods which attempted to delineate the possible combinations of variables which might produce given results.

All this having been said, it must be admitted that much of the 1980s research involved a rather more established rediscovery of the tradition of 'case-study' research. This period saw the proliferation and sophistication of the local case study, in both historical (see Gilbert 1992 and the overview in Savage 1990 and Savage 1996) and geographical research. Much of the research in the SCELl programme used fairly conventional methods, ie linked surveys carried out in seven different localities. Nonetheless, it cannot be doubted that this period saw a significant challenge to the hegemony of the national sample survey.

Finally, the fourth factor behind the rediscovery of the locality was the development of new data sets which were able to provide detailed data on small spatial units. Unquestionably, this trend was based on the proliferation of Geographical Information Systems, and more generally on the rapid development of electronic media which permitted the construction and storage of data on very small spatial units. One example of the uses to which data could be put to was the utilisation of NOMIS data to look at changes in local labour markets. A variety of papers produced by CURDS at Newcastle University epitomised the new potential of such data by linking together a variety of economic and social indicators of the economic prosperity of various places within Britain.

To summarise, during the 1980s four different trends made researchers more interested in the 'locality' and the rise of local research that was a major feature of that period. For much of the 1980s these trends appeared to be working hand in hand and pointing in the same sorts of direction. Looking back from the vantage points of the later 1990s, however, it would appear that the promise of this research has not been realised. In hindsight it is clear that the four reasons for looking at the locality always remained different and never congealed into a coherent body of research with enduring lessons for social research. I now move on to explain some of the reasons for this in order to show how interests moved towards an interest in movement and shifts *through* rather than *in* space.

3 THE COLLAPSE OF LOCALITY

Perhaps the clearest indication of the shift from locality was the research reported by the Social Change and Economic Life Initiative. The five volume series exploring various facets of economic and social change contained a variety of findings, but about one aspect they were largely unanimous. There were actually only relatively minor differences between the various localities they studied, despite the fact that they had been chosen as contrasting and differing places. This conclusion came out particularly strongly in their volume on unemployment where Marsh and Vogler (1994: 61) stated baldly that 'The six study towns, which ended the long post-war boom looking strikingly different and bearing the birthmarks of their very different early histories, emerged from the recession in the 1980s looking much more similar in economic terms'. They noted that economic trends in the 1980s had tended to create more homogeneous manufacturing profiles across Britain, that the service sector tended to have relatively uniform employment, and that unemployment rates tended to have converged. A few of the volumes from the SCELl series did include local case studies, but these were nearly always designed as illustrations of the broad picture revealed by national level analysis. And it

was striking that some of the reviews of these studies actually suggested that it would have been better to have provided a national survey rather than to rely on local studies (Scase 1995).

The research of the 16-19 Initiative also reported broadly similar outcomes. Despite the fact that earlier research had mapped considerable local variation in the transition from school to work, Banks et al (1992: 179) concluded that 'local labour market conditions had little impact on proportions staying in full-time education or leaving at 16'. Insofar as there were significant differences, they seemed due to institutional differences - such as the differences between Scottish and English education systems, as social and economic ones.

The only research initiative which championed the cause of locality with any enthusiasm was the Changing and Urban Regional System initiative. Cooke (1990: 296), took the results of his research to proclaim that 'localities are actively involved in their own transformation, though not necessarily as masters of their own destiny. Localities are not simply people or communities: they are the sum of social energy and agency resulting from the clustering of diverse individuals, groups and interests in space'. But Cooke's own proclamations were not entirely borne out by the case studies reported, where the dominant impression was of localities being restructured by all powerful outside interests and with relatively little significant local intervention. Admittedly Pickvance (1990) used some of the case study material to show that local initiative was important in finding ways to influence growing numbers of bodies, including those at the European level. But, in fact many of the case studies reported by Harloe et al (1990) also testified to the declining powers of local authorities and the concluding sentence of the book where Urry contrasted the extensive local political action of the US with that of Britain seems clear enough:

'The danger is that in late twentieth century Britain there are inhospitable "conditions", combined with insufficient "resources" which will mean that it is well nigh impossible to devise locally specific policies, at least under the sway of local authorities' (Urry 1990: 204).

In general, then, the return to locality appeared to have produced a series of answers suggesting that actually, and with various qualifications, localities were not very important. What remains, then, of the 'return to locality'? The important point to bear in mind is that the optimistic air of the 1980s, when descriptive, theoretical, methodological and data considerations all seemed to be pointing in the same direction have increasingly 'come apart' and now tend to point in very different and contradictory directions. Let me review recent developments under each of these four heads.

Starting with descriptive issues, the most striking trend in the 1990s has been the erosion of local differences. In precisely the sorts of areas where there had been growing spatial variation in the 1980s, so the 1990s saw growing national homogeneity. This was largely due to the fact that the brunt of the recession which broke in 1989 was felt in London and the Home Counties. It was precisely the region which had prospered most in the 1980s and which had hence 'pulled away' from other parts of Britain, which suffered greatest in the early 1990s and so came increasingly into line with the rest of the country. Equally, there were some distinct signs that the expansion of the service sector which had been so dramatic in the South East during the early 1980s had to a modest extent 'trickled out' to other parts of Britain (Thrift and Leyshon 1992). This trend towards economic equalisation was acutely felt in unemployment rates, where rates in London and the South East soon came to mirror national trends fairly closely, and also in house prices, since the largest falls in house prices were felt in these areas. The 1992 General

Election saw a considerable 'nationalisation' of voting patterns, with the Conservatives falling back relatively in their heartland but improving their performance in places such as Scotland where they had previously been extremely weak.

In short, Britain in the 1990s was becoming rather more uniform, at least as measured by some of the indicators which had appeared to mark regional differences most markedly. Of course, one should not exaggerate. It is possible to find some areas of growing polarisation, such as rates of long term unemployment. Nonetheless, the general picture in the 1980s is that Britain was becoming more uniform, an impression strengthened by trends such as the spread of familiar multiple retailers to all parts of Britain, of which the northward march of Sainsbury's was one of the most visible reminders. In retrospect, what social scientists had taken to be an epochal change in the nature of the nation state, with the breakdown of Fordist modes of regulation, may in part simply have been a short term cyclical phenomenon.

These shifts took away some of the interest in issues of spatial difference. Yet in some respects they were less important than trends in social theory. For there is no question that the later 1980s and early 1990s saw an explosion of interest in spatiality in social theory. However, as I want to show, this renewal of interest in space led to very different conceptions of space to those which were developed in Giddens' s work around the notion of locale, and ultimately offered a much more radical and unsettling analysis which was difficult to reconcile with empirical social research. Or, to put this point another way, the interest in space was partly a critique of notions of locality. I will try to summarise an extremely complex body of theory to emphasise how this came about.

A number of points can be made here. Firstly, it became clear that Giddens's structurationist approach did not mesh well with 'locality' research. The main problem was that of scale. Giddens focused on the locales of co-present, face-to-face individuals rather than the larger social spaces comprising towns or localities. Furthermore, Giddens's focus on locales did not necessarily involve fixed places - the passengers on a moving airplane could still be part of a distinct locale even though they were demonstrably not part of a specific locality.

This issue was related to wider theoretical doubts about how structure and agency could be investigated using local case studies. Attempts to find distinctive local mechanisms which showed how specific features of the local social structure gave rise to particular social, political or cultural affects ran aground. A major problem was that counterfactuals could readily demonstrate that similar local social structures might have very different effects. Thus, a homogeneous, working class town might be expected to be a radical, militant, town as workers could unite together readily, but, equally, it might be expected that the lack of an obvious class enemy in a homogenous town might lead it to a more passive political outlook (see Savage 1996 for a full discussion of this point).

More generally, it also became clear that local mechanisms were not the only ones producing local outcomes, and that frequently it was the way that social processes stretched across space that was decisive. Thus, many cases of working class militancy were not due simply to the nature of local social relations, but the ability of workers to find ways of linking local processes to wider ones, for instance by organisations that went beyond the locality (such as trade unions, political parties etc). In recent times it can be argued that local political actions depend very strongly on the ability of activists to tap into networks which stretch well beyond any particular locality - such as is evident in campaigns against road building, local environmental campaigns, and so forth.

All this leads to the recognition that space is important not as fixed places with clear boundaries, but as fluidity and mutability. This relates to the renewed interest in space that came from claims that the development of post-modernism marked a new attention to space over time (Soja 1986). Spatiality was held to be of crucial importance in an increasingly globalised world. In one of the best known accounts of recent economic and political change, David Harvey (1988) examined the changing space-time relationships of post-Fordist capitalism, and examined how the hypermobility of capital over space led to new forms of cultural dislocation.

The most important aspect of this renewal of interest was the argument that spatiality undermined place, if places were to be understood as spaces with fixed boundaries, as they conventionally had been in most 1980s locality research. Spaces with fixed, Euclidean boundaries were seen as being a particular cultural product of Western modernity, and emphasis was placed instead on flows and mobility as being central to spatial dynamics. One interesting example of this point came from the elaboration of a 'post-colonial' paradigm. Here, writers sympathetic to the Third World struggles noted that although the formal, Euclidean spaces of Third World countries had been formally de-colonised, nonetheless myriad forms of domination and control continued. More radically, in place of the Western 'centred' view of space, which focused upon the agency of co-present individuals in fixed spaces, writers such as Gilroy (1993) emphasised the themes of exile, of places of loss and absence, which characterised those whose history was one of dispossession and denial. Gilroy's book, *The Black Atlantic* (1993) thus offered an account of black culture which refused to locate it in a distinct 'homeland' of active, co-present individuals, but placed it in the Atlantic, between the European, North American and African axes which together and in articulation had formed black cultures.

In short, the problem was that the concept of locality, or case study area, to use the more innocuous term of Duncan (1989), appeared to draw attention to a fixed, bounded, spatial arena, when in fact it was the way that meanings transcended fixed boundaries which has increasingly captured the imagination. Theoretically the most important source of inspiration for this view is the work of Henri Lefebvre, whose book *The Production of Space* was translated into English in 1991 and had a major effect (see also Shields 1991). Lefebvre's main point was to point to the way that abstract space was related to the commodification inherent in capitalist social relations, and that it was necessary to discover other, imaginary, ways of thinking about space as a means of resisting the logic of capitalist society.

It is clear, therefore, that contemporary appreciations of space render it in ways which try to avoid the notion that spaces are empirically observable, bounded places in which things happen. It is the possibilities of space, the flux and mobility of people, ideas, goods etc which are thought to be central for the constitution of identity. Giddens's views of the locale as a site where structure and agency met look increasingly banal and simplistic and no longer appear to offer a particularly appealing account of these relations. Structuration theory itself has been exposed to devastating critique (Mouzelis 1995; Archer 1995) by writers showing that Giddens's attempt to replace the 'dualism' of structure and agency with a 'duality' ends up by eliding issues of power and domination.

These points link into a third problem in the 'locality' approach, concerned with the methodological issues in studying fixed spatial units. The problem, put simply, was the methodological difficulty in putting socially significant boundaries round particular places. The problem of using administrative boundaries to define places were, of course, familiar, since they tended to be socially arbitrary, and in the course of the 1980s researchers spent much time exploring how more meaningful boundaries might be drawn. Undoubtedly the best attempt was the idea of 'travel-to-work areas', or local labour markets. These places attempted to delineate

boundaries around areas where people typically lived and worked, and therefore in a sense, live their daily lives. Researchers such as Champion et al (1988) provided a sophisticated atlas of British society which used local labour markets as their bases. However, for all the attractions of this idea, problems remained.

For one thing, the idea of the local labour market worked far better for small towns with distinctive rural hinterlands than they did for large urban sprawls. In other words, they worked better for older types of urban settlement than for newer ones. One main problem was the South-East of Britain which included the Greater London area but also a large semi-rural and suburban swathe. Was this all one large 'local' labour market of 20 million people? It was notable indeed that all the locality research in the 1980s avoided the South East of England (excepting some unusual outliers such as the Isle of Thanet).

For another thing, different social groups have different local labour markets. At one extreme, professional and managerial workers are able and prepared to travel long distances to work, and indeed, many of the buoyant areas of the 1980s, such as the 'M4 corridor' were characterised by high commuting distances. On the other hand, women employees in semi or unskilled employment, especially if they have children, and without access to a car may have a local labour market which is very small (see generally on these points Duncan and Savage 1989). There is, therefore, no one local labour market which delimits an entire local population, but rather, a fragmented series of local labour markets for different social groups.

For all these reasons, the notion of discrete spatial units, such as localities, with given effects has become difficult to sustain. For both theoretical and methodological reasons it appears more helpful to examine specific social processes of interest (eg health outcomes, poverty, political alignments) and then consider the spatial dimensions of this, which may very well have different spatial characteristics to other issues. There is also more advantage to be had by focusing on space as forms of movement rather than spaces as fixed places. It can be argued that forms of spatial fixity are both temporary and also the product of factors which stretch beyond any particular place. In this respect the lessons of 'locality' research appear clear - place localities in wider networks and do not fetishise places themselves.

CONCLUSIONS

We can see, then, that the place of space in social research is currently in turmoil. In the 1980s all sorts of intellectual currents led researchers to become interested in the 'local' level. In the 1990s the interests of different communities of researchers have fractured. Theorists have criticised notions of fixed places with clear boundaries and have moved towards an interest in the cultural production of space. Researchers have also found it increasingly difficult to define the boundaries of places. Descriptively, local differences seem less important. The challenge, then, has become one of exploring spatiality less as the fixed properties of places defined in terms of euclidean spaces, and more in exploring processes of movement through space.

REFERENCES

- Adam B (1990), *Time and Social Theory*, Oxford; Polity
- Anderson P (1974), *Lineages of the Absolutist State*, London; Verso
- Banks M, I Bates, G Breakwell, J Bynner, N Emls, L Jamieson, K Roberts, *Careers and Identities*, Milton Keynes, Open University Press
- Champion AG, AE Green, DW Owen, DJ Ellis and MG Coombes, *Changing Places*, London; Edward Arnold
- Collier B (1993), *Critical Realism*, London; Verso
- Cooke P, 'The local question - revival or survival?', in Cooke (ed)
- Cooke P (ed), *Localities*, London; Unwin Hyman
- Dex S (1985), *Women's Occupational Mobility*, Basingstoke; MacMillan
- Duncan SS (1989), 'What is a locality?', in R Peet and N Thrift (eds), *New Models in Geography*, London; Unwin Hyman
- Duncan SS and Savage M (1989), 'Spaces, Scale and Localities', *Antipode*, 21, 3, 176-206
- Esping-Anderson G (1990), *The Three Worlds of Welfare Capitalism*,
- Gallie D, Marsh C and Vogler C (1994), *Social Change and the Experience of Unemployment*, Oxford; Clarendon
- Gregory D and Urry J (eds) (1985), *Social Relations and Spatial Structures*, Basingstoke; MacMillan
- Giddens A (1979), *Central Problems in Social Theory*, London; MacMillan

MULTILEVEL APPROACHES TO MODELLING CONTEXTUALITY: from Nuisance to Substance in the Analysis of Voting Behaviour¹

Kelvyn Jones

INTRODUCTION: CONTEXT AND COMPOSITION

Research on voting behaviour has long been concerned to assess the importance of context (Ennis, 1962; Wright 1977). A major question is whether observed differences between places are merely an artifact of within-place characteristics, that is the social and demographic composition of a place. To help place ideas, consider voting behaviour in the South Wales Valleys. It is well-known that people there have a strong and long-term tendency to vote Labour. This may be interpreted as a contextual effect so that there is something about the social and economic milieu of this area that produces a distinctive political culture. It could also be argued, however, that this is nothing more than the result of the class composition of the area. Individuals of low social class have a strong tendency to vote Labour *wherever* they live.² People in the Valleys are predominantly drawn from the lower social classes, and the high-level of Labour support in the area merely reflects this.

Some are convinced that any apparent contextuality is merely the result of composition:

‘contextual variables have little or nothing to add to explanations of individual political behaviour based on individual variables’ Tate (1974, 1662);

while others have their doubts:

‘election surveys have to face up to the challenge posed by ecological accounts of voting patterns.....survey researchers cannot afford to treat the ecological evidence as an aside in explaining electoral behaviour’ Scarbrough (1987, 241).

It is important to realize that these differences are in part a result of research design and technique, with those dismissing the importance of contextual effects basing their empirical support on large-scale, cross-sectional surveys of individual voters.³

¹ This paper was originally prepared for an ESRC seminar on the Role of locality and spatial effects held at the University of Manchester and also at the Political Geography Speciality Group meeting - "A critical examination of methodology and theory in electoral geography" held at the Association of American Geographers Annual Meeting, Charlotte, April 1996.

² It may also be argued that these political differences are created not through some sort of social miasma but through differential mobility; a longitudinal design would be needed to assess the claims of such an explanation.

³ Another major reason for the differing findings is the common practice by those who use survey data to include attitudinal variables at the individual level as an explanation of actual voting. To claim that contextual effects then disappear is, to say the least, problematic. It is like saying there is no geography of death when we take account of those who are gravely ill.

FROM NUISANCE TO SUBSTANCE

For reasons of cost and efficiency, most if not all large-scale probabilistic surveys adopt a multistage or clustered design. Typically this involves a three-stage design, so that constituencies are first selected, then wards, and only then individuals. Such a design generates a 3-level hierarchical structure with individuals at level 1 nested in wards at level 2, which are in turn nested in constituencies at level 3. Individuals living in the same ward can be expected to be more alike than a random sample, that is they are autocorrelated, and consequently such clustered samples do not contain as much 'information' as simple random samples of similar size. It is well known (Skinner *et al*, 1989) that ignoring this autocorrelation results in incorrect estimates of precision, standard errors, confidence limits and tests. There is an increased risk of finding differences and relationships where none exist, and of building unnecessarily complicated models. From this perspective the convenience of a hierarchical design becomes a nuisance in the analysis, and much effort has been spent in both measuring this 'design effect' and correcting for it.

The multilevel perspective, to be developed here, is radically different (Goldstein, 1987, 1991b, 1995; Jones, 1993a, Hox and Kreft, 1994). The hierarchical structure is seen not as a result of multistage sampling, but rather the *population* itself is conceptualised to have a complex (often hierarchical) structure. Individuals, wards, and constituencies are seen as distinct structures in the population which may be measured and modelled. As a result of supra-individual contextual processes operating at a ward or constituency level, individuals within a unit will tend to be more alike than those in different units. The differing levels are then seen as an integral part of the population structure that needs to be properly modelled. In the terminology of survey research, with multilevel models the standard errors are automatically adjusted for the design effect and associated autocorrelation or intra-class correlation that results from the hierarchical structure. But there is more to it than just technical improvements as a multilevel analysis is able to get some purchase on modelling *contextuality*. Thus, such models are not only able to model between-individual variation (at level 1) but also between-place variation (between-ward and between-constituency variation at levels 2 and 3 respectively). Consequently, this higher-level variation is not a nuisance but of key substantive importance.

Previously, researchers have been on the horns of a dilemma. They have had to work at *either* the level of the aggregate (as adopted by many geographers) *or* the individual (the preferred choice of many political scientists). Choosing to work at the aggregate level lays one open to the charge of the ecological fallacy (Robinson, 1950) and aggregation bias (Roberts and Burstein, 1980), while choosing to work at the individual level risks being found guilty of the atomistic fallacy (Alker, 1969). The latter approach misses the context in which individual behaviour occurs, while the former fails to recognise that it is individuals who act, not aggregates. Working at the individual level misses the context of local cultures, while working at the aggregate level fails to capture individual variation fully. The standard statistical approaches to aggregate analysis (such as the calculation of rates and averages and regression modelling) and to dis-aggregate analysis (such as cross-tabulation, logistic and log-linear modelling) cannot deal with these problems because they operate at a single level. Clearly if we ignore any level in an analysis we can say nothing substantive about it, but this is what researchers have been forced to do.

Multilevel models were explicitly developed to resolve this dilemma by working at more than one level simultaneously, so that an overall model can handle the micro-scale of people and the macro-scale of places. Most importantly by distinguishing different levels, multilevel procedures allow relationships to vary according to context. Indeed, there are now a growing

number of examples where multilevel models have revealed contextuality and complexity which would have been hidden by standard procedures. This talk aims to outline and develop this multilevel approach by outlining three graphical typologies which are illustrated by analyses of voting behaviour in recent British general elections.

Figure 1 Varying relationships between voting and income

Figure 2 Between-place heterogeneity

Figure 3 Dotplots of the higher-level distributions underlying Figure 1

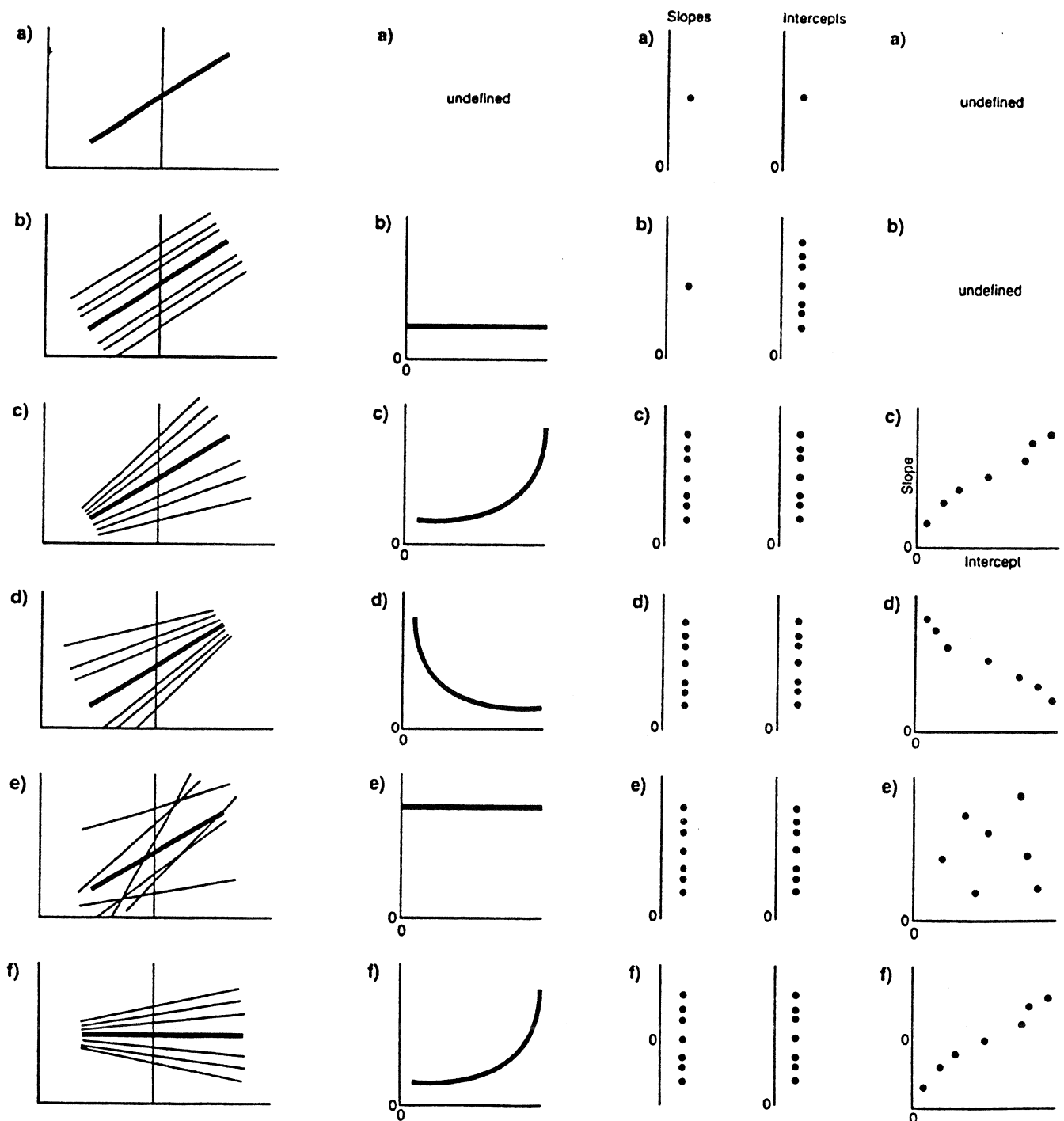
Figure 4 Scatterplots of the higher-level distributions underlying Figure 1

Figure 1

Figure 2

Figure 3

Figure 4



THREE GRAPHICAL TYPOLOGIES

I Graphs of varying relationships

To introduce the basic concepts, I will begin with a two-level model, individuals at level-1 and constituencies at level-2, and consider just two variables. The response variable is underlying propensity of voting for the Conservatives and the individual predictor variable is income (centred around the survey average).⁴ Figure 1 gives a range of possible models. In 1(a) the general voting/income relation is shown as a straight line with a positive slope; 'rich' people vote for the 'right'. In this graph there is no context; place does not matter for voting is conceived only in terms of individual characteristics. This is remedied in 1(b) with each of the different places (six in this example) having its own relation represented by a separate line at a varying 'distance' from the general underlying relationship shown by the thicker line. The parallel lines imply that, while the voting/income relation in each constituency is the same, some places have uniformly higher rates of support for the right than others.

The situation becomes more complicated in 1(c) to (f) as the steepness of the lines varies from place to place. In 1(c) the pattern is such that place makes very little difference for the 'poor', but places have very different Tory support from the 'rich'. In contrast, 1(d) shows relatively large place-specific differentials in voting by the 'poor'. The next graph, 1(e), with its criss-crossing, represents a complex interaction between income and place. In some places there is support for the Tories by the 'poor' while in others the well-off support Labour. The final plot, 1(f), shows that the poor are similar in all constituencies in terms of voting Conservative, but the intentions of the 'rich' vary from place to place. This is similar to 1(c), but this time this difference is achieved by some constituencies having a high rate of Tory support from the 'rich', while in others it is the well-off who vote for the left.

Another way of portraying these varying relationships is to plot, in Figure 2, the between-place heterogeneity, that is the variation at level 2. In (a) all places have the same relationship so that there is no variation between places; in (b) there are differences between places but this is unchanging with income. In (c), the differences between places increases rapidly with income, as they do in (f), while in (d) they decrease with income. The complexity of (e) is characterised by a between-place difference that is relatively large at all levels of income. The differing patterns of Figure 1 and 2 are achieved by varying the slopes and intercepts of the lines. The slope measures the increase in right-wing voting associated with a unit increase in income; since the vertical axis in these graphs is centred at the mean of income, the intercept is the probability of voting Tory for a person of average income. The key feature of multilevel models is that they specify the potentially different intercepts and slopes for each place as coming from a distribution at a higher level. Figures 3 and 4 show the higher-level distributions for the slope and intercept that correspond to the different graphs of Figure 1. Figure 3 shows a 'dotplot' for the distributions of the slopes and intercepts separately, while Figure 4 plots the 'scatter' of the joint distribution. These distributions concern places, not individuals, and result from treating constituencies as a sample drawn from a population.

⁴ The underlying propensity to vote Conservative is usually modelled through a non-linear multilevel model (Goldstein, 1991a).

It can then be seen that:

- 1(a) is the result of a single, or fixed, non-zero intercept and slope
- 1(b) has a single fixed slope, but intercepts are allowed to vary or treated as random terms;
- 1(c) to (f) have sets of intercepts and slopes, that is both the slopes and intercepts are allowed to vary or treated as random terms.

The different forms of 1(c) to (f) are a result of how the intercepts and slopes are associated. In (c) the voting/income relation is strongest in places where there is strong right-wing support by people of average income; a steep slope is associated with a high intercept. Put another way, there is positive association between the intercepts and slopes, as shown in Figure 4(c). In contrast, in Figure 1(d), places where there is strong Tory support by people with average incomes, have a weak voting/income relationship. A high intercept is associated with a shallow slope. Consequently, Figure 4(d) shows negative association between the slopes and intercepts. The complex criss-crossing of 1(e) is the result of the lack of pattern between the intercepts and slopes shown in 4(e). The degree of Tory support from people of average income in a particular constituency tells us nothing about the marginal increase in right-wing voting with income in that community. The distinctive feature of the final plot, 1(f), results from the slopes varying about zero so that in the 'typical' constituency there is no relation between voting and income; in some the slope is positive, in others it is negative. In this latter case, a single-level model would reveal no relationship whatsoever between income and right-wing voting; this average relationship, however, would occur nowhere.

These graphs provide a valuable 'technical apparatus' for discussing geographical variations. In particular, the plots of the intercepts and slopes demonstrate that we can achieve quite different and complex variations by straightforward changes of the underlying structure. The plots of Figures 3 and 4 which characterise the higher-level distribution refer specifically to places and not people. While it is possible to directly measure which way people vote, place differences have to be estimated in a modelling framework for they are not directly observable. A particular strength of the multilevel approach is that higher-level units such as wards and constituencies remain in the analysis as identifiable entities that are not lost in the statistical soup of aggregate analysis, or assumed away as in an individual-level analysis. If the multilevel analysis reveals distinctive contextual influences in particular places, it would be possible to adopt qualitative, intensive approaches to try and uncover the social processes that are operating there. Thus in terms of Figure 4(f), the analysis would reveal the constituency with the strongest positive relationship between income and right-wing vote, and also the place where the relation is the inverse.

The basic concept underlying multilevel modelling is the specification of models at each level and then their combination into an overall model (Jones, 1991a). More specifically, there is an individual-level, micro-model which represents the within-place equation, and an ecological, macro-model in which the parameters of the within-place model are the responses in the between-places models. This simultaneous specification allows for the separation, in a quantitative sense, of the compositional from contextual (Mason *et al*, 1984). The central empirical question concerning contextual variation becomes does the level-2 variation remain significant when a range of appropriate and relevant individual variables (such as income, class,

employment status) are included in an overall model?⁵ At the same time it must be stressed that it is always possible to argue that apparent contextual effects are a result of the mis-specification of individual effects (Hauser, 1970). However, if higher-level variation remains substantial after taking into account 'many' individual factors, then it is not unreasonable to conclude (albeit and always provisionally) that there are genuine contextual differences.

It is important to realise that these substantive advantages are made within a robust technical framework. From the graphs in Figure 1 it appears as though a separate line is fitted in each constituency. This would be equivalent to procedures based on traditional single level OLS regression in which the fixed part of the model is expanded to include a slope and intercept term for each individual constituency (that is 'Analysis of Covariance'). If there were 200 constituencies, however, this approach would involve fitting a model with 400 parameters and a very large sample size would be needed to obtain reliable estimates.⁶ Traditional quantitative approaches to contextual analysis are, therefore, highly inefficient. In contrast, multilevel techniques involve estimating the statistical characteristics of the higher-level intercept and slope distributions for the population using the constituencies as a sample. Consequently, it is the random part of the model that is expanded and, in the example above, a multilevel analysis would involve estimating only two fixed part terms giving the average intercept and slope across all 400 places and three random terms summarising the variability between specific places. It should be noted, however, that *predictions* of place-specific intercepts and slopes can be obtained once the overall between-place variance has been estimated. Since these predictions are made using the entire sample of places they are more efficient than those from a traditional approach in which each place is estimated separately.⁷ Since the multilevel approach involves estimating more than one random term, traditional OLS estimation strategies cannot be used and special multilevel modelling software is required (such as *Mln*, Rasbash and Woodhouse, 1995).

II Cross-level interactions

The distinctive feature of the second graphical typology is that an additional predictor variable is included in the model that refers not to individual characteristics, but to the nature of the constituencies. Again there is a two-level model (individuals in constituencies) but this time the response is voting Labour, while the individual predictor variable identifies the working- as opposed to the middle-class, and the constituency variable is the ecological characteristic, the percent of the voters who are working-class. A very wide range of differing results involving these three variables are possible of which a selection are shown in Figure 5. The vertical axis

⁵It is also possible (Jones and Bullen, 1993) for genuine contextual effects to be hidden or masked by not allowing for social and demographic composition. Such a result can occur, for example, when a place with a genuinely low Tory support rate has relatively high numbers of high-status individuals who nationally have a high probability of voting Conservative.

⁶ It would also require for effective estimation that within each constituency, there was a 'reasonable' range of values on the predictor variable; a constituency specific OLS slope parameter would be impossible to estimate if all individuals in a constituency had the same income.

⁷ It must be stressed that this formulation of place differences as random effects implies that each place is not assumed to be a 'separate entity' but rather is seen as coming from a distribution. If it was believed that a particular place (or set of places) were not part of a single 'national' distribution but was in some way untypical, this could be accommodated by including appropriate differentiating factors as fixed terms in the model. There would be no 'pooling' of information across the typical and untypical sets of places. In practice, this approach can also be adopted when a particular place is an outlier thereby unduly inflating the size of the between-place variance.

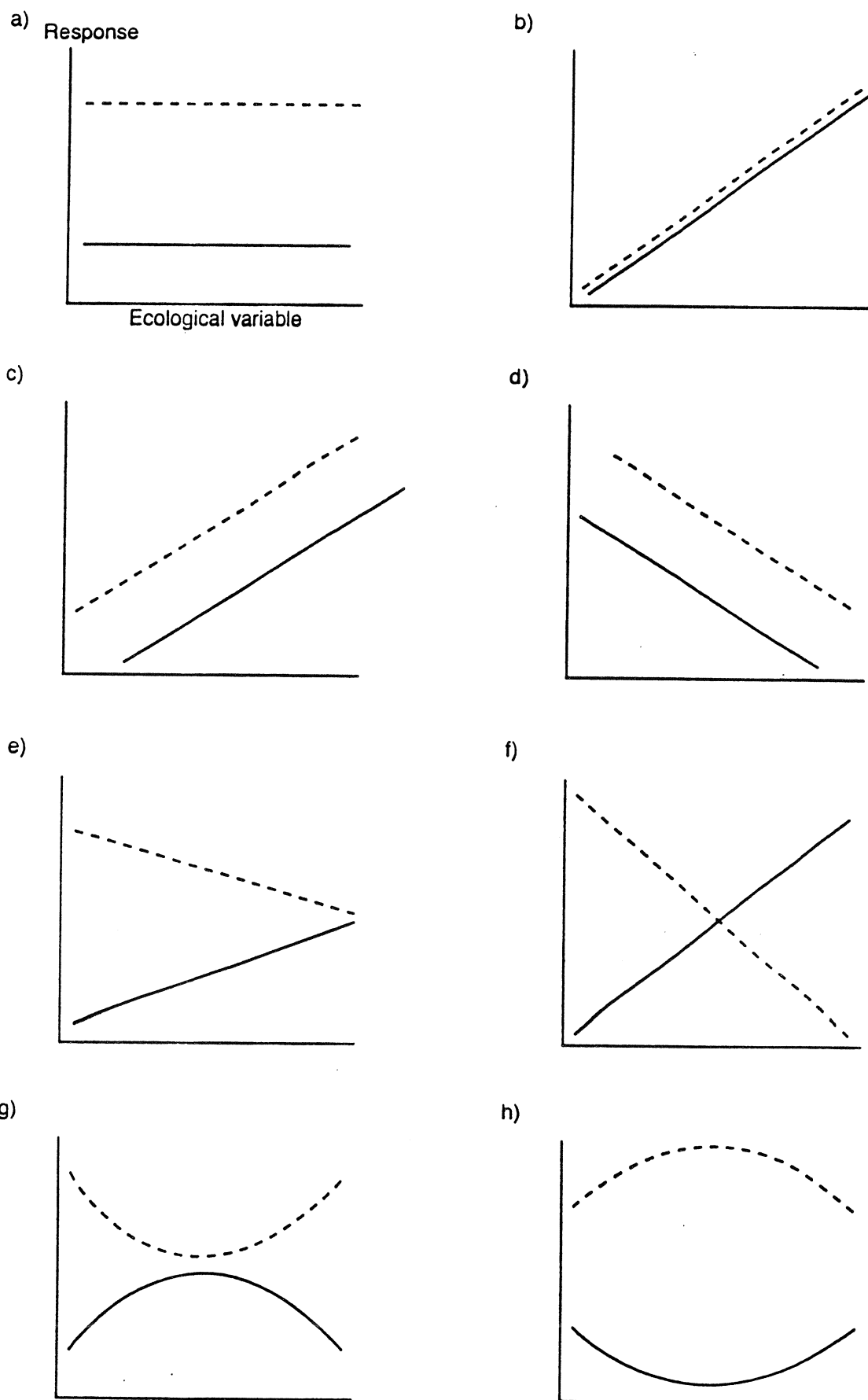
represents the response, the horizontal axis the ecological variable, while the lines on the graph represent different types of individual (the dotted line and continuous lines represents low-status and high-status individuals respectively). Thus, Figure 5(a) shows that there are marked differences between individuals but no ecological effect, while (b) represents the converse: little difference between types of people but a substantial ecological effect. The parallel lines of Figure 5(c) and 5(d) represent the cases when both the individual and ecological effects are marked: the former represents what Miller (1978,266) calls a 'consensual environmental effect', while the latter shows a 'reactive effect'. In the consensual case, both the ecological and individual effect of class are operating in the same 'direction' so as to reinforce each other. The strongest Labour support coming from lower-class individuals in 'lower-class' areas. But in the reactive case, while it is generally lower-class individuals who vote for the left, they are less likely to do so when resident in areas with a high percentage of lower-class. Examples of both types of effects are provided by Huckfeldt (1984, 400). Figure 5(e) represents what is known in the political science literature as the Przeworski environmental effect. The graph shows a model in which the environmental effect is reactive for the lower-class but consensual for the middle class (Przeworski and Soares, 1971). Figure 5(f) represents the case where the cross-level interactions are strong enough to invert the individual-level effects; while in (g) and (h), non-linear interaction terms are of importance so that either the smallest or largest ecological effects are found at 'middling' levels of the ecological variable. While such concepts of environmental influences have long been of interest to empirical researchers (Van den Eeden and Huettner, 1982) and are the subject of much conceptual interest (Jones, 1993b), it is only with the development of the multilevel model with its micro and macro equations that the full complexities of such relationships can be effectively analyzed.

III A multitude of multilevel models

The third and final graphical typology (Figure 6) depicts a wide range of multilevel structures (Jones and Duncan, 1996). The two-level structure of 6(a) can be readily extended to the three-level structure of 6(b) with individuals at level 1 nested within households at level 2 and places at level 3. Variables can be included at each level so that the relations between an individual's voting and age, for example, can be examined in the context of household income. The extension of the framework to many levels (individuals, polling districts, wards, constituencies, regions...) is important for, as pointed out by McAllister (1987b, 47):

'The level at which contextual effects are supposed to operate has never been satisfactorily resolved. The bulk of the British literature assumes that the effect is confined to a parliamentary constituency (containing on average some 80,000) people, but at least part of this choice reflects the spatial level at which disaggregation is possible in a sample survey, and the availability of census statistics'.

Figure 5 Individual and ecological cross-level relationships: a range of people and place interactions (see text for explanation)



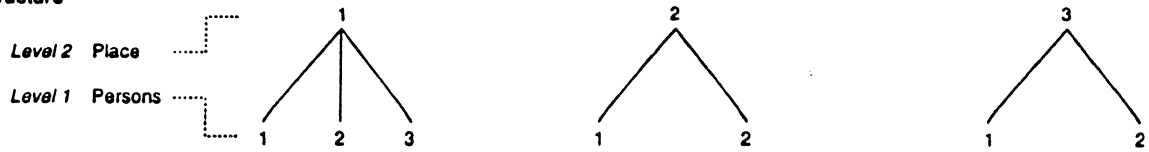
Until very recently the number of levels in available software has been restricted to only 2 or 3. However, with the release of *Mln*, the number of levels can be set dynamically. Consequently, the richness and size of available surveys is likely to be a greater restriction than software. For example, in the British Electoral Studies only one ward has been sampled in each constituency, so that the variation of these two distinct levels is confounded in the survey design.

Changing contexts and changing behaviour are important too, and the framework handles temporal setting as well as place as context. Two possibilities arise depending on the level of unit that is repeatedly measured. If repeated cross-sectional surveys are undertaken then constituencies could be monitored every year, producing a structure with individuals at level 1, years within constituencies at level 2, and constituencies at level 3. This is shown in Figure 6(c). When individuals are repeatedly measured in a panel design, the behavioural measurement taken at different times is level 1. These are nested within individuals at level 2, which in turn nest within a further higher level unit such as the constituency. This structure is shown in Figure 6(d). The first case permits the examination of trends within settings having allowed for their compositional make-up. The second case allows the assessment of individual change within contextual settings. Substantively, the multilevel approach allows the flexible specification of variance and covariance structures through which it becomes possible to assess both which sort of individuals and which sort of constituencies 'change' their behaviour. Technically, multilevel analysis is not affected by the restrictive data requirements that have hampered conventional repeated measures analyses (Ware, 1985). Within a multilevel structure both the number of observations per unit *and* the spacing among observations may vary. This flexibility enables efficient use to be made of all the data available. Such models could be used to explore what Warde (1986) has termed the '*electoral mystery*' of increasing individual class de-alignment and simultaneous constituency class polarization (McMahon and Heath 1992), and to develop arguments regarding the geographical variations in class de-alignment (Johnston and Pattie, 1992).

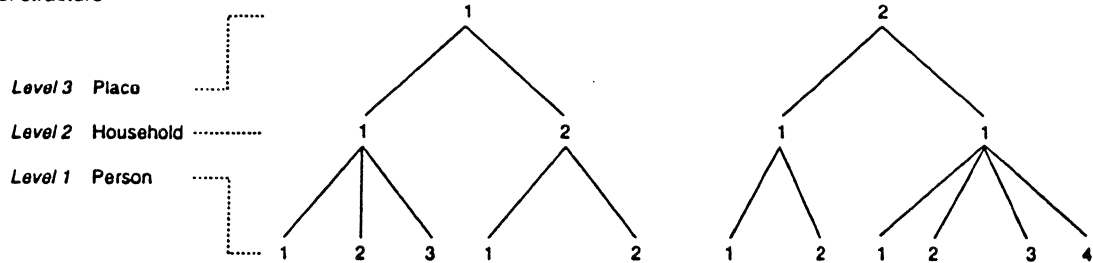
Standard approaches are not well-suited to modelling several response variables simultaneously. Yet, it would be reasonable to examine simultaneously several measures of voting intention (such as intended vote in local and general elections). By extending the multilevel framework to a multivariate model it becomes possible to assess the degree to which different behaviours are connected (Duncan *et al*, 1993a). If we collect information for individuals on each behaviour then we can produce a multivariate, multilevel structure in which level 1 is a set of response variables, one for each behaviour, which nest within individuals at level 2, who nest within constituencies at level 3. This form of multilevel structure is given in Figure 6(e). In substantive terms, two main benefits arise from a multilevel, multivariate approach. First, the behaviours are directly comparable in terms of how each is related to individual-level characteristics. Answers to complex questions can be given: for example, is voting in local elections related to income and socio-economic status in the same way as in general elections? Second, the residual covariance matrix between the set of responses can be estimated at any level, so that it is possible to assess the 'correlation' of voting behaviours both between individuals and between constituencies, and to do so conditional on other variables. For example, if this correlation at the higher level is negative, there will be places which support Labour in the local elections but not in the general, and *vice versa*.

Figure 6 A range of multilevel data structures

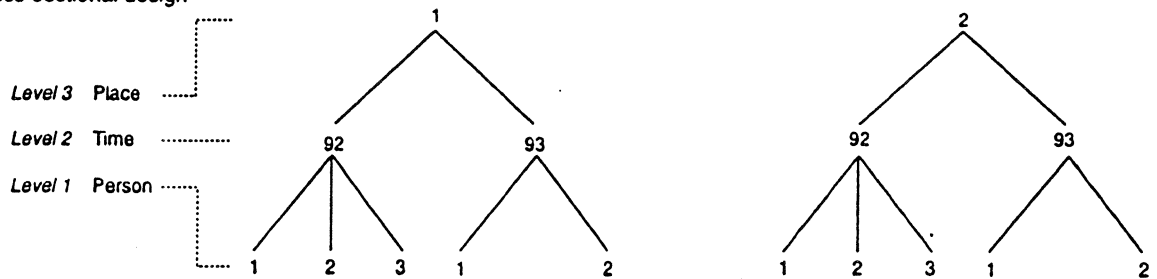
a) Two-level structure



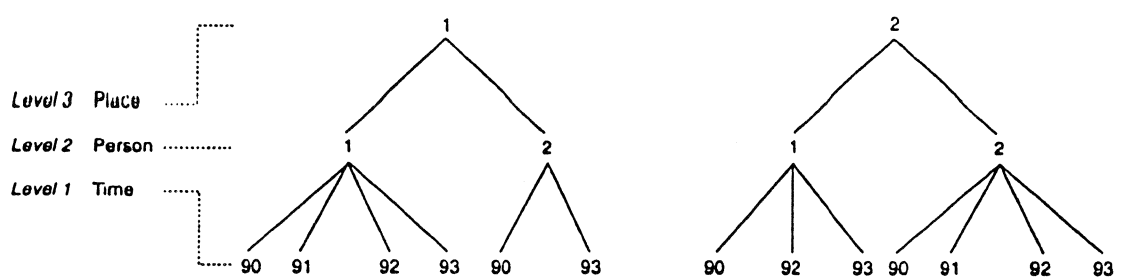
b) Three-level structure



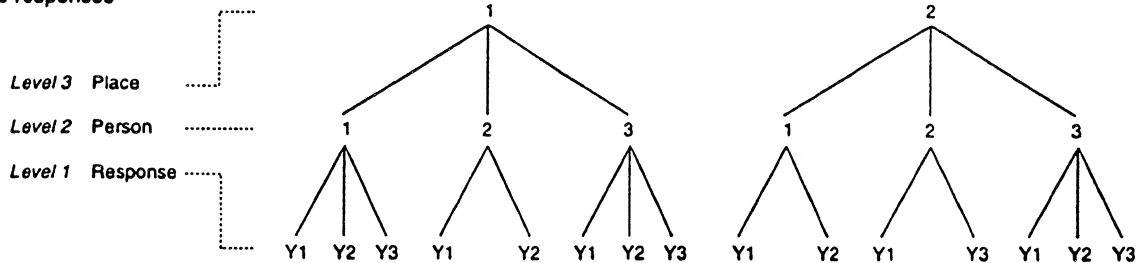
c) Repeated cross-sectional design



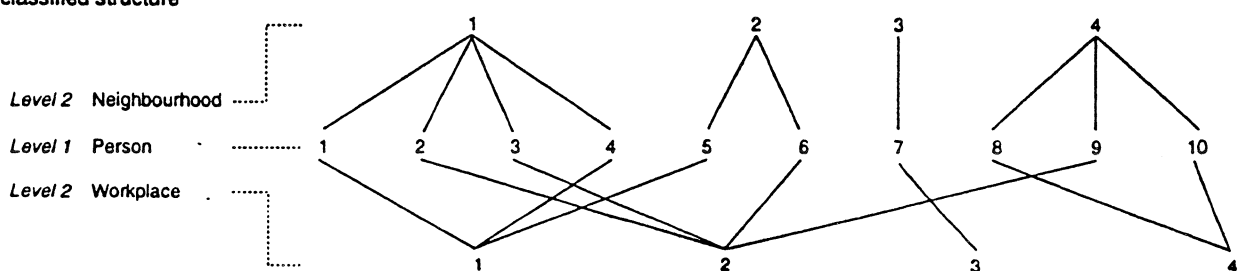
d) Panel design



e) Multivariate responses



f) Cross-classified structure



All multilevel models can use both continuous and categorical data. Importantly, therefore, this multivariate multilevel framework can be applied to continuous response variables, categorical response variables, and also a combination of the two. For example, the multivariate response could be the choice set of voting intentions (Labour, Conservative, Liberal-democrats, Other). This produces what is termed a multinomial, multivariate model and provides a way of modelling multiple response categories (Multilevel Models Project, 1993). Alternatively, when the response is more than one choice set, there could be two sets of binary outcomes: supporting Labour in comparison to other parties; and actually voting for Labour in comparison to other parties. There may be a different set of relationships for one choice set to another. Finally, the responses can be a mixture of both categorical and continuous variables. (Duncan *et al*, 1996). Technical benefits flow in terms of efficiency if the responses are correlated and there are many missing responses as in matrix sample designs.

All these examples have so far been strictly hierarchical so that each lower unit nests exactly into one, and only one, higher-level unit. But it is likely that there will not be a single context but many. The possible existence of multiple contexts can be examined through the cross-classified multilevel model. The appropriate structure is given in Figure 6(f) with individuals at level 1 nested in residential neighbourhoods at level 2, and in workplaces also at level 2. This approach is extremely valuable as it can identify contextual settings which are having a confounding influence. In the example above it may be discovered that what appears as between-workplace variation is in fact really between-neighbourhood variation (Goldstein, 1994). Such models could be used to examine debates in the literature about regional effects (Warde, 1986) and whether these differences reflect 'geography', that is continuous 'blocks of territory' or 'functional regions' that is groups of constituencies with similar characteristics, irrespective of location such as that used by Johnston *et al* (1988).

IMPLICATIONS FOR SURVEY DESIGN

Given this power and flexibility of the multilevel approach, there are now obvious questions about what sort of survey design should be used, how large the number of higher-level units, and how modest the sample size within a higher unit (Jones, 1994). In terms of sample design, Goldstein (1984) has argued that the multistage design is the most efficient one for studying contextual effects. To take an example, and using census wards as the higher-level unit, if a national study was conducted on 10,000 respondents chosen according to simple random sampling, we would anticipate that only one voter would be found on average in each of the wards. The within-ward variation would then be totally confounded with the between-ward variation and no separate estimates of these distinct components would be possible. To obtain reliable estimates of both the within- and between-group variation, we need a compromise between the number of lower- and higher-level units. This is, of course, what is achieved by a multistage design. For a given total sample size, if we allow the number of higher-level units to increase each unit will contain fewer individuals, and we approach the situation of a single-level model, where we are unable to model contextual effects.

To get reliable estimates of place differences we need lots of places. Having many individual respondents provides information on the voting/income relation within a place, but many places are needed to assess the differences between places. While it is difficult to be specific about the required sample size (much depends on the magnitude of the higher-level random effects) there is some guidance from educational research, the area in which multilevel modelling has been most applied. Bryk and Raudenbush (1992, 203) find that with 60 students per school and 160 schools it is possible to have a total of four coefficients random at the school level. Paterson and

Goldstein (1992) suggest a minimum of 25 individuals in each of 25 groups to do useful work; in preference 100 groups are needed. What is not appropriate, and yet this has often been the case in researching locality effects, is hundreds of respondents in five or ten higher-level units. Of course, in some situations it is impossible to follow this 'rule of 25' such as when dealing with households and repeated measures of voting behaviour. In such cases while it is vital that multilevel approaches are applied (because of high levels of anticipated autocorrelation), it would not be sensible nor useful to make inferences about particular households or individuals.⁸ Finally, if there is a need to derive sample-based aggregate variables, then a higher degree of sample clustering than is normal may be employed. For example, if a researcher is interested in cross-level interactions between voting and individual and neighbourhood income, then a sample based estimate of the latter would require a sizeable number of respondents in each neighbourhood.

EXEMPLIFICATION: THE GEOGRAPHY OF VOTING BEHAVIOUR

'Class is the basis of British politics; all else is mere embellishment and detail'; so wrote Pulzer in 1967. If this is correct and still valid, there is no space for geography. This argument has been made in its most sophisticated form by McAllister (1987a,b), and Rose and McAllister (1990) who have maintained that the apparent differences between places in their voting behaviour is simply the result of differential social and demographic make-up, the composition versus context argument outlined earlier. Much of the work in this area has been undertaken using single-level analyses, but this debate can only be adequately addressed by recognising that individuals, constituencies and regions form different levels in a hierarchical structure.

At the same time as the 1987 General Election, a social survey (Heath et al, 1991) was conducted on voting and voters' characteristics; this allowed Jones et al (1992) to explore how voting varies from place to place in a multilevel model. The structure of the data is given in Table 1. The basic finding from this work is that there are substantial differences in voting behaviour between the regions and constituencies even after the differing compositions of the areas are taken into account. Thus, the South Wales region is strongly pro-Labour even when allowance is made for the class, tenure, employment and demographic characteristics of the voters who live there. To put it another way, people of similar characteristics vote differently in different places.

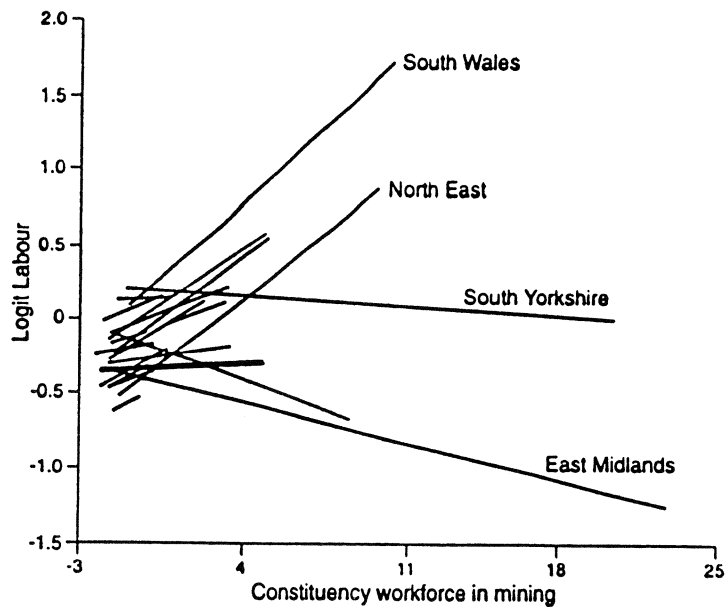
The variations are not, however, adequately described by a simple model in which only the intercepts are allowed to vary, that is a model in which certain places are uniformly pro-Labour for all types of voter (as in Figure 1a). Some of the complexity of the results is conveyed in Figure 7. In the model that lies behind this figure, the constituency-level relationship between mining and voting Labour is allowed to be random at the regional level; the parameter for a level-2 variable is allowed to have a distribution at level 3. Consequently each line on the graph represents the predicted regional relationship between the percentage of miners in each constituency and constituency support for Labour (after allowing for constituencies' composition in terms of age, class and tenure, and its economic prosperity). Traditionally, mining constituencies have been seen as Labour's heartland, and while this is generally true as shown

⁸ That is the higher-level variances can and should be estimated, but the *predictions* for specific individuals or households will have large confidence intervals.

Table 1 Data structure used by Jones, Johnston and Pattie (1992)

- Level 1:** 2281 individual voters
response: choice of Labour as opposed to Conservative
predictors: age, occupational-class, housing tenure, employment status
- Level 2:** 250 constituencies
predictors: unemployment, employment change, workers in the mining industry
- Level 3:** 22 economic regions as defined by **The Economist** in their presentation of election results

Figure 7 Constituency vote/mining relationships varying over region



by the majority of slopes being positive, there are some noticeable exceptions. The most marked contrast is between South Wales and the East Midlands. South Wales is a pro-Labour area, and this support increases as the economy of the constituency is more involved with mining as measured by the 1981 census. The opposite is found in the East Midlands, the most anti-Labour areas are the coal-mining constituencies. The other two notable relationships identified on the graph are the North East which shows a similar relationship to South Wales but from a lower base, and South Yorkshire which shows generally strong support for Labour that does not change radically in the coal-mining areas. Places that appear to be outwardly the same (they all were coal-mining areas) are shown to act quite differently when the analysis is sensitive to place differences. The 1987 election was held quite soon after the National Union of Mineworkers' strike which had received virtually no support in the northern East Midlands mining area known as the Dukeries. Griffiths and Johnston (1991) develop a complex argument for why this place differs from other coalfields in terms of the history of work traditions, cultural practices and social relations.

Table 2 Data structure used for multilevel analysis of 1992 General election

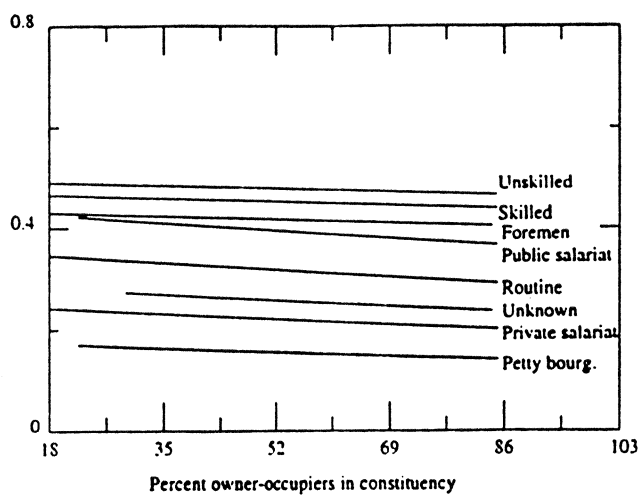
Level 1:	2302 individual voters
	response: choice of Labour as opposed to Conservative
	predictors: age, gender, occupational-class, housing tenure, employment status, income, educational qualifications
Level 2:	218 constituencies
	predictors: % unemployment,
	% employers and managers,
	% local-authority tenure

Another survey was undertaken at the time of the 1992 General Election (Heath et al 1993) and this has been used to examine the nature of people and place effects in a model that allows for cross-level interactions between individual class positions and three constituency characteristics (Jones, Tonkin and Wrigley 1996). The underlying structure of the data set is given in Table 2. The response is again the choice between Labour and Conservative and the model includes a large number of terms for individual characteristics in addition to those used by Jones et al (1992); these include age-sex interactions, income, and educational qualifications. In a random-intercepts model, the between-constituency variance remains large and significant even after these additional variables are included, thus adding further empirical support to the importance of place characteristics. Attempts were then made to account for these level-2 differences by including ecological and interaction variables. Figure 8 shows the obtained results in graphical form; it must be remembered that these graphs portray the results after allowing for the individual characteristics.

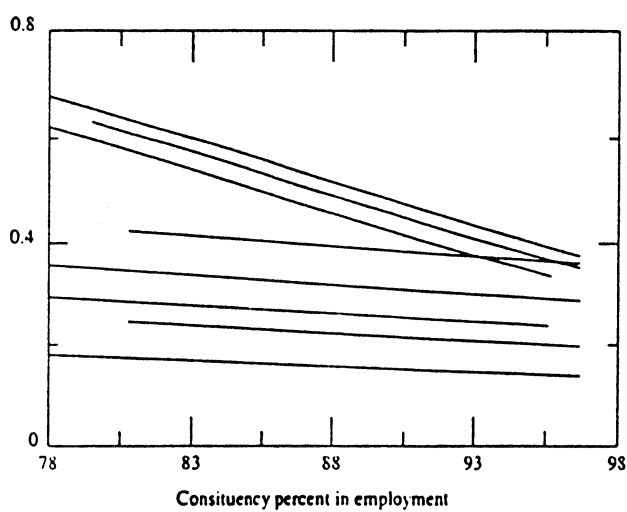
Three constituency ecological variables were included in the model: tenure, measured as percent of owner-occupiers; employment measured as 100 percent minus unemployment rate; and class, measured as percent employers and managers. In order to provide for reliable estimates only linear models were used, and cross-level interactions were based on two categories of class: working class (defined as foremen, skilled and unskilled manual); and non-

Figure 8 Cross-level interactions estimated for the 1992 General Election: individual and constituency interactions

a) Class and housing interactions



b) Class and employment interactions



c) Class and percent employers and managers interactions

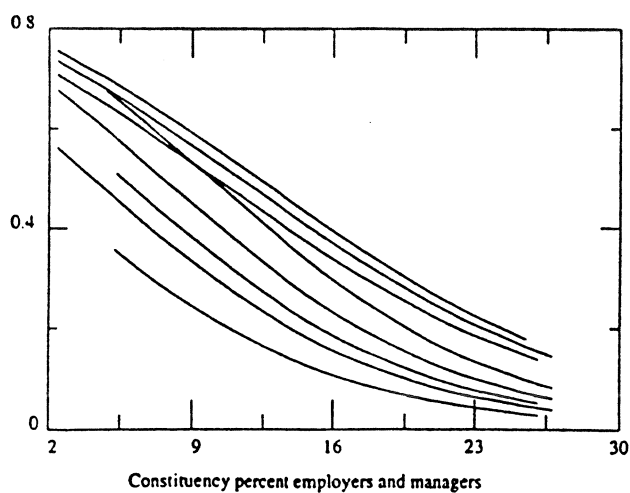
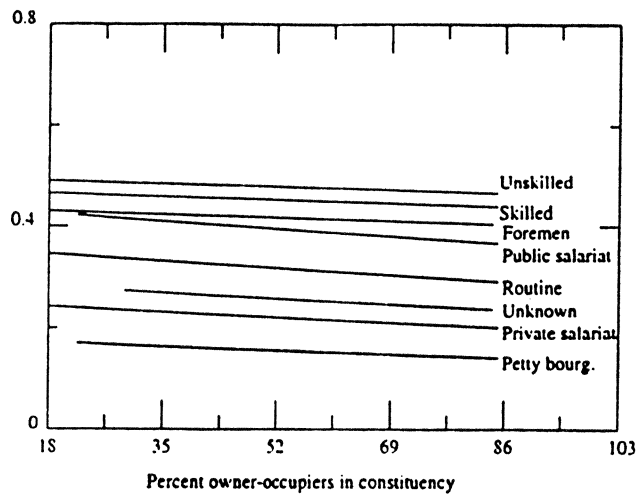
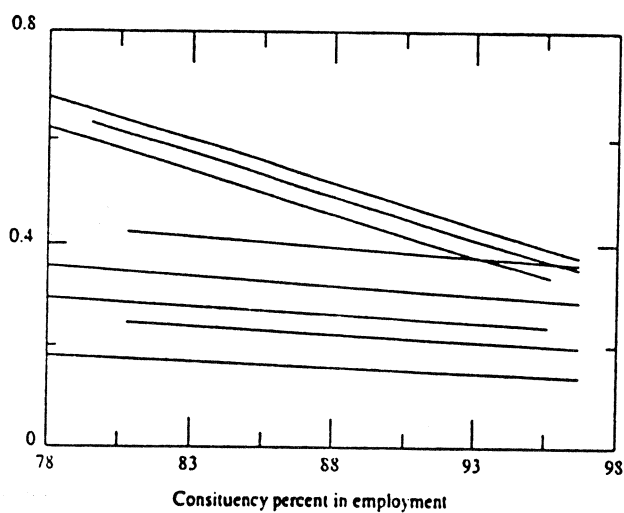


Figure 8 Cross-level interactions estimated for the 1992 General Election: individual and constituency interactions

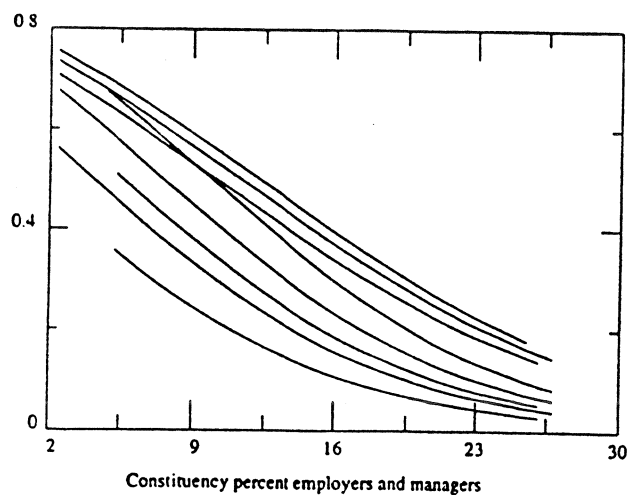
a) Class and housing interactions



b) Class and employment interactions



c) Class and percent employers and managers interactions



working class (public and private-sector salariat, routine non-manual, petty-bourgeoisie and 'unknown'). To aid comparison, all the results in Figure 8 are drawn on probability axes with the same range. Constituency housing tenure is seen to have very little effect. There are marked differences between classes but these do not change in relation to the tenure characteristics of the constituency: the ideal type of Figure 5a is found. These results, go against the theoretical arguments of Savage (1987), but strongly support Curtice's (1988, 143) study of political attitudes:

'working-class people are not influenced by whether they live on a council estate or not. While their own individual tenure position is important, it is their neighbour's class composition which matters, not their tenure position'.

Figure 8b, that for (un-)employment shows a more complex picture. Allowing for the other constituency variables; the probability of voting Labour is related to employment levels in the constituency more markedly for working-class individuals than non-working class individuals. Such results can be related to egocentric and sociotropic voting (Weatherford, 1986). The voters' decision is based on the voters' evaluation of government's performance; voters satisfied with incumbent government both in relation to wider society (sociotropic) and in their own interests (egocentric) are likely to vote for its return. While the non-working class are differentiated by class (personal interests), they remain somewhat immune to the economic situation of the local area. In contrast, the working class are additionally affected by the local economic environment; there is an element of sociotropic voting at the local level. (It must be remembered that individual employment status has also been included in this model.)

The final graph, Figure 8c shows the results for percent employers and managers. There is a very strong ecological effect which takes a consensual form even when individual employment and tenure are included in the model. All classes tend to be anti-Labour in constituencies with high rates of employers and managers; but the most anti-Labour voting is by individuals of high occupational status. There does not appear to be any marked difference in the steepness of the relationships between the working and non-working class. In summary, and contrary to those single-level studies which found no contextual effect, these results suggest that a strong ecology of voting remains after a wide range of demographic and socio-structural variables are included in the models. This ecology most strongly relates to the class character of a constituency defined in terms of employers and managers. Where this group forms a sizeable proportion of the population, more or less everyone, irrespective of their individual class, votes Conservative. Constituencies are more polarised politically than people. These results support the findings of aggregate analyses (Miller, 1978, 1979; Waller, 1983).

The final illustration concerns another debate in the literature about the importance of contextual effects. It focuses on the relative importance of social and geographical contexts in accounting for voting outcomes, and particularly, on whether the differences are more marked for a 'geographical' regionalisation, than for a 'functional' one. According to Johnston et al (1988), the former refers to:

'contiguous blocks of territory'

while the latter represents

‘groups of constituencies with similar characteristics, irrespective of location’.

They argue that both types of grouping are

‘relevant contextual variables for voting behaviour: the first linked to Britain's growing North-South divide ... and the latter to a growing urban: rural divide with the urban sector divided into different functional types’.

Savage, in contrast, argues that (1987, 66) functional differences are becoming of increasing importance in comparison to geographical differences

‘whereas in the past constituencies of a similar type often had different political alignments because of the salience of their local political cultures, this is becoming much less apparent and constituencies of a similar type are behaving in similar ways whatever part of the country they are in’.

In order to assess these different regionalisations, Jones et al (1996a,b) used a cross-classified multilevel model with the structure given in Figure 9 and the variables given in Table 3. Individual voters (at level 1) are nested within constituencies (at level 2) which are nested within geographical regions (at level 3) and functional regions (also at level 3). There are 31 types of constituency based on a large-scale cluster analysis of census variables, and 24 regional groupings of constituencies based essentially on the non- and metropolitan division of Standard Regions.

The results from a random intercept three-level model show that the largest higher-level variance is the between-functional regional variance and the smallest is the between-geographical regional variance. To give some appreciation of the size of these effects consider the estimated probability of voting Labour for the type of person whose characteristics are the most commonly occurring in the sample. This is a 46-year old woman without educational qualifications who lives in owner-occupied household whose head is in employment, receives a 'middle' annual income (£12-20,000) and who is classified as unskilled working class. Nationally this type of person has a 0.49 probability of voting Labour. For the geographical regions, the predicted probability at the extremes ranges from 0.36 in the 'South Coast' to 0.63 in the 'Industrial North East'. For the functional regions this range is from 0.25 ('Scottish rural areas') to 0.75 ('Areas with poorest domestic conditions'). The differences between functional regions remain substantial despite taking account of a wide range of predictor variables. For comparison the greatest fixed effects for individual class are the difference between unskilled manual (0.49) and petty bourgeoisie (0.15) when all the other predictor variables are held at their 'stereotypical' value.

A further model allows the size of the higher-level effects to be differentiated by class. It is found that geographical regions are not differentiated by class, but for the functional regions the results suggest that where the working class vote for Labour is high, there is a tendency for the upper class differential to be negative. In terms of probabilities, the range of Labour support amongst the working class goes from 0.28 in 'Scottish rural' areas to 0.75 in 'areas with poorest

domestic conditions'. For the private salariat, the range is from 0.13 in 'very high status' areas to 0.42 in the 'Scottish industrial' constituencies. The biggest difference between the classes is found in 'Textile' areas where the probabilities for the working class and private-sector salariat are 0.62 and 0.25 respectively. The smallest difference is in 'Agricultural areas' where the probabilities are 0.31 and 0.17. At the constituency level, a similar pattern is found.

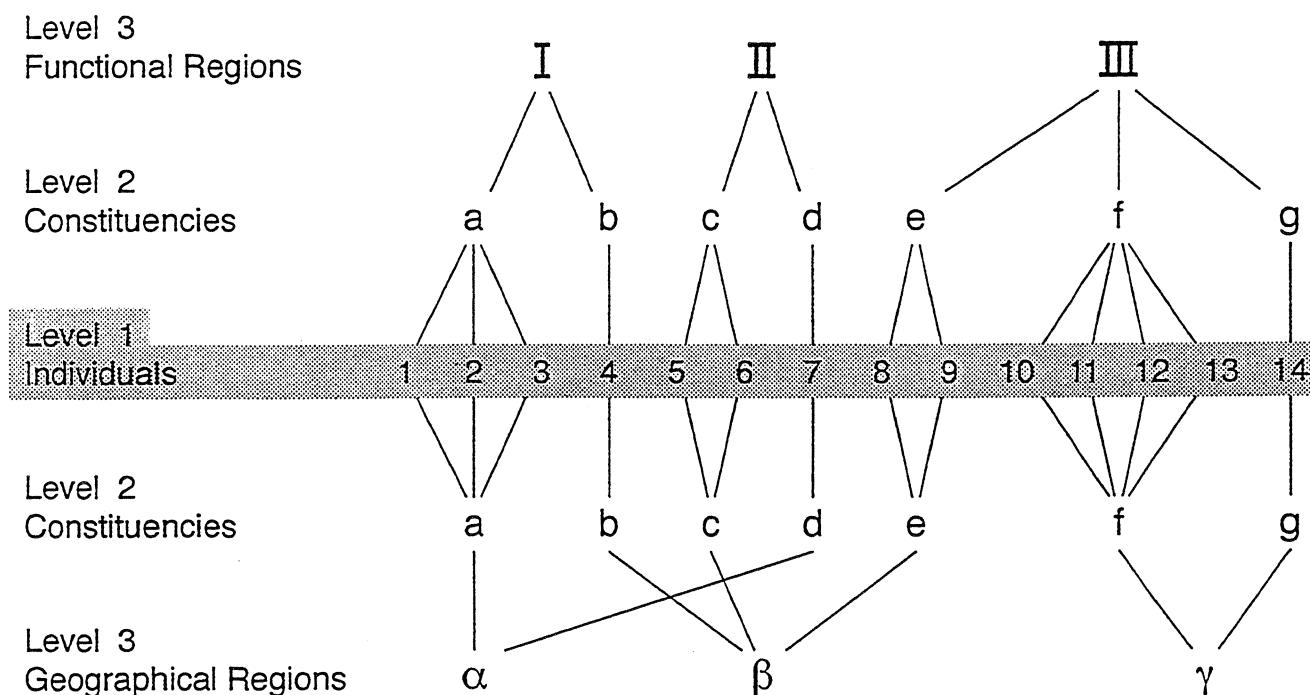
Table 3 Data structure used for multilevel cross-classified analysis of 1992 General election

Level 1: 2275 individual voters
 response: choice of Labour as opposed to Conservative
 predictors: age, gender, occupational-class, housing tenure, employment status, income, educational qualifications

Level 2: 218 constituencies

Level 3: 31 functional regions and 24 geographical regions

Figure 9 A multilevel cross-classified structure for analysing 'regional' variation



This suggests a complex geography of constituency preference even after allowing for demographic and social characteristics of individuals, with a tendency for places that are pro-Labour for the working class to be relatively even more anti-Labour for the upper class. In summary, there is evidence that contextuality is complex and differentiated. While differences between geographical regions are not great, the differences between functional regions and constituencies are substantial and the size of effect is differentiated by class.

THREE CAVEATS

As with all technical developments, there are always real dangers of 'overuse' and over-sell'. While there are a range of technical problems that remain in the use of multilevel models (Jones, 1991b), I want to point out three important substantive problems that beset current practice.

The first problem relates to defining the higher-level units. Multilevel models were first extensively developed for pupils in schools, so that the institution was the readily-identifiable higher-level to which individuals belonged. Applying the approach to places (Duncan *et al* 1993b) is, of course, fraught with difficulties, for as Massey (1991) points out:

'localities are not simply spatial areas you can easily draw a line around';

as they must be must be seen as

'the intersection of sets of locales'.

While it may be possible to use Openshaw's (1977) automatic zoning procedure in tandem with cross-classified multilevel models in an overall GIS framework, this is not a problem that can be overcome by a technical fix. While the levels of constituency (and individuals) seem entirely reasonable ones for voting behaviour, it is a much more open question how to define neighbourhoods and regions.

The second problem relates to the fact that the overwhelming majority of electoral studies have adopted a static rather than a dynamic approach. Consequently, such work has been unable to distinguish the 'breeder' hypothesis (that social milieux breeds social attitudes) from that of the 'drifter' hypothesis, whereby differentials in geographical mobility create spatial patterns of voting behaviour. While the type of structure discussed earlier (Figure 6d) allows the examination of repeated measures of voting, this is not enough to tackle changing contexts. Voters may be socialised in their youth in a particular context and subsequently move to another contexts to have these views challenged or reinforced. This necessitates a non-hierarchical structure. As yet there are no examples of such research. This is due to a lack of suitable data sets and because, until recently, the calibration of cross-classified models has been somewhat intractable. Recent developments in multilevel computational estimation strategies (Jones *et al* 1996b; Rasbash and Goldstein, 1994) and the collection of long-term multistage panel studies such as the British Household Panel Survey will allow such problems to be tackled in the future.

The third and final problem to be considered here is to recognize that apparent 'place' heterogeneity may really be 'people' heterogeneity. The models discussed above (and indeed the published literature) have focused on the elaborations of the higher-level random terms so as

to capture between-place heterogeneity. In contrast, the micro-models have, in the main, been simple with a single random part attempting to summarise between-people differences. This formulation presumes that people differ by a 'fixed' amount, but have the same variability; an unlikely presumption but more or less universally made in single-level regression modelling. To take a specific example, it may be that working-class voters are not only more likely to vote for Labour, but are also more variable in their voting behaviour than the non-working class.⁹ Crucially, the heterogeneity between levels may be confounded so that when simple level 1 models are used, there may be an overestimate of the higher-level variation. Multilevel estimation procedures and software can deal with such complexity (Bullen et al, 1996), and such developments must now routinely be put into practice.

CONCLUSIONS

Multilevel procedures have a number of features that make them attractive. Technically, by taking into account the complexity of the data, they overcome difficulties associated with autocorrelation. Consequently, significance tests and confidence intervals will more properly reflect the data structure, and there is reduced risk of mis-estimated precision and inferential error. Moreover, by pooling data they ameliorate the small number problem by making use of precision-weighted estimation.¹⁰ In substantive terms, multilevel modelling allows us to explore some of the complexity that we know exists in reality (Rose, 1974), and in so doing, provides for improved empirical description that is sensitive to context.

As Skinner *et al* (1989, 289) put it

'the issues raised are not mere technical issues which have only a marginal esoteric interest... On the contrary our experience suggests that the use of analytical procedures which take account of the population structure and the sample selection mechanism can change the objectives of the analysis and can have a substantial impact on the subsequent interpretation of the results'

while Eagles' (1995) book on 'Spatial and contextual models in political research' concludes (p.282)

'capturing the effects of context is likely to involve influences from a variety of spatial scales... developments in multilevel modelling ... will enormously facilitate this important research agenda in political science'.

⁹ In non-linear models of the propensity to vote Labour, this heterogeneity between individuals is called 'extra-binomial' variation (Goldstein, 1991a)

¹⁰ These technical issues have been at best dealt with here sketchily; for further discussion see Goldstein (1991b, 1995); Jones and Bullen (1994).

ACKNOWLEDGEMENTS:

Peter Tonkin of the Centre for Survey Analysis, University of Southampton helped with the analysis of the 1992 general election; while at Portsmouth, Robert Watt provided general research assistance. The data for both the 1992 and 1987 General Elections was made available through the ESRC Data Archive at the University of Essex. The research on which this paper is based was partly supported by ESRC grant H519255005 under the Analysis of Large and complex Data Programme.

REFERENCES

- Alker, H S (1969): 'A typology of ecological fallacies' in Dogan M and Rokkan, S (eds) *Quantitative Ecological Analysis*, MIT Press, Mass.
- Bryk, A S and Raudenbush, S W (1992): *Hierarchical linear models: applications and data analysis methods*, Sage, Newbury Park.
- Bullen, N I, Duncan, C and Jones, K (1996): Modelling complexity: analysing between-individual and between-place variation, *Environment and Planning*, Series A, forthcoming.
- Curtice J (1988): One nation? in Jowell R et al *British Social Attitudes; the 5th Report* Gower, Aldershot
- Duncan, C, Jones, K and Moon, G (1993a): 'Multivariate multilevel models for blood pressure, age and gender' In *ML3: a Guide to New Users*, 2nd edition (ed Woodhouse, G.), pp.55-82. University of London: Institute of Education.
- Duncan, C, Jones, K and Moon, G (1993b): 'Do places matter? A multilevel analysis of regional variations in health-related behaviour in Britain'. *Social Science and Medicine*, 37, 725-733.
- Duncan C, Jones K, Moon G (1996): 'Health-related behaviour in context: a multilevel approach'. *Social Science and Medicine*, forthcoming
- Eagles M (1995): The promise of spatial and contextual analyses in political research in Eagle M (ed) *Spatial and contextual models in political research*, Taylor and Francis, New York.
- Ennis, P H (1962): The contextual dimension in voting, in McPhee, W N and Glaser W A (eds) *Public opinion and congressional elections*, Free Press, Glencoe, 180-211.
- Goldstein (1984): 'The methodology of school comparisons', *Oxford Review of Education*, 10,69-74.
- Goldstein, H (1987): *Multilevel models in educational and social research*, Charles Griffin, London.
- Goldstein H (1991a): Nonlinear multilevel models, with an application to discrete response data *Biometrika* 78 45-51
- Goldstein, H (1991b): 'Multilevel modelling of survey data' *The Statistician*, 40,235-244.
- Goldstein, H.(1994): Multilevel cross-classified models. *Sociological Methods and Research* 22, 364-375.
- Goldstein, H (1995): *Multilevel Statistical Models*, Edward Arnold, London.
- Griffiths, M J and Johnston, R J (1991): 'What's in a place?' An approach to the concept of place as illustrated by the British National Union of Mineworkers' strike, 1984-5 *Antipode* 185-213.

- Hauser, R M (1970): Context and consex: a cautionary tale, *American Journal of Sociology* 75 645-664.
- Heath, A et al (1991): *Understanding political change*, Pergamon Press, Oxford.
- Heath, A et al (1993): *British election study 1992: user guide*, ESRC Data Archive, University of Essex, Colchester.
- Hox J J and Kreft I G (1994): Multilevel analysis methods. *Sociological Methods and Research* 22, 283-299.
- Huckfeldt R R (1984): Political loyalties and social class ties: the mechanisms of contextual influence, *American Journal of Political Science* 8 398-417.
- Johnston, R J, Pattie, C J and Allsopp, J G (1988): *A Nation Dividing?*, Longman, London.
- Johnston R J and Pattie C J (1992): Class de-alignment and the regional polarization of voting patterns in Great Britain, *Political Geography* 11 73-86.
- Jones, K (1991a): 'Specifying and estimating multilevel models for geographical research' *Transactions of the Institute of British Geographers*, 16, 148-159.
- Jones, K (1991b): 'Multilevel models for geographical research' *Concepts and Techniques in Modern Geography*, Environmental Publications, Norwich.
- Jones, K (1993a): 'Using multilevel models for survey analysis', *Journal of the Market Research Society*, 35, 249-265.
- Jones K (1993b): 'Everywhere is nowhere': *Multilevel perspectives on the importance of place*, Inaugural Lecture, University of Portsmouth, 12 pp.
- Jones, K (1994): 'Analysing and designing extensive studies in which place matters: a comment on neighbourhood incivilities and the study of crime in place' *Area*, 26, 70-81.
- Jones, K and Bullen, N J (1993): 'A multilevel analysis of the variations in domestic property prices: southern England, 1980-87' *Urban Studies*, 30, 1409-1426.
- Jones, K and Bullen, N J (1994): 'Contextual models of urban house prices : a comparison of fixed- and random-coefficient models developed by expansion', *Economic Geography*, 70, 252-272.
- Jones K and Duncan C (1996): People and places: the multilevel model as a general framework for the quantitative analysis of geographical data. In Longley, P. and Batty, M. (eds.) *Spatial Analysis: Modelling in a GIS Environment*, Longman.
- Jones K, Gould, M I, Watt, R (1996a): The Labour vote in the British General Election of 1992: an exploratory cross-classified analysis, *Sistemi Urbani*, forthcoming.
- Jones, K, Gould, M I and Watt, R (1996b): Multiple contexts as cross-classified models: the Labour vote in the British general election of 1992, *Mimeo*, Department of Geography, University of Portsmouth.
- Jones, K, Johnston, R J and Pattie, C J (1992): People, places and regions: exploring the use of multilevel modelling in the analysis of electoral data *British Journal of Political Science*, 22, 343-380.
- Jones, K, Tonkin, P and Wrigley, P (1996): 'Individual and constituency characteristics: analysing the geography of voting behaviour within a multilevel modelling framework' *Transactions of the Institute of British Geographers*, forthcoming.
- Massey, D (1991): 'The political place of locality studies' *Environment and Planning*, Series A 23, 267-281.

- Mason, W M, Wong, G Y and Entwistle, B (1984): 'The multilevel model: a better way to do contextual analysis' in Leinhardt, S (ed) *Sociological Methodology*, Jossey-Bass, San Francisco.
- McAllister I (1987a): Social context, turnout and the vote: Australian and British comparisons *Political Geography Quarterly*, 6, 17-30.
- McAllister I (1987b): Comment on Johnston, *Political Geography Quarterly*, 6, 45-50.
- McMahon D and Heath A (1992): Class and party in Britain: preliminary results with a multilevel logit model, *Multilevel Modelling Newsletter* 4 5-8.
- Miller W L (1978): Social class and party choice in England: a new analysis, *British Journal of Political Science* 8, 257-284.
- Miller W L (1979): Class, region and strata at the British General Election of 1979, *Parliamentary Affairs*, 32, 376-382.
- Multilevel Models Project (1993): A Guide to ML3, *Macros-Multilevel Multinomial Response Models*, University of London, Institute of Education.
- Openshaw, S (1977): A geographical solution to scale and aggregation problems in region-building, partitioning and spatial modelling, *Transactions of the Institute of British Geographers*, NS2, 459-72.
- Paterson, L and Goldstein, H (1992): 'New statistical methods for analysing social structures: an introduction to multilevel models', *British Educational Research Journal*, 17, 387-393.
- Przeworski A and Soares G A D (1971): Theories in search of a curve: a contextual interpretation of the Left Vote, *American Political Science Review*, LXV, 51-68.
- Pulzer, P (1967): *Political representation and elections in Britain*, Allen and Unwin, London.
- Rasbash J and Goldstein H (1994): 'Efficient analysis of mixed hierarchical and cross-classified random structures', *Journal of Educational and Behavioural Statistics* 19, 337-350.
- Rasbash J and Woodhouse G (1995): *Mln Command Reference: Version 1.0*, Multilevel Models Project, Institute of Education, University of London.
- Roberts, K H and Burstein, L (eds) (1980): *Issues in aggregation*, Jossey-Bass, San Francisco.
- Robinson, W S (1950): 'Ecological correlations and the behaviour of individuals' *American Sociological Review*, 15, 351-7.
- Rose R (1974): Britain: simple abstractions and complex realities in Rose R ed, *Electoral behaviour: a comparative handbook*, Free, Press, New York.
- Rose, R and McAllister, I (1990): *The loyalties of voters*, Sage Publications, London.
- Savage M, (1987): Understanding political alignments in contemporary Britain; do localities matter, *Political Geography Quarterly* 6 53-76.
- Scarborough E, (1987): The British electorate twenty years on: electoral change and election surveys, *British Journal of Political Science* 17 219-246.
- Skinner, C J, Holt, D and Smith, T F (eds.) (1989): *Analysis of Data from Complex Surveys*, John Wiley, Chichester.
- Tate C N (1974): Individual and contextual variables in British voting behaviour: an exploratory note, *American Political Science Review*, 68, 1656-62.
- Van den Eeden, P and Huettner, H J M (1982): 'Multilevel research', *Current Sociology*, 30, 1-181.
- Waller R (1983): *The almanac of British politics*, Croom Helm, London.

- Warde A (1986): Electoral mysteries, *International Journal of Urban and Regional Research* 10 289-94.
- Ware, J H (1985): 'Linear models for the analysis of longitudinal studies', *American Statistician*, 39, 95-101.
- Weatherford, M S (1986): Economic determinants of voting in Long, S (ed), *Research in micropolitics*, JAI Press, Greenwich CT.
- Wright G C (1977): Contextual models of electoral behaviour: the Southern Wallace vote, *American Political Science Review*, 71 497-508.

Context and Culture in the Health of Ethnic Groups

Chris Smaje

INTRODUCTION¹

It is appropriate that my paper is scheduled last in the day's proceedings, since my concerns combine elements of all of the previous presentations. Specifically, I am interested in examining spatial effects in the health of ethnic groups. This arises both from a feeling that, empirically, there is a good *prima facie* case for suspecting the presence of such effects, and that, theoretically, thinking through the implications of such effects may galvanise sociological understanding, not only of locality but also of ethnicity.

I should state at the outset that the work I have so far completed on this topic (reported in Smaje, 1995a) represents only the beginnings of an engagement with the relevant issues, and one which, empirically, falls short of the kind of analysis enabled by the techniques we have been discussing today. Moreover, sociological theorisations of ethnicity remain undeveloped, and indeed sometimes hostile to the line of enquiry opened up by a concern with the spatial dimensions of health among ethnic groups. I therefore offer my thoughts below in the spirit of scholarly debate, rather than in the conviction that any definitive perspective has yet emerged on the issue.

What I propose to do in this paper is, first, very briefly summarise the evidence concerning the role of spatial effects in the health of ethnic groups. I will then examine how such effects may be explained and how these putative explanations may inform our thinking on the links between ethnicity, 'culture' and health. Finally, I shall consider how these problems can be delineated by further research, focusing particularly but not exclusively upon the possibilities and pitfalls of existing census and survey data.

HEALTH, ETHNICITY AND PLACE

Until very recently, research on the health of minority ethnic groups in Britain was dominated by epidemiological approaches which tended to focus upon the ethnic patterning of particular diseases and to seek explanation - where one was sought at all - in the distribution of known underlying risk factors. With the inclusion of questions on both health and ethnicity in the 1991 census the opportunity has arisen to examine the social patterning of morbidity within and between ethnic groups in terms which permit more sociologically based analyses. While arguments proliferate about the appropriateness of the ethnic categories chosen in the census, and the extent to which its question on limiting long-term illness returns valid results across ethnic groups (see Smaje, 1995b), the opportunity to present a broad characterisation of health on a consistent national basis is surely to be welcomed.

Dunnell (1993) has shown that, after controlling for age, all minority ethnic groups with the exception of the Chinese report higher average levels of limiting long term illness than the white population. However, if cross-ethnic comparison of this sort is to be analytically useful we need to be aware both of the confounding factors which may distort the determination of an 'ethnic effect' implicit in such a comparison, and of the possibility that the relationship between such factors and health status may vary between different ethnic groups.

¹ I would like to thank Sara Arber for reading an earlier draft of this paper, and participants in the seminar at the University of Manchester for their comments.

Age, gender, socio-economic status, parental status, social roles, social support and area of residence are likely to be among the most important of these factors, yet researchers have made rather little progress in examining how they mediate the health experience of different ethnic groups. Indeed, most studies have simply controlled for age, gender and - less successfully - socio-economic status in reporting results across ethnic groups (Smaje, 1995b). However, with the availability both of census data and other data sets such as the Policy Studies Institute's Fourth National Survey of Ethnic Minorities it has become possible empirically to examine some of the spatial and social effects mentioned above.

An understanding of spatial effects is particularly important if we wish to move beyond the kind of descriptive analyses which have hitherto typified ethnic health research. The idea that the area in which one lives can exert an influence upon health independently of one's individual circumstances has been investigated using a variety of outcome measures, statistical techniques and data sets (e.g. Pocock et al, 1980; Fox et al, 1984; Humphreys and Carr-Hill, 1991; Morris and Carstairs, 1991; Macintyre et al, 1993; Duncan et al, 1994; Sloggett and Joshi, 1994). These studies have produced varied results, but on balance would seem to suggest that independent area effects have some bearing upon health, albeit that they are outweighed by individual circumstances. This is a particularly important result when the specific focus is upon the health of minority ethnic groups, since these groups have a very distinctive and highly concentrated residential distribution. Nearly 70 per cent of people who described their ethnic group as other than 'white' in the 1991 census lived in the urban areas of Greater London, the West Midlands, Greater Manchester and West Yorkshire, where only around 20 per cent of the white population lived (Owen, 1994). Substantial concentration existed even within these broad areas. Looking at this rather crudely, with the exception of the Chinese, more than half of all people from each of the 'non-white' groups enumerated in the 1991 census lived in electoral wards where the total 'non-white' population exceeded 44 per cent, even though together the 'non-white' groups constituted less than 6 per cent of the national population (Owen, 1994; see also Smith, 1989, for an account of the social and political processes involved in this process of concentration). Thus, in view of the fact that most people from minority ethnic groups are concentrated into relatively few areas, if significant spatial effects upon health exist it is possible that apparent 'ethnic effects' in health status are an artefact of residential patterns.

This possibility has been little investigated in Britain. However, initial work appears to indicate varying levels of an independent 'ethnic effect' in areas of different character, even after controlling for aggregate differences between areas in the socio-economic composition of their populations (hereafter called 'compositional differences'). This finding, if it proves to be robust, is interesting in that it indicates spatial variation in the *degree* of ethnic health differentials which cannot be accounted for in terms of compositional effects. To the extent that, implicitly, we regard minority ethnic status as betokening some commonality of experience by grouping people under this category and seeking correlations with other dimensions of experience such as health, this contextual variability therefore imposes upon us the duty to look more carefully at the commonalities assumed among people in the same putative ethnic group. Macintyre (this volume) has usefully schematised the range of mechanisms which may underlie area effects in health, distinguishing in particular between contextual (mainly physical) effects, collective (socio-cultural) effects, and compositional effects (see also Macintyre et al, 1993). In this case, it would appear that we need to look for collective effects - that is, for mechanisms within the organisation of ethnic groups themselves, or in the nature of their encounters with the broader population - as well as contextual and compositional factors.

The significance of collective effects has been the focus of a rather separate literature, which has not always been concerned with the significance of area as such, but has examined the impact of 'social consistency' defined in a number of possible ways, upon health status. For example, several studies from around the world have found that people living in areas where their own ethnic or class characteristics are predominant tend to enjoy better health than those whose characteristics are atypical of their area of residence. The most compelling evidence for such an 'ethnic density' effect has come in studies of mental well-being (Halpern, 1993). However, there is also some evidence on the impact of ethnic density upon infant mortality rates. LaVeist (1993) undertook an ecological analysis of data from 176 US cities, finding that black residential concentration and black poverty were positively and independently associated with the black infant mortality rate. Conversely, the degree of black political power - measured by the racial composition of the city council indexed to the composition of the electorate - was associated with lower mortality after controlling for other factors. Political power was nevertheless also associated with concentration. LaVeist suggested that the impact of political power could not be explained in terms of direct resource flows to the black population achieved by local black politicians, but is indicative instead of an underlying process of 'community integration', which explains both the greater degree of political mobilisation and lower infant mortality. Thus, greater community integration in segregated black populations partially offsets the negative health consequences of segregation itself.

These results would appear to suggest that ethnic density may conflate two distinct processes with divergent implications for health. On the one hand, the areas into which minority ethnic groups are typically concentrated may enjoy poorer housing quality, employment opportunities, recreational facilities and so on, which contribute to poorer health (Smith, 1989; Massey and Denton, 1993). On the other, the processes of community integration enabled by residential concentration may be a resource for positive health, particularly with regard to mental well-being.

These arguments have not been subjected to systematic empirical analysis in Britain, but are not inconsistent with the findings of existing studies based in Glasgow and London (Ecob and Williams, 1991; Smaje, 1995a)². Clearly, ethnic density effects are only one among many factors bearing upon health, and differentials between ethnic groups are probably affected much more strongly by average group differences in individual material circumstances, and perhaps also by spatial effects in terms of the physical properties of the areas in which minority ethnic groups are concentrated. However, elaborating the significance of ethnic density effects is useful not just as a refinement of our empirical grasp of ethnic patterns in health, but also for two other reasons. First, it will help develop our understanding of the way material and psycho-social mechanisms mediate health experience. The concept of 'community integration' as a result of spatial concentration - redolent of Tönnies's *Gemeinschaft* or the remnants of the Chicago school's urban ecology - is at best rather crude and demands that researchers proceed to excavate convincing mechanisms linking space, 'community' and health. Second, pursuing studies in this area may help tell us something more general about the nature of ethnicity in contemporary Britain. Certainly, the notion of cultural process may add a layer of complexity to sociological theories which are often informed by rather over-generalised critical perspectives.

² In both cases some evidence was found for group density effects on the basis of concentration of the 'South Asian' ethnic group (although such a group, as Ecob and Williams in particular show, clearly comprises several identifiable 'sub-groups' which have a non-random spatial distribution). Smaje's study also examined the 'black' (i.e. Caribbean and African) population, finding no significant effects.

EXPLAINING SPATIAL EFFECTS ON THE ETHNIC PATTERNING OF HEALTH

If spatial effects of the kind described above are firmly established, we need to understand the processes underlying them. Doing so inevitably involves defining a standpoint on four issues of increasingly general scope. First, it is necessary to determine the extent to which such effects may be a consequence of social selection or, more directly, of some feature of the physical or socio-cultural environment. In either case, it follows that, second, some relationship exists between the spatial and the social. In other words, geographic space, far from being the neutral canvas upon which social life is painted, is in fact implicated in the construction of social life. This in turn suggests that, third, at least implicit in these ideas is some notion of ethnic 'communities'. As I suggest below, the concept of a community is problematic in sociological terms, and the concept of an ethnic community still more so. It is therefore important to be explicit about the nature of the claims that are being made. Finally, in broaching the subject of ethnic communities we necessarily confront at the broadest level questions about the meaning of 'ethnicity' itself.

This is not the place for a thorough review of all of these questions, but a few comments on each of the above points are in order.

Social Selection or Socio-Cultural Causation?

A recurrent question in research on the social patterning of health is whether associations between health and social variables result from some feature of the experience of being in a particular social group, or, conversely, whether health status allocates social position. There is, for example, some evidence that the excess incidence of schizophrenia among people of lower socio-economic status results from downward social mobility consequent upon experiencing the illness (Cochrane and Bal, 1987), although the extent of the selection effect remains debatable (Fox, 1990). It is possible that a selection process of this sort underlies the ethnic density effect described above; here, people who were downwardly mobile or in poor health would drift from areas of high minority ethnic residence to areas of lower density residence, thus accounting for the apparently positive effect of ethnic density. But this does not seem very likely. Many areas of low ethnic density are of higher social status than the urban zones which typically contain the greatest concentrations of people from minority ethnic groups. Here, one would expect a drift of *upwardly* mobile individuals. Moreover, there does not appear to be any compelling evidence of a drift of unhealthy people away from areas of high minority ethnic residence.

Another selection process which particularly affects minority ethnic groups in Britain due to their origins in recent migrations is the 'healthy migrant effect'. Those individuals who tackled the not inconsiderable obstacles of migration to a distant country - normally for economic betterment - are unrepresentative of either the populations from which they originate or to which they migrate, typically being considerably healthier. This, it has been suggested, may explain the often favourable health experience of migrants when compared to the general population³. But while such an effect may plausibly explain *between* group differences, it cannot of itself explain *within* group differences of the sort suggested by the ethnic density hypothesis, since everyone involved in a particular migration is likely to be subject to the same selective process (see Williams, 1993 for a more detailed consideration of this point). It may, however, be the case that ethnic density effects result from a 'wearing off' of the healthy migrant effect which is spatially ordered.

³ Although little direct evidence exists for this assertion.

Two points can be made here. First, the 'wearing off' of a selection effect implies the action of other (non-selective) processes. Thus, any hypothesis couched in these terms suggests the need to look beyond selection factors alone towards social causation. Second, if we were to hypothesise that such a 'wearing off' process underlies ethnic density effects, the direction typically taken by the latter would lead us to the suggestion that areas of high ethnic density act as centres for receipt of recent migrants - and are hence typified by the good health of recent 'healthy migrants' - while the more established migrants, among whom the effect had 'worn off', were the ones moving to areas of lower ethnic density. There can be little doubt that such a process occurs, but its overall significance is much more questionable. The picture it conjures may have been more applicable in the earlier phases of post-war 'chain' migration and subsequent entrenchment, but urban areas of high minority ethnic residence today contain long established populations of early migrants⁴. Moreover, the decline of large-scale primary migration from the early 1960s onwards calls into question the notion that high density areas are significant 'reservoirs' of recent migrants. Finally, it is unlikely that anything quite so simple is occurring as a unidirectional movement of longer-established migrants from initial residence in areas of high density to subsequent residence in lower density areas. While minority ethnic groups do exhibit greater spatial mobility than the majority population, these movements are complex and, as Robinson (1991) has shown, often involve urbanisation - as with Pakistanis moving from the textile towns of the north to the larger West Midlands conurbation - as well as counter-urbanisation, which is more typical of the Indian populations of the South-East. Thus, while selection processes may be important, it would appear unlikely that an explanation based solely upon the spatial sorting of migrants by length of residence can explain the kind of results described earlier.

It is not possible to test the 'wearing off' hypothesis directly using available cross-sectional census data. However, in Table 1 the individual Sample of Anonymised Records (SAR) from the 1991 census is used to examine the proportion of minority ethnic residents in areas of different ethnic concentration who were born abroad. This is an extremely poor substitute for a variable indicating length of residence in the UK, but if there are systematic differences in place of birth between people in the same age group and ethnic group living in areas of different concentration, this may indicate the presence of selection processes. The 278 SAR areas were ranked in terms of the proportion of people describing their ethnic group as other than white, and then aggregated into five groups, Group 1 representing areas of lowest concentration and Group 5 representing the areas of highest concentration. The proportion of people born in the UK (or Ireland) in four age groups within each of three 'non-white' groups ('Black', 'Asian' and 'Other', in accordance with the OPCS four-category aggregation of its 1991 census groupings) was then calculated. As would be expected, there is a very strong age gradient (seen by looking down the columns in the table), with the younger age groups displaying much higher proportions of UK born individuals. The critical dimension, however, is the gradient across concentration categories within particular age and ethnic groups (seen by looking along the rows). Here, the table shows little systematic variation, with the exception of Black people aged 35-49, who display a very striking gradient towards a lower proportion of UK birth in areas of higher concentration. The numbers in some of the cells for the areas of low concentration are quite small so these results should be interpreted with some caution. On the face of it, however, they do not reveal striking spatial variation by country of birth within ethnic groups.

⁴ In fact, different age groups within the same ethnic group in a given area may undergo very different selection processes, which considerably complicates the analytical picture.

Table 1: Proportion of Selected Ethnic Groups Born in the UK by Area of Residence

		Proportion UK Born by Area of Residence (Concentration Quintiles 1 = Lowest, 5 = Highest) (%)				
	Age	1	2	3	4	5
Black	<16	87	82	82	87	89
	16-34	66	66	60	68	66
	35-49	48	40	22	15	8
	50+	14	28	10	10	6
Asian	<16	83	90	84	88	89
	16-34	30	39	33	39	35
	35-49	5	2	4	4	3
	50+	2	7	4	3	3
Other	<16	77	83	73	80	80
	16-34	32	29	27	31	30
	35-49	12	19	11	10	8
	50+	18	15	11	13	7

Source: 2% Sample of Anonymised Records from the 1991 Census, Crown Copyright

Another indirect way of examining spatial sorting using the SAR is to consider patterns of internal migration by comparing areas of previous and current residence in those individuals who have changed address. Unfortunately, this information is only available rather crudely at the regional level. Looking this time at all individuals within the Caribbean, Indian and Pakistani groups, the proportion of people in each group who changed their region of residence within the UK was 8.2 per cent, 7.4 per cent and 7.3 per cent respectively. Among Caribbean people, of those who *had* changed region, some 67 per cent came from inner and outer London, and the West Midlands. Looking just at this sub-set of people, the same three regions were the destination of over 90 per cent of them. A very similar pattern obtains for the Indian population when the same three regions of origin/destination are chosen, and for the Pakistani population in its three major regions of residence: the West Midlands, the North West and Yorkshire/Humberside.

Clearly, analysis at the regional level may obscure a more finely grained sorting from areas of high ethnic concentration to lower density areas *within* regions; the apparent movement of Indian people from inner to outer London may be an example of this. However, at the regional level, there is little evidence of major dispersion as these minority populations mature demographically. Much of the movement of individuals seems to be between areas of established high minority ethnic residence (see also Robinson, 1991; Ballard, 1994).

Turning then to social causation, there is a good deal of evidence to suggest that socio-cultural and psycho-social factors are implicated in health differences between social groups. The extent and the nature of people's social roles and social support are associated with health status, as are perceptions of 'self-efficacy' or 'locus of control'. These relationships may often be indirect, mediated by structural or material factors associated with the social states described, but factors such as social support may also have a direct psycho-social effect - probably via neuroimmunological mechanisms - by providing a 'buffer' against social stressors (Kaplan, 1991; Aneshensel, 1992). There has been some discussion of how these factors interact with gender, socio-economic status and race or ethnicity (Dressler, 1988; Hughes and Demo, 1989; Hibbard and Pope, 1993), though most of the evidence on the latter is confined to the USA.

Dressler (1988), for example, has examined the impact of status consistency upon depressive symptoms in a black population in the Southern USA. In essence, status consistency refers to the degree to which an individual's perceived or idealised status and their actual status are congruent. Dressler identified several status dimensions, finding that 'lifestyle incongruity' - in which the possession of a range of consumer items and the participation in various consumption activities exceeds that predicted by occupational ranking - was associated with poor mental well-being. Drawing upon Veblen's (1918) classic discussion concerning the conspicuous consumption of the leisure class, Dressler argued that, amongst other things, the "symbolic discordance between one's self-definition in terms of lifestyle and the status attributed by the community on the basis of occupational ranking" underlay his findings (1988, p.87). Since his study was confined to a particular ethnic group, he was unable to examine the extent to which lifestyle incongruity varies between groups, but he suggested that it was likely to be particularly problematic for discriminated-against minorities. The black population, for example, faces considerable material constraint but also tends substantially to share in the "overall ideology of American culture" which is "one of open access to avenues of upward mobility" (1988, p.87-88). In this context, Dressler suggests several reasons why lifestyle incongruity is likely to be more problematic for black people than for whites. Again, it is important to emphasise that this kind of research should not be construed as explaining overall ethnic differences in mental well-being - the extent of which in any case remain unclear (Cockerham, 1990). There may be some force to the criticism that research into the mental health of minority ethnic groups has concentrated too much upon the response of individual people from minorities and not enough upon broader structural and institutional mechanisms (Francis, 1993). However, my present purpose is not to estimate the significance of the various factors affecting the overall mental health of people from minority ethnic groups, but merely to establish that there may be some plausible socio-cultural or psycho-social mechanisms directly linking social structure with health.

It is doubtful whether the implications of Dressler's research can be translated wholesale to Britain. The black-white dichotomy routinely employed in US research is increasingly being called into question in the USA (e.g. LaVeist, 1994) and is clearly inappropriate in Britain. Moreover, as we have just seen, most minority ethnic groups in this country comprise a large number of migrants of diverse original backgrounds, and the experiences and motivations associated with migration considerably complicate analysis of the socio-cultural mediation of health among these groups. This does not, of course, mean that such effects are absent.

Linking the Spatial and the Social

Few of the studies mentioned above explicitly incorporate concepts of space or community into their analysis. However, Halpern (1993) suggests that the positive association between ethnic concentration and health arises through the protection from prejudice afforded by concentration and through the provision of social support it affords, which helps minority ethnic individuals cope with disadvantage. Evidence, contrary to the majority population, of a *negative* relationship between socio-economic status and mental well-being among Indian women in Britain (Cochrane and Stopes-Roe, 1981) is consistent, in Halpern's view, with this suggestion since high socio-economic status in this group generally reflects upward social mobility, which, he argues, in turn is often associated with movement to 'higher status' areas where the buffering of the group density effect would no longer obtain.

Ethnic group identity is invoked by Halpern as an *a priori* given, around which affective ties in space are necessarily organised. Thus, he suggests that the superior mental well-being of South Asian groups in Britain relative to the Caribbean population may reflect the greater segregation of the former; the strength of the group or community here seems indexed purely to concentration in space - the more concentrated, the more social support. Other accounts, such as the work by Dressler described above, draw out some of the mechanisms by which socio-cultural factors may affect health within particular communities but rest upon a very abstract notion of 'community'. In societies such as the UK and USA, in which a heavy historical investment in ideologies of racial 'difference' can be discerned, it could be that the health effects reported above do indeed reflect an intensification of community support associated with concentration but unmediated by the effectivity of ethnicity or space itself. Concentration, in other words, may act simply as an independent variable with no significant interactions between it, ethnic identity and the nature of communality. There may, on the other hand, be something more complex occurring. This is suggested by evidence of substantive differences between groups in the nature of social institutions. For example, Halpern, in discussing an Irish counter-example to the normally positive direction of the group density effect, describes "the apparently schizophrenogenic aspects of rural Catholic Irish culture" (1993, p.600; see also Williams, 1992). In the USA, Williams et al (1992) have shown that ethnic differences exist in the mental health consequences of divorce or a spouse's death, suggesting that marriage may have different social implications in different groups, while Dressler himself has argued that health behaviours are modified by ethnic differences in concepts of the individual's incorporation or embeddedness in society (1993).

Thus, it may be that there are ethnic differences in communal norms. The key question, however, is the extent to which such differences are informed by spatial distribution itself. For example, Moore (1992) has argued that the concentration of the Bangladeshi population in poor quality housing in Spitalfields, East London, arose largely through local government policy, but enabled the reproduction of familiar economic and domestic patterns within a community infrastructure which also helped its members guard themselves against racist attacks. Thus, while processes of residential concentration cannot themselves constitute populations as ethnic groups, their consequences may be to help define a self-conscious and spatialised ethnic 'community'.

Ethnic Communities?

There appear, then, to be differences between ethnic groups in their distribution through space and in their social orientation. An obvious question is the extent to which these dimensions of difference are interrelated. The classic formulation of this problem was in the urban ecology of

the Chicago school sociologists. For them, congruences between space and communality were regarded not only as obvious but as paradigmatic. However, the associations between status, identity and spatial order they sought are now rejected as overly simplistic. More recent theorists have moved away from tracing communities absolutely in space, and concentrated upon articulating the way that people identify themselves symbolically with others (Bell and Newby, 1971; Cohen, 1985). More generally, there has been an 'erosion of classic norms' (Rosaldo, 1993), whereby the concern with structure in classical social theory has been replaced with a contemporary tendency towards anti-structure which emphasises the plurality of social understandings within even narrowly delimited contexts. Older concerns with class and community have thus given way to ideas of transience, of cultural borderlands which people constantly negotiate in their everyday lives. Under the influence of post-modern theories of consumer culture and the globalisation of communication, concepts of locality have lost much of their sociological force. For writers like Appadurai (1990), the contemporary world is characterised by simultaneous processes of local hybridisation and global homogenisation which neutralise the specific effects of the local. Similarly, it is difficult to imagine the post-modern consumer described by Baudrillard (1988), trapped in the self-referential world of 'hyperreality', caring much about the people in the surrounding streets, or indeed caring much about anything at all.

An argument about the unmediated effect of local community upon something as concrete as health therefore directly confronts these contemporary approaches. Empirically, the evidence adduced by Halpern constitutes one kind of argument that space *does* matter, and that not everyone inhabits the sort of world described by Baudrillard. In addition, perhaps the finding that marked ethnic segregation persists - in some areas, ED-level indices of dissimilarity exceed 90 per cent for certain ethnic groups - suggests that the notion of local hybridisation invoked by Appadurai is overstated. But such observations do not in themselves allow us to return to the geographical reifications of Chicago. The presumption underlying the concept of an 'ethnic community' is that ethnic groups are undifferentiated internally but defined by socio-cultural boundaries which coincide with their distribution in physical space. This presumption needs, however, to be demonstrated rather than assumed. Robinson, marshalling the empirical evidence concerning the effectivity of space, concludes,

The spatial distribution of ethnic groups *produces* inequalities in access to services, employment, desirable housing and ultimately life chances; it *shapes* patterns of social interaction whether these be positive (e.g. whom you marry) or negative (e.g. whom you attack); and it *contributes* to the development of attitudes and stereotypes (1987, p.194).

But here - as in the debates about the extent to which ethnic concentration is a result of choice, constraint or the 'contingent' impact of socio-economic status - the theoretical status accorded the various terms of the argument - ethnicity, space, 'life chances' - remains unclear. A conceptual as well as an empirical approach to the question is required.

Theorising Ethnicity

Establishing the appropriate basis on which to detach putatively 'ethnic' phenomena from other social distinctions is a pervasive problem in broader theorisations of ethnicity which elsewhere I describe in terms of a dilemma between reductionism and reification (Smaje, 1996a). Liberal sociology has been castigated by commentators writing from a variety of critical perspectives

for constructing essentialist notions of ethnic difference (e.g. Bourne, 1980; CCCS, 1982), but these writers themselves tend towards either an ontic reification of race,

We need to analyse race in terms of its *specific* form at *different* periods of time in order to see how it articulates - or not - with other social relations (CCCS, 1982, p.35).

or formulations of such generality that it is difficult to envisage how any empirical purchase can be gained,

Structures of class, racism, gender and sexuality cannot be treated as 'independent variables' because the oppression of each is inscribed within the other - is constituted by and is constitutive of the other (Brah, 1992, p.132).

Thinking through the relationships between ethnicity, space and community is one way to try to move beyond this kind of generality. With a few notable exceptions (Harvey, 1989), Marxist traditions of class analysis upon which much of the more radical writings are based have never been interested in the social effectivity of space, albeit that there are some suggestive possibilities in the concept of 'struggles' by the 'black community'. Nor, however, can we be satisfied with a kind of Weberian primordialism which simply assumes some prior ethnic identity instantiated in 'communities' of co-ethnics cutting across class (Rex, 1986). There is empirical work to be done here in examining the way that notions of status and communality operate within and across social and spatial lines of demarcation. With respect to the construction of ethnicity in contemporary Britain, such work has barely begun, although suggestive lines of enquiry have been opened by analysis of South Asian business (Lyon and West, 1995) and ethnic cultural identity (Acland and Siriwardena, 1989; James, 1993; Ballard, 1994; Modood et al, 1994). But it is also necessary to find a theoretical language that enables us to make useful generalisations about the relationships between ethnicity, class, culture and space while refusing the essentialism of any particular term⁵.

Here, recent critiques in the sociology of consumption may point the way. Campbell (1995) argues that most sociological theories of consumption - from Veblen to Baudrillard - are grounded in the metaphor of communication, in which the sociologist imputes a *post hoc* 'meaning' to the act of consumption. "Consumption," he continues, "is then seen as involving an 'attempt' by the consumer to 'adopt a lifestyle' or 'create an identity' when there are few grounds for any such assumption" (1995, p.117). Miller's (1994) arguments against the global homogenisation thesis are also apposite. On the basis of ethnographic work in Trinidad he asserts, "Trinidad is not becoming more like anywhere else except in the most superficial sense that it is using products of the global economy" (1994, p.319). Alternatively, he develops a complex argument which, rather than merely seeing existing groups or societies as constructing their own forms of representation from the material available to them, regards both the contours of these groups and the cultural orders associated with them as 'objectifications' emergent as historical forms of a broader process of human self-creation.

⁵ Almost all accounts of ethnicity, particularly those concerned with aspects of everyday experience such as health, adopt an essentialist shorthand in talking of 'ethnic groups'. This is not always inappropriate: often membership of such groups is quite clear for many purposes. But this does not negate the fact that, at some level of theoretical abstraction, ethnic categories are social fictions. Indeed, the examination of social and cultural process necessitated by research of the sort suggested in this paper can, it seems to me, help us formulate more precisely *how* such constructions proceed. See Smaje (1996a).

Turning then to social causation, there is a good deal of evidence to suggest that socio-cultural and psycho-social factors are implicated in health differences between social groups. The extent and the nature of people's social roles and social support are associated with health status, as are perceptions of 'self-efficacy' or 'locus of control'. These relationships may often be indirect, mediated by structural or material factors associated with the social states described, but factors such as social support may also have a direct psycho-social effect - probably via neuroimmunological mechanisms - by providing a 'buffer' against social stressors (Kaplan, 1991; Aneshensel, 1992). There has been some discussion of how these factors interact with gender, socio-economic status and race or ethnicity (Dressler, 1988; Hughes and Demo, 1989; Hibbard and Pope, 1993), though most of the evidence on the latter is confined to the USA.

Dressler (1988), for example, has examined the impact of status consistency upon depressive symptoms in a black population in the Southern USA. In essence, status consistency refers to the degree to which an individual's perceived or idealised status and their actual status are congruent. Dressler identified several status dimensions, finding that 'lifestyle incongruity' - in which the possession of a range of consumer items and the participation in various consumption activities exceeds that predicted by occupational ranking - was associated with poor mental well-being. Drawing upon Veblen's (1918) classic discussion concerning the conspicuous consumption of the leisure class, Dressler argued that, amongst other things, the "symbolic discordance between one's self-definition in terms of lifestyle and the status attributed by the community on the basis of occupational ranking" underlay his findings (1988, p.87). Since his study was confined to a particular ethnic group, he was unable to examine the extent to which lifestyle incongruity varies between groups, but he suggested that it was likely to be particularly problematic for discriminated-against minorities. The black population, for example, faces considerable material constraint but also tends substantially to share in the "overall ideology of American culture" which is "one of open access to avenues of upward mobility" (1988, p.87-88). In this context, Dressler suggests several reasons why lifestyle incongruity is likely to be more problematic for black people than for whites. Again, it is important to emphasise that this kind of research should not be construed as explaining overall ethnic differences in mental well-being - the extent of which in any case remain unclear (Cockerham, 1990). There may be some force to the criticism that research into the mental health of minority ethnic groups has concentrated too much upon the response of individual people from minorities and not enough upon broader structural and institutional mechanisms (Francis, 1993). However, my present purpose is not to estimate the significance of the various factors affecting the overall mental health of people from minority ethnic groups, but merely to establish that there may be some plausible socio-cultural or psycho-social mechanisms directly linking social structure with health.

It is doubtful whether the implications of Dressler's research can be translated wholesale to Britain. The black-white dichotomy routinely employed in US research is increasingly being called into question in the USA (e.g. LaVeist, 1994) and is clearly inappropriate in Britain. Moreover, as we have just seen, most minority ethnic groups in this country comprise a large number of migrants of diverse original backgrounds, and the experiences and motivations associated with migration considerably complicate analysis of the socio-cultural mediation of health among these groups. This does not, of course, mean that such effects are absent.

Linking the Spatial and the Social

Few of the studies mentioned above explicitly incorporate concepts of space or community into their analysis. However, Halpern (1993) suggests that the positive association between ethnic concentration and health arises through the protection from prejudice afforded by concentration and through the provision of social support it affords, which helps minority ethnic individuals cope with disadvantage. Evidence, contrary to the majority population, of a *negative* relationship between socio-economic status and mental well-being among Indian women in Britain (Cochrane and Stopes-Roe, 1981) is consistent, in Halpern's view, with this suggestion since high socio-economic status in this group generally reflects upward social mobility, which, he argues, in turn is often associated with movement to 'higher status' areas where the buffering of the group density effect would no longer obtain.

Ethnic group identity is invoked by Halpern as an *a priori* given, around which affective ties in space are necessarily organised. Thus, he suggests that the superior mental well-being of South Asian groups in Britain relative to the Caribbean population may reflect the greater segregation of the former; the strength of the group or community here seems indexed purely to concentration in space - the more concentrated, the more social support. Other accounts, such as the work by Dressler described above, draw out some of the mechanisms by which socio-cultural factors may affect health within particular communities but rest upon a very abstract notion of 'community'. In societies such as the UK and USA, in which a heavy historical investment in ideologies of racial 'difference' can be discerned, it could be that the health effects reported above do indeed reflect an intensification of community support associated with concentration but unmediated by the effectivity of ethnicity or space itself. Concentration, in other words, may act simply as an independent variable with no significant interactions between it, ethnic identity and the nature of communality. There may, on the other hand, be something more complex occurring. This is suggested by evidence of substantive differences between groups in the nature of social institutions. For example, Halpern, in discussing an Irish counter-example to the normally positive direction of the group density effect, describes "the apparently schizophrenogenic aspects of rural Catholic Irish culture" (1993, p.600; see also Williams, 1992). In the USA, Williams et al (1992) have shown that ethnic differences exist in the mental health consequences of divorce or a spouse's death, suggesting that marriage may have different social implications in different groups, while Dressler himself has argued that health behaviours are modified by ethnic differences in concepts of the individual's incorporation or embeddedness in society (1993).

Thus, it may be that there are ethnic differences in communal norms. The key question, however, is the extent to which such differences are informed by spatial distribution itself. For example, Moore (1992) has argued that the concentration of the Bangladeshi population in poor quality housing in Spitalfields, East London, arose largely through local government policy, but enabled the reproduction of familiar economic and domestic patterns within a community infrastructure which also helped its members guard themselves against racist attacks. Thus, while processes of residential concentration cannot themselves constitute populations as ethnic groups, their consequences may be to help define a self-conscious and spatialised ethnic 'community'.

Ethnic Communities?

There appear, then, to be differences between ethnic groups in their distribution through space and in their social orientation. An obvious question is the extent to which these dimensions of difference are interrelated. The classic formulation of this problem was in the urban ecology of

the Chicago school sociologists. For them, congruences between space and communality were regarded not only as obvious but as paradigmatic. However, the associations between status, identity and spatial order they sought are now rejected as overly simplistic. More recent theorists have moved away from tracing communities absolutely in space, and concentrated upon articulating the way that people identify themselves symbolically with others (Bell and Newby, 1971; Cohen, 1985). More generally, there has been an 'erosion of classic norms' (Rosaldo, 1993), whereby the concern with structure in classical social theory has been replaced with a contemporary tendency towards anti-structure which emphasises the plurality of social understandings within even narrowly delimited contexts. Older concerns with class and community have thus given way to ideas of transience, of cultural borderlands which people constantly negotiate in their everyday lives. Under the influence of post-modern theories of consumer culture and the globalisation of communication, concepts of locality have lost much of their sociological force. For writers like Appadurai (1990), the contemporary world is characterised by simultaneous processes of local hybridisation and global homogenisation which neutralise the specific effects of the local. Similarly, it is difficult to imagine the post-modern consumer described by Baudrillard (1988), trapped in the self-referential world of 'hyperreality', caring much about the people in the surrounding streets, or indeed caring much about anything at all.

An argument about the unmediated effect of local community upon something as concrete as health therefore directly confronts these contemporary approaches. Empirically, the evidence adduced by Halpern constitutes one kind of argument that space *does* matter, and that not everyone inhabits the sort of world described by Baudrillard. In addition, perhaps the finding that marked ethnic segregation persists - in some areas, ED-level indices of dissimilarity exceed 90 per cent for certain ethnic groups - suggests that the notion of local hybridisation invoked by Appadurai is overstated. But such observations do not in themselves allow us to return to the geographical reifications of Chicago. The presumption underlying the concept of an 'ethnic community' is that ethnic groups are undifferentiated internally but defined by socio-cultural boundaries which coincide with their distribution in physical space. This presumption needs, however, to be demonstrated rather than assumed. Robinson, marshalling the empirical evidence concerning the effectivity of space, concludes,

The spatial distribution of ethnic groups *produces* inequalities in access to services, employment, desirable housing and ultimately life chances; it *shapes* patterns of social interaction whether these be positive (e.g. whom you marry) or negative (e.g. whom you attack); and it *contributes* to the development of attitudes and stereotypes (1987, p.194).

But here - as in the debates about the extent to which ethnic concentration is a result of choice, constraint or the 'contingent' impact of socio-economic status - the theoretical status accorded the various terms of the argument - ethnicity, space, 'life chances' - remains unclear. A conceptual as well as an empirical approach to the question is required.

Theorising Ethnicity

Establishing the appropriate basis on which to detach putatively 'ethnic' phenomena from other social distinctions is a pervasive problem in broader theorisations of ethnicity which elsewhere I describe in terms of a dilemma between reductionism and reification (Smaje, 1996a). Liberal sociology has been castigated by commentators writing from a variety of critical perspectives

for constructing essentialist notions of ethnic difference (e.g. Bourne, 1980; CCCS, 1982), but these writers themselves tend towards either an ontic reification of race,

We need to analyse race in terms of its *specific* form at *different* periods of time in order to see how it articulates - or not - with other social relations (CCCS, 1982, p.35).

or formulations of such generality that it is difficult to envisage how any empirical purchase can be gained,

Structures of class, racism, gender and sexuality cannot be treated as 'independent variables' because the oppression of each is inscribed within the other - is constituted by and is constitutive of the other (Brah, 1992, p.132).

Thinking through the relationships between ethnicity, space and community is one way to try to move beyond this kind of generality. With a few notable exceptions (Harvey, 1989), Marxist traditions of class analysis upon which much of the more radical writings are based have never been interested in the social effectivity of space, albeit that there are some suggestive possibilities in the concept of 'struggles' by the 'black community'. Nor, however, can we be satisfied with a kind of Weberian primordialism which simply assumes some prior ethnic identity instantiated in 'communities' of co-ethnics cutting across class (Rex, 1986). There is empirical work to be done here in examining the way that notions of status and communality operate within and across social and spatial lines of demarcation. With respect to the construction of ethnicity in contemporary Britain, such work has barely begun, although suggestive lines of enquiry have been opened by analysis of South Asian business (Lyon and West, 1995) and ethnic cultural identity (Acland and Siriwardena, 1989; James, 1993; Ballard, 1994; Modood et al, 1994). But it is also necessary to find a theoretical language that enables us to make useful generalisations about the relationships between ethnicity, class, culture and space while refusing the essentialism of any particular term⁵.

Here, recent critiques in the sociology of consumption may point the way. Campbell (1995) argues that most sociological theories of consumption - from Veblen to Baudrillard - are grounded in the metaphor of communication, in which the sociologist imputes a *post hoc* 'meaning' to the act of consumption. "Consumption," he continues, "is then seen as involving an 'attempt' by the consumer to 'adopt a lifestyle' or 'create an identity' when there are few grounds for any such assumption" (1995, p.117). Miller's (1994) arguments against the global homogenisation thesis are also apposite. On the basis of ethnographic work in Trinidad he asserts, "Trinidad is not becoming more like anywhere else except in the most superficial sense that it is using products of the global economy" (1994, p.319). Alternatively, he develops a complex argument which, rather than merely seeing existing groups or societies as constructing their own forms of representation from the material available to them, regards both the contours of these groups and the cultural orders associated with them as 'objectifications' emergent as historical forms of a broader process of human self-creation.

⁵ Almost all accounts of ethnicity, particularly those concerned with aspects of everyday experience such as health, adopt an essentialist shorthand in talking of 'ethnic groups'. This is not always inappropriate: often membership of such groups is quite clear for many purposes. But this does not negate the fact that, at some level of theoretical abstraction, ethnic categories are social fictions. Indeed, the examination of social and cultural process necessitated by research of the sort suggested in this paper can, it seems to me, help us formulate more precisely *how* such constructions proceed. See Smaje (1996a).

The work of writers like Campbell and Miller points to a theoretical project which gives full weight to the complexity of meanings, intentions and interconnections in people's senses of self, culture and community. I am unable to offer any clearly articulated argument which can link this level of theoretical abstraction to the finding of ethnic group density effects in health with which I began my discussion, but I remain convinced that we need to try to work such arguments through. A good starting point seems to me to acknowledge, like Rosaldo (1993), that while there is much to be learned from recent critiques of classical sociology's obsession with structure, the concerns of a voguish anti-structuralism contain their own distortions of social life. While communities may, in sociological terms, be 'imagined' (cf. Anderson, 1991), their imagining is real enough to affect people's health in accordance with how individuals are positioned vis-a-vis communities.

Like Brah, I am unhappy with the notion that ethnicity and other dimensions of human difference can ultimately be treated simply as independent variables. But the argument that they are 'inscribed' within each other does not get us very far. We therefore need to think about how the empirical relationships found in census and survey data between their imperfect operationalisations of everyday social categories like 'health', 'ethnicity', 'class' and 'area' can help us refine theoretically what we think these concepts mean.

AN EMPIRICAL AGENDA

This returns us to the main theme of discussion for the day - exploiting survey data to examine spatial effects. I want to conclude by reflecting upon how the themes emerging from the discussion above can be addressed in empirical research.

Essentially, the task is twofold. First, the existence, direction and geographic distribution of spatial effects in the health of different ethnic groups must be established. Second, it is necessary to determine whether these are indicative of group density effects, and thence to identify the possible mechanisms involved. As I have suggested above, this may in turn contribute among other things to more mature theoretical reflections on the nature of ethnicity itself.

Census and survey data are indispensable for the first of these tasks, and of considerable utility for the second. With regard to the first, it seems clear that the ability of multilevel statistical techniques to model simultaneously for contextual and compositional effects in health data offers the key to mapping the distribution of ethnic density effects. Other people presenting papers today have commented with more authority than I possess upon the detail of these techniques and I have nothing to add here. I do, however, wish to comment on some of the difficulties which arise in trying to examine the questions outlined above empirically using existing survey data.

Mapping Contextuality

On the first question of contextuality, the difficulties mentioned above can be subsumed under three headings,

- choosing appropriate health outcome measures
- choosing appropriate areal units of analysis
- choosing appropriate measures of socio-economic status

Much of the literature on group density effects is concerned with mental health, and it is here where we might expect to find the most powerful relationships. It is interesting to note, however, that Duncan et al (1994) found little contextual variation in GHQ scores at ward and regional level in their multilevel analysis of *Health and Lifestyle Survey* data, though their study was not concerned specifically with ethnicity. Social stressors are also associated with physical health outcomes (Hibbard and Pope, 1993), but there is likely to be a more complex mediation in this case. Clearly, one has little choice of dependent variable when using data from the 1991 census, which only contains information on limiting long-term illness and permanent sickness. It may, however, be the case that these are inappropriate variables for discerning contextuality. Unfortunately, more detailed sample surveys such as the *Health and Lifestyle Survey* rarely contain sufficient numbers of people from minority ethnic groups to offer a realistic chance of examining ethnic group density effects. It is possible that more fruitful possibilities may be contained in recent surveys which have specifically sampled minority ethnic groups, such as the Health Education Authority's recent health and lifestyle survey (HEA, 1994), and the Policy Studies Institute's forthcoming Fourth National Survey of Ethnic Minorities.

A second issue is choosing the appropriate unit of areal analysis. Distinct ethnic concentration can be discerned at every geographic level in Britain from regions down to enumeration districts. However, the small absolute size of minority ethnic groups is such that analysis at larger geographical scales may not prove particularly useful, certainly if the interest is in ethnic group density effects. Thus, while nearly 60 per cent of all 'non-white' individuals enumerated in the 1991 census live in the Southeast, only 20 per cent of the population of the Southeast are non-white (Owen, 1994). As we move from region to local authority, ward, enumeration district and street we typically find ethnic concentration increasing. It is important then to choose an areal level of analysis consistent with the hypothesis we are examining. If we are proposing the existence of an ethnic density effect, and bearing in mind the nature of ethnic concentration in Britain, this level - as Halpern (1993) indicates - is likely to be fairly small. There are also clearly problems in mapping a presumed concept of 'community' onto generally asocial geographic co-ordinates. The most pressing practical problem, however, is in attaching microdata to area information of sufficiently small scale to be sociologically useful. For confidentiality reasons SAR data are available only at essentially local authority level; analysis of these data may entirely miss contextual effects at more local levels. Although no such constraints exist in small area and local base statistics, one is then faced with problems of ecological inference. To examine ethnic density effects, it may be possible partially to circumvent the problem through multilevel log-linear analysis of aggregate data tabulated at the desired spatial level; this would enable comparison between areas which took account of compositional differences without making inferences about individuals (Duncan et al, 1993).

Turning to socio-economic status, the problems of indicators such as social class for socio-economic status both within and between ethnic groups are well known, and will not be discussed further here. Other data available from the census such as employment status, car ownership and housing characteristics may be more appropriate. However, some care needs to be taken over the way these are incorporated into the relevant models. Two points can be made here. First, since at least as long ago as John Rex's controversial discussion of 'housing classes' (Rex and Moore, 1967) it has been clear that much of salience about ethnic residential concentration revolves around housing quality and employment opportunities. While there will clearly be differences on these measures between individuals living in a given area, to include them only as measures of individual circumstances may be to confound the compositional with the contextual. Second, there may also be significant gender effects. For example, it has been shown that the class homogeneity of geographic areas apparent on the basis of male occupation

does not always hold good when the position of women is brought into the picture (Pratt and Hanson, 1988). 'Position' in this context can mean female occupational status, social mobility or status within the household. The net effect of this could be that neither measures such as social class nor measures based on household characteristics capture women's individual circumstances satisfactorily. It is also possible that the spatial patterning of social networks varies significantly by gender which may complicate the putative relationship between place and the construction of community.

Explaining group density effects

This brings us on from questions about mapping contextual effects to a more direct focus upon identifying the mechanisms presumed to underlie ethnic group density effects, if the latter are indeed detected. Here, again there are three issues which warrant some discussion,

- eliminating health selection explanations
- distinguishing between the physical and socio-cultural effects of area
- identifying which aspects of cultural process are protective of health

Cross-sectional data typically available from censuses and surveys do not allow the elimination of selection explanations. The only rigorous way of doing so is to adopt a prospective or at least a panel study design, the OPCS Longitudinal Survey being perhaps the most obvious possibility. The significance of selection effects was discussed earlier, and it was suggested that there may be some role for the 'wearing off' of healthy migrant effects which are spatially sorted, albeit that this still implies social causation. A major question here is whether there is something particularly positive about ethnic concentration or whether, in a racist society, the effect is located in the set of problems associated with upward mobility or living in areas of low ethnic concentration. One way of tackling this is to examine between-group differences *within* areas of minority ethnic concentration to determine whether the more concentrated group enjoys superior health, but results of this sort are likely to be confounded by other possible between-group differences. The best approach is to advance particular hypotheses about the nature of the observed density effect and attempt to test them directly.

This is also the best strategy for unpicking the possibly contradictory effects of the specific physical properties of the areas in which people from minority ethnic groups are concentrated, and the effects of concentration itself - a distinction I have elsewhere termed the extrinsic and intrinsic effects of concentration (Smaje, 1995a). The figure below is a simplified version of the model presented in that analysis. As the figure indicates, if ethnic concentration promotes social support a positive effect on health may be detected, but if concentration occurs in 'undesirable' areas with unhealthy physical properties the former effect may be 'washed out' by the negative consequences of the latter. It is therefore better to concentrate on the way that the relationship between area and health may be mediated by these two sets of factors, than to infer specific effects from the gross relationship between area and health.

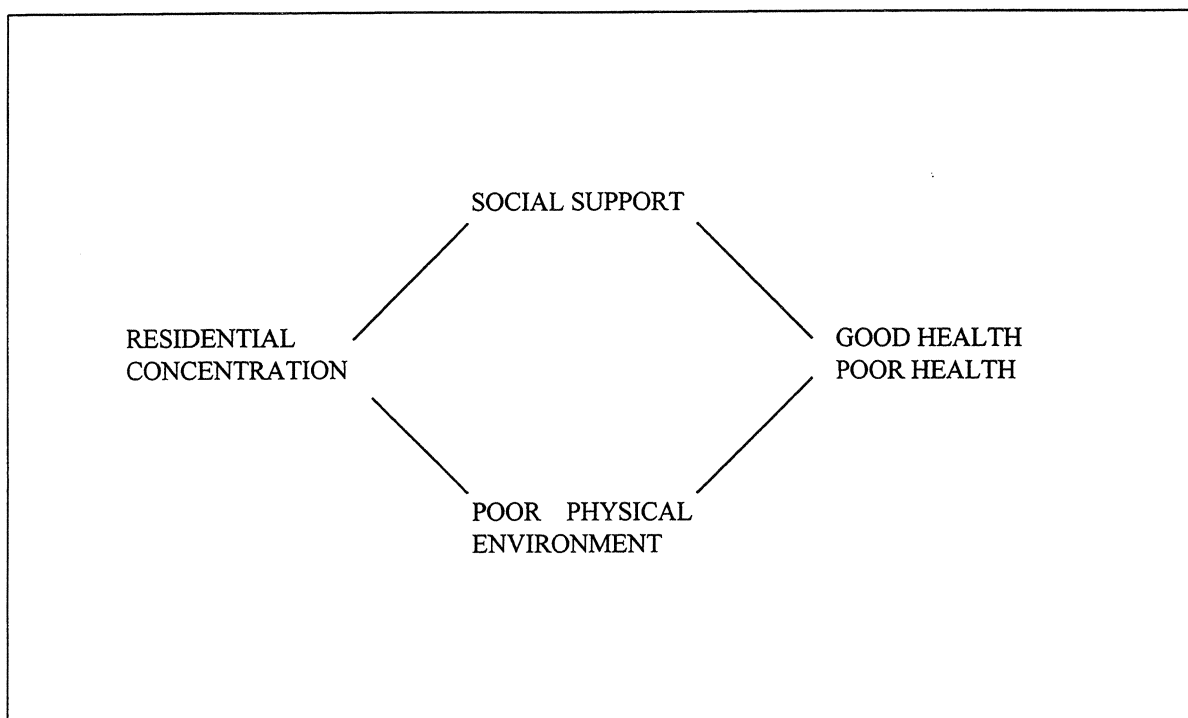


Figure 1: Ethnic residential concentration and health

All of this therefore indicates the need to operationalise the concepts of social support, community integration and so on to which group density effects point circumstantially. Clearly, this cannot be done from census data, although some possibilities may be delineated by examining geographic patterns in household and kinship structures. Richer information on activities, participation in 'ethnic' organisations and family support is available in the HEA and PSI surveys; a good starting point for analysis would certainly be to examine the correlations between these variables and health status within ethnic groups and between areas (though progress here would be limited by sample size considerations, since neither survey is particularly large). However, one is left wondering whether these somewhat crudely quantitative measures might fail to capture a more organic sense of what being in a community is like. On this point, qualitative and in particular comparative ethnographic approaches may have more to contribute, but they would clearly be unable to provide any focused answer as to which mechanisms link ethnic concentration to health.

A mixed methods approach of this kind might nevertheless enable us to build a fairly compelling case for some kind of community integration or support hypothesis. It may also throw interesting light on the point made earlier about differences in the nature of communality between ethnic groups. It is perhaps impossible to broach this subject in health or social policy contexts without inviting the charges of ethnocentrism, inferiorisation, reification and victim blaming to which many commentators have not unreasonably been sensitised by disreputable projects in policy analysis and social science of which perhaps the best known example was the Moynihan Report in the USA (Rainwater and Yancey, 1967). But, as I have argued elsewhere (Smaje, 1996b) a blank refusal to engage with cultural explanations for the ethnic patterning of health is increasingly problematic. The whole question seems much less controversial in more general learned discussion, not least in so far as many of those who are developing arguments about the distinctive features of black culture are themselves black scholars of radical political persuasions (e.g. Henry, 1992; hooks, 1992; Gilroy, 1993; Miller, 1993). There are also some fascinating discussions of the historical development of Caribbean and South Asian kinship and

social forms (e.g. Miller, 1994; Lyon and West, 1995) which increasingly call into question simplistic characterisations of minority ethnic culture(s) as oriented unidimensionally either to communality or resistance.

But finally it could be that more systematic analysis will indicate that *no* contextual or group density effects exist. This would be a surprising result on the basis of existing evidence, even though one may not expect group density effects to explain much in absolute terms about the overall health of particular ethnic groups. Nevertheless, recent discussions of a variety of South Asian populations indicate that, sometimes in spite of spatial dispersion, particular groups remain structurally cohesive but also internally segmented (Ballard, 1994; Anwar, 1995; Lyon and West, 1995). This indicates some of the potential complexities which may confound any expectation of finding a generalisable ethnic density effect associated neatly with levels of concentration, and again underlines the need for developing fairly precise hypotheses about the nature of the relationship between ethnicity, community and health. Thus, in view of the level of *de facto* ethnic residential concentration in Britain it is important that researchers concerned with the health of minority ethnic groups pay attention to developments in research on area effects in health, while at the same time exploring the significance of the social and cultural resources within ethnic groups which may not be spatially ordered in any simple way.

REFERENCES

- Acland, T. and Siriwardena, S. (1989) 'Integration and segregation in an Asian community', *New Community*, vol 15, no 4, pp 565-576.
- Anderson, B. (1991) *Imagined Communities*, London: Verso.
- Aneshensel, C. (1992) 'Social stress: theory and research', *Annual Review of Sociology*, vol 18, pp 15-38.
- Anwar, M. (1995) 'Social networks of Pakistanis in the UK: a re-evaluation' in A. Rogers and S. Vertovec (eds) *The Urban Context: Ethnicity, Social Networks and Situational Analysis*, Oxford: Berg.
- Appadurai, A. (1990) Disjuncture and difference in the global cultural economy. *Theory, Culture and Society*, 7, 295-310.
- Ballard, R. (1994) *Desh Pradesh: The South Asian Presence in Britain*, London: Hurst.
- Baudrillard, J. (1988) *Selected Writings*. Stanford: Stanford University Press.
- Bell, C. and Newby, H. (1971) *Community Studies*. London: George Allen and Unwin.
- Bourne, J. (1980) 'Cheerleaders and ombudsmen: the sociology of race relations in Britain' *Race and Class*, 21, 4, pp 331-52.
- Brah, A. (1992) 'Difference, diversity and differentiation' in Donald, J. and Rattansi, A. *'Race', Culture and Difference*, London: Sage.
- CCCS [Centre for Contemporary Cultural Studies] (1982) *The Empire Strikes Back*, London: Hutchinson.
- Campbell, C. (1995) 'The sociology of consumption' in Miller, D. *Acknowledging Consumption*, London: Routledge.
- Cochrane, R. and Stopes-Roe, M. (1981) 'Social class and psychological disorder in natives and immigrants to Britain' *International Journal of Social Psychiatry* 27, pp 173-82.
- Cochrane, R. and Bal, S. (1987) 'Migration and schizophrenia: an examination of five hypotheses', *Social Psychiatry*, 22, pp 181-91.

- Cockerham, W. (1990) 'A test of the relationship between race, socio-economic status and psychological distress' *Social Science and Medicine* 31, 12, pp 1321-6.
- Cohen, A. (1985) *The Symbolic Construction of Community*, London: Routledge.
- Dressler, W. (1988) 'Social consistency and psychological distress', *Journal of Health and Social Behaviour*, vol 29, pp 79-91.
- Dressler, W. (1993) 'Type A behaviour: contextual effects within a southern black community', *Social Science and Medicine*, 36, 3, pp 289-95.
- Duncan, C., Jones, K. and Moon, G. (1993) 'Modelling ecologies: the multilevel model as a general framework for analysing census data' Paper presented for the 1991 Census Conference, Newcastle, September 1993.
- Duncan, C., Jones, K. and Moon, G. (1994) 'Regional variations in psychiatric morbidity: a multilevel approach' Paper presented for the Annual Conference of the BSA Medical Sociology Group, York, September 1994.
- Dunnell, K. (1993) 'Sources and nature of ethnic health data' in NE/NW Thames Regional Health Authority, *The Health of the Nation: The Ethnic Dimension*, London: NE/NW Thames Regional Health Authority.
- Ecob, R. and Williams, R. (1991) 'Sampling Asian minorities to assess health and welfare', *Journal of Epidemiology and Community Health*, vol 45, pp 93-101.
- Fox, A., Jones, D. and Goldblatt, P. (1984) 'Approaches to studying the effect of socio-economic circumstances on geographic differences in mortality in England and Wales', *Medical Bulletin*, vol 40, no 4, pp 309-314.
- Fox, J. (1990) 'Social class, mental illness and social mobility: the social selection-drift hypothesis for serious mental illness' *Journal of Health and Social Behaviour*, 31, pp 344-53.
- Francis, E. (1993) 'Psychiatric racism and social police: black people and the psychiatric services', in W. James and C. Harris (eds) *Inside Babylon: The Caribbean Diaspora in Britain*, London: Verso.
- Gilroy, P. (1993) *The Black Atlantic: Modernity and Double Consciousness*. London: Verso.
- HEA [Health Education Authority] (1994) *Health and lifestyles: black and minority ethnic groups in Britain*, London: Health Education Authority.
- Halpern, D. (1993) 'Minorities and mental health', *Social Science and Medicine*, vol 36, no 5, pp 597-607.
- Harvey, D. (1989) *The urban experience*, Baltimore: Johns Hopkins University Press.
- Henry, C. (1992) 'Understanding the underclass: the role of culture and economic progress', in Jennings, J. (ed) *Race, politics and economic development*, London: Verso.
- Hibbard, J. and Pope, C. (1993) 'The quality of social roles as predictors of morbidity and mortality' *Social Science and Medicine* 36, 3, pp 217-25.
- Hooks, b. (1991) *Yearning: Race, Gender and Cultural Politics*. London: Turnaround.
- Hughes, M. and Demo, D. (1989) 'Self-perceptions of black Americans: self-esteem and personal efficacy' *American Journal of Sociology* 95, 1, pp 132-59.
- Humphreys, K. and Carr-Hill, R. (1991) 'Area variations in health outcomes: artefact or ecology', *International Journal of Epidemiology*, vol 20, no 1, pp 251-258.

- James, W. (1993) 'Migration, racism and identity formation: the Caribbean experience in Britain' in James, W. and Harris, C. *Inside Babylon: The Caribbean Diaspora in Britain*. London: Verso.
- Kaplan, H. (1991) 'Social psychology of the immune system: a conceptual framework and review of the literature' *Social Science and Medicine* 33, 8, pp 909-23.
- LaVeist, T. (1993) 'Segregation, poverty and empowerment: health consequences for African Americans', *Milbank Quarterly*, vol 71, no 1, pp 41-64.
- LaVeist, T. (1994) 'Beyond dummy variables and sample selection: what health service researchers ought to know about race as a variable', *Health Services Research*, 29, 1, pp 1-16.
- Lyon, M. and West, B. (1995) 'London Patels: caste and commerce' *New Community*, 21, 3, pp 399-419.
- Macintyre, S., Maciver, S. and Soomans, A. (1993) 'Area, class and health: should we be focusing on places or people?', *Journal of Social Policy*, vol 22, no 2, pp 213-234.
- Massey, D. and Denton, N. (1993) *American Apartheid* Massachusetts: Harvard University Press.
- Miller, A. (1993) 'Social science, social policy and the heritage of African-American families'. In Katz, M. (ed) *The 'Underclass' Debate: Views From History*. Princeton: Princeton University Press.
- Miller, D. (1994) *Modernity: An Ethnographic Approach*, Oxford: Berg.
- Modood, T. (1994) *Changing Identities* London: Policy Studies Institute.
- Moore, R. (1992) 'Labour and housing markets in inner city regeneration', *New Community*, 18, 3, pp 371-86.
- Morris, R. and Carstairs, V. (1991) 'Which deprivation? a comparison of selected deprivation indices', *Journal of Public Health Medicine*, vol 13, pp 318-26.
- Owen, D. (1994) 'Spatial variations in ethnic minority groups populations in Great Britain', *Population Trends*, no 78, pp 23-33.
- Pocock, S., Shaper, A., Cook, D., Packham, R., Lacey, R., Powell, P. and Russell, P. (1980) 'British regional heart study: geographic variations in cardiovascular mortality, and the role of water quality', *British Medical Journal*, 24th May 1980, pp 1243-1249.
- Pratt, G. and Hanson, S. (1988) 'Gender, class and space' *Environment and Planning D: Society and Space* 6, pp 15-35.
- Rainwater, L. and Yancey, W. (1967) *The Moynihan Report and the Politics of Controversy*. Cambridge, Massachusetts: MIT Press.
- Rex, J. (1986) 'The role of class analysis in the study of race relations: a Weberian perspective'. In Rex, J. and Mason, D. (eds) *Theories of Race and Ethnic Relations*. Cambridge: Cambridge University Press.
- Rex, J. and Moore, R. (1967) *Race, Community and Conflict*, London: Oxford University Press.
- Robinson, V. (1987) 'Race, space and place: the geographical study of UK ethnic relations 1957-1987', *New Community*, vol 14, no 1/2, pp 186-197.
- Robinson, V. (1991) 'Goodbye yellow brick road: the spatial mobility and immobility of Britain's ethnic population, 1971-81' *New Community*, 17, 3, 313-30.
- Rosaldo, R. (1993) *Culture and Truth: The Re-making of Social Analysis* London: Routledge.

- Sloggett, A. and Joshi, H. (1994) 'Higher mortality in deprived areas: community or personal disadvantage?', *British Medical Journal*, vol 309, pp 1470-4.
- Smaje, C. (1995a) Ethnic residential concentration and health: evidence for a positive effect? *Policy and Politics*, 23, 3, 251-69.
- Smaje, C. (1995b) Health, 'race' and ethnicity: making sense of the evidence, London: King's Fund.
- Smaje, C. (1996a) *Not Just a Social Construct: Theorising Racial Ontology* Guildford: University of Surrey (Occasional Paper, Department of Sociology).
- Smaje, C. (1996b) 'The ethnic patterning of health: new directions for theory and research' *Sociology of Health and Illness*.
- Smith, S. (1989) *The Politics of 'Race' and Residence*, Cambridge: Polity.
- Veblen, T. (1899) *The Theory of the Leisure Class*, New York: Macmillan.
- Williams, R. (1992) 'The health of the Irish in Britain' in Ahmad, W. *The Politics of 'Race' and Health* Bradford: Race Relations Research Unit.
- Williams, R. (1993) 'Health and length of residence among South Asians in Glasgow: a study controlling for age', *Journal of Public Health Medicine*, 15, 1, pp 52-60.
- Williams, D., Takeuchi, D. and Adair, R. (1992) Marital status and psychiatric disorder among blacks and whites. *Journal of Health and Social Behaviour*, 33, 140-57.