Environment and Planning B: Planning and Design 1999, volume 26, pages 855 – 878

Towards a new digital data infrastructure for urban analysis and modelling

P A Longley, R J Harris
School of Geographical Sciences, University of Bristol, University Road, Bristol BS8 1SS, England;
e-mail: paul.longley@bristol.ac.uk; r.j.harris@bristol.ac.uk
Received 18 February 1999; in revised form 12 May 1999

Abstract. Formal models of urban systems have the potential to reveal a lot about the form and functioning of urban settlements, yet much of this potential has still to be realised. In this paper we focus on the extent to which this has reflected the dearth of digital data that are rich, relevant, and disaggregate. Geodemographic classifications have made important and enduring contributions to small-area analysis. Yet, on the one hand, reliance upon census data makes them outdated and irrelevant and, on the other, fragmentation and diversification of social areas in cities has made the 'mosaic metaphor' of small-area analysis untenable. As part of the quest for a new perspective on data modelling, we investigate in this paper the potential of 'lifestyles' data sets for creating richer, more relevant digital models of human activity patterns in cities.

1 Introduction

Simple powerful theories and models are the most established and enduring in social science. Yet they provide only normative sketches which are increasingly irrelevant to the understanding of the messy irregularity that characterises the patterning of the real world. Empirical generalisation requires quantitative data collected according to rigorous research designs, yet most such data are too infrequently collected, too coarse, and/or insufficiently relevant to the functioning of fast-changing systems. In this paper we begin to gather a new perspective on the practical foundations to model building and to seek a rapprochement between social scientific orthodoxies and the practice of generalising individual and household activity patterns. We develop our arguments in the context of the analytical tradition of measuring and modelling the spatial patterning of social groupings in city regions, extended to encompass the so-called 'lifestyles' analysis of household consumption and activity patterns. We argue that our ability routinely to measure, share, and concatenate rich digital data sources creates opportunities to develop relevant and timely depictions of what is going on right across urban systems but that social science has so far held back from embracing such sources. The reasons for this are valid, yet if ultimately determinate they will stifle the creativity of model building in the digital age. As one of us has argued elsewhere (Goodchild and Longley, 1999; Longley, 1998), the 'new digital infrastructure' to GIS-based analysis can be fundamentally unsystematic in design but this need not necessarily preclude all systematic analysis. In our discussion we will make much of the roles of improved *data models* of spatial distributions in fostering the development of systematic and thematic *urban models* of social systems.

2 Models of social patterns in cities

Measurement of the patterning of urban social areas is a long and rich tradition in urban geography, from the work of the Chicago human ecologists through to the development of computer-based methods for social area analysis in the 1950s and 1960s. The classic exemplar is the Burgess model, which was devised to depict the way in which a clearly identifiable process (rapid immigration to the core of the city)

became manifest in a differentiated mosaic of residential land uses, as waves of 'invasion and succession' swept through successive inner-city neighbourhoods. Although generations of students have since seen measurement of pattern as a goal in itself, it is important to remember that the original rationale for such measurement lay in a simple theory of urban dynamics and that an early objective was to relate generating process to spatial pattern.

The dynamics of Chicago's rapid growth in the 1920s present an unusually straightforward context for analysis of the evolution of urban form—a fast-growing city, the change dynamics of which could be traced simply to in-migration. The application nevertheless spawned the much more general 'social area analysis' tradition to analysis of residential differentiation, based on a developing range of principal component and factor analysis techniques. Reviews of this research (see Clarke and Gleave, 1973; also Timms, 1969) illustrate the way in which inductive generalisation about the similarities between residential areas, harnessed to the 'mosaic metaphor' (Johnston, 1999), took place in innumerable case studies. Much of this research was avowedly technocentric and arose out of the development of computers to handle, by the standards of the time, large and complex data sets. Yet even as taxonomies of social areas and the groups resident within them became more detailed and sophisticated, the approach came to be seen as increasingly irrelevant to any understanding of the way in which towns and cities functioned (see Harvey, 1973). Over time, it became apparent just how dependent statistical classifications were upon the particular cocktail of variables that were used to generate them, and the constructs and labels that were appended to statistical 'dimensions' (such as 'stage in family life cycle') themselves came under greater scrutiny (see Stapleton, 1980). At the same time, within mainstream quantitative analysis, there was heightened awareness of ecological fallacy and modifiable areal unit effects in geographical analysis (see Openshaw, 1984). If the characteristics *of* areas could misleadingly be confounded with the characteristics of individuals *in* areas, then there were clear problems in overreliance upon publicly available data sets available only for (often very coarse) areal aggregations (see Cole, 1993). In short, from the mid-1970s onwards, data models of multivariate spatial distributions (and thence the thematic urban models that were built upon them) became viewed with increasing suspicion.

Data models of spatial distributions provided the foundations to thematic models of urban systems (Batty, 1981), and disillusionment with the measurement paradigm no doubt contributed significantly to the demise of urban modelling in the late 1970s. Not only were the units of analysis too coarse and spatial attributes too ill defined, but the state of computation was also too rudimentary and the scope of urban models too ambitious to capture the richness and diversity of urban systems (Birkin, 1995; Sayer, 1979). Such problems were compounded in any attempt to model the dynamics of increasingly rapid change. The growth of Burgess's Chicago was undoubtedly an anomaly in the history of urban dynamics, in that the single process which fuelled its short-term growth was clearly defined and spatially manifest. Yet even here the evolving physical layout and land-use configuration was not the regular and idealised 'city of pure geometry' (Batty and Longley, 1994) of textbook illustrations. The urban modelling tradition of the 1970s was able to come to terms neither with the myriad forms of human agency, nor with the jagged irregularity of urban morphology that arises out of urban growth dynamics in the real world. The 1960s and 1970s saw the morphology of urban land use affected by changes in affluence, increases in car ownership, and fragmentation of consumption patterns—in short, the lifestyles of those resident within the physical carcass of the city became increasingly diverse. As this took place, so the representation of urban dynamics by using crude surrogate data models, and crude spatial partitions, became increasingly irrelevant to the understanding of city systems.

In this context, the quest to relate form to function, patterning to social process, was largely abandoned (Batty and Longley, 1997). Since that time, urban geography has arguably been overwhelmed by the task of representing the statics and dynamics of spatial structure, to the point at which the discipline appears to have all but withdrawn from the task of generalisation. From the innovation of behavioural geography through to the depictions of individuals in cultural geography, the balance of intellectual activity has shifted from system-wide generalisation to richer yet haphazard depiction of disparate and fragmented subgroups within society. This has had knock-on effects in terms of the confidence of the discipline to suggest prescriptions and prognoses for urban change. Generalisation is a cornerstone to rational planning policy, and an urban geography which eschews system-wide generalisation is likely to become relegated to the sidelines of all but academic discourse. The demise of applied geography is particularly apparent in this context (see Pacione, 1999), as is the reduced esteem in which 'predict and provide' planning is presently held.

3 Geodemographics and the emergence of 'lifestyles'

Academic qualms about the validity, scope, and applicability of social area analysis have had few implications for applied marketing geography. Within the United Kingdom, for example, classifications of residential areas have become an established marketing tool ever since digital census data first appeared following the 1971 Census (Beaumont, 1991). This applications field has become known as 'geodemographics', defined by Brown (1991, page 221) as "a shorthand label for both the development and the application of area typologies that have proved to be powerful discriminators of consumer behaviour and aids to 'market analysis'". In the United Kingdom and USA, data models of geodemographic distributions of entire populations have been built by retail consultancies, by applying techniques of cluster and principal components to census data. Qualitative analysis of the results of data reduction leads to the assigning of labels to the different groups (such as 'affluent achievers', 'have nots', 'thriving greys', etc) which marketeers have associated successfully with particular product and service niches in retailing (Goss, 1995). In an important review paper, Batey and Brown (1995) trace the transfer of the techniques of social area analysis to applications in marketing through a range of 'near-market' research activities. Geodemographics has no core theory beyond the notion that 'birds of a feather flock together' (Flowerdew and Leventhal, 1998) yet experience has shown that this provides no practical barrier to successful application. Various refinements have been carried out to the classification methodologies and their marketing to clients, including the use of supplementary noncensus data to label census classifications (Batey and Brown, 1995).[1] By the mid-1990s geomdemographics had become a successful and standard tool of the marketeer. Different proprietary systems have used different cocktails of census counts and classifying algorithms (with the SuperProfiles system having perhaps the best academic pedigree; Openshaw, 1996) yet successive geodemographic systems produced from 1971, 1981, and 1991 UK Census data represent applications of a core technology which have enjoyed repeat purchase by a range of business clients.

However, fundamental problems remain with census-based classifications, which may be illustrated with respect to the UK case. The raw data of census counts provide at best imperfect indicators of likely consumption behaviour because crucial information is not collected (notably income data). In turn, composite indicators of consumer

[1] Supplementary descriptors include the use of the National Readership Survey and data from large-scale public-sector surveys such as the General Household Survey and Family Expenditure Survey. The ascription of labels from the coarser and varied geographies of such sources introduces the risk of invoking additional ecological fallacies in the classification process.

behaviour are thus also dependent upon crude surrogate data. Moreover, current geodemographic systems are frequently reliant upon census data that are over a decade old, and this represents an increasing handicap in fast-changing, consumer-led markets.[2] Census data are comprehensive in terms of coverage of the UK population—notwithstanding problems of (illegal) nonresponse by the 'missing millions' in the 1991 UK Census (Marsh, 1993)—yet this is rarely a key criterion for market area analysis, with its usual focus on subgroup behaviour. Outside of the USA, many public-sector agencies have found it necessary to introduce aggressive data-pricing regimes in order to recover some of the costs of data creation, and this presents a further disincentive to the use of census sources. Taken together (and contrary to the claims of some of those who sell geodemographic systems) it is clear that the data infrastructure provided by the census does not present a panacea for analysis of consumption and activity patterns. It is certainly well founded in survey research terms, but its content is increasingly marginal to the understanding and prediction of what is going on in modern Britain.

At the same time, the capture of digital data by using a range of new technologies has become commonplace. This has led to the advent of a wide range of so-called 'lifestyles' data, originating from such diverse sources as consumer product-guarantee returns, store loyalty programmes, and recorded travel behaviour. A wide definition of lifestyles data would emphasise their chief facets: they 'capture' (measure) some of the varied consumption choices, shopping habits, and practices of identifiable individuals. The wider definition thus includes a range of nonsurvey-based sources of lifestyles data, such as guarantee-card returns, electronic point-of-sale (EPOS) from retail purchases, loyalty card data, share ownership records, and country court judgments, many of which have been available since the beginning of the 1990s. Companies such as ICD, Claritas, Experian, and Pyschographics have built up huge 'data warehouses'. CACI Information Services claim that by using their 'LifestylesUK' product it is possible to target over 44 million individuals by using 300 lifestyle variables (source: promotional brochure). Claritas UK's 'Micromarketing' offers information on 75% of UK households (source: promotional brochure).

In the empirical sections of this paper (and following Harris, 1998, page 2) we will take a narrow definition of GB lifestyles data, as "data obtained and stored at the non-aggregated level of a named individual and geo-referenced by their address; the data [are] collated from the return of (consumer) questionnaires mailed directly to eligible voters, as recorded upon the Electoral Register for Great Britain". An example of one such lifestyle questionnaire is shown in figure 1.

Lifestyles data sets elicit information on a far wider range of themes than the census, from household structure and demographic characteristics (often including income) to consumption habits and recreational pursuits. As such, today's lifestyles data provide a range of relevant (direct and/or indirect) indicators of individual and household propensities to consumer particular goods and services, as well as detailed information about individual and household activity patterns. Data pertain to the

[2] This said, there is a counterargument that geodemographic classifications remain effective long after the data from which they are derived have become outdated. This is because, the argument goes, the classificaitons are 'driven' or influenced by variables that reflect directly the structure of the property market. If we assume that typologies capture property *attributes* as well as household *characteristics* then, as time goes by, although the residents will change, the property market acts as an effective filter which tends to have the effect of making those who replace them *likely* to share many essential characteristics of their predecessors. In marketing terms, this means that targeting the same areas remains likely to reach the same *kinds* of consumers, or that mail shots will continue to reach people who share similar tastes, aspirations, and patterns of consumer behaviour—subject to changing small-area fashions, as the 1980s phenomenon of gentrification illustrates.

ABOUT THIS SURVEY

Consumer Surveys will keep you fully up to date with new offers and information from reputable companies as they become available. The following questions have been asked by the organisations listed below who may use your answers for marketing purposes: q.10 The Open University, q.25 Accident Assist Ltd., q.33 Hayter plc, qs.77/78 and 134/135 Norwich Union Financial Services, qs. 85, 87 and 89 RBS Advanta, qs.103 to 117 Imperial Tobacco Ltd., qs.120/121 Kimberly-Clark Ltd., q.141 Stewart Assurance Company Ltd., q.142 Berkeley Jacobs Financial Services Ltd., qs.76 and 158 MKD Holdings Ltd., q.170 Micro Star, q.171/172 AOL Bertelsmann Online, q.173 Lloyds Bank Insurance Direct.

PLEASE HELP US TO HELP YOU

GUIDANCE NOTES

1. Please answer questions on behalf of yourself, your partner or your household as appropriate. Of course, you can always discuss your answers with other members of your family.
2. Please feel free to ignore particular questions if you wish. Your remaining answers are still of value.
3. Please write in CAPITAL letters – tick boxes like this ☑
4. If you make a mistake, simply cross it out and continue.

Name and address changes

Please check to see if your name and address details are on the front page of this form. Complete this only if the details are not shown or are incorrect.

Your name: Mr ☐ Mrs ☐ Miss ☐ Ms ☐

Initials

Surname

Address

Town

County

Postcode

Are you?

Married 01 ☐ Single 02 ☐ Divorced/Separated 03 ☐ Widowed 04 ☐ Living Together 05 ☐

Your partner's name: Mr ☐ Mrs ☐ Miss ☐ Ms ☐

Initials

Surname

What are the dates of birth of: You D D M M Y Y Partner D D M M Y Y

GENERAL INTERESTS

1. Please tick the leisure interests and activities which you and your partner enjoy regularly: *(Please tick all that apply)*

	You	Ptnr		You	Ptnr
Art	01 ☐	31 ☐	Home Computing	61 ☐	91 ☐
Arts & Crafts	02 ☐	32 ☐	Knitting/Sewing	62 ☐	92 ☐
Antiques	03 ☐	33 ☐	Live Sports Events	63 ☐	93 ☐
Astrology/Horoscopes	04 ☐	34 ☐	Model Making	64 ☐	94 ☐
Betting	05 ☐	35 ☐	Motorcycles	65 ☐	95 ☐
Book Reading	06 ☐	36 ☐	Motoring	66 ☐	96 ☐
Bingo	07 ☐	37 ☐	Music-Classical/Jazz	67 ☐	97 ☐
Camping	08 ☐	38 ☐	Music-Pop/Rock	68 ☐	98 ☐
Caravanning	09 ☐	39 ☐	National Lottery	69 ☐	99 ☐
Catalogue Shopping	10 ☐	40 ☐	National Trust	70 ☐	100 ☐
Cigarette Smoking	11 ☐	41 ☐	Nature	71 ☐	101 ☐
Cinema	12 ☐	42 ☐	Nightclubs	72 ☐	102 ☐
Coin/Stamp Collecting	13 ☐	43 ☐	Package Holidays	73 ☐	103 ☐
Collectibles	14 ☐	44 ☐	Personal Finance	74 ☐	104 ☐
Competitions	15 ☐	45 ☐	Pets	75 ☐	105 ☐
Computer Games	16 ☐	46 ☐	Photography	76 ☐	106 ☐
Concerts	17 ☐	47 ☐	Pipe/Cigar Smoking	77 ☐	107 ☐
Cookery	18 ☐	48 ☐	Religious Activity	78 ☐	108 ☐
Crosswords/Puzzles	19 ☐	49 ☐	Science/Technology	79 ☐	109 ☐
Current Affairs	20 ☐	50 ☐	Self Improvement	80 ☐	110 ☐
DIY	21 ☐	51 ☐	Slimming	81 ☐	111 ☐
Eating Out	22 ☐	52 ☐	Theatre	82 ☐	112 ☐
Fashion	23 ☐	53 ☐	The Environment	83 ☐	113 ☐
Football Pools	24 ☐	54 ☐	Visiting Pubs	84 ☐	114 ☐
Foreign Travel	25 ☐	55 ☐	Voluntary Work	85 ☐	115 ☐
Further Education	26 ☐	56 ☐	Watching TV	86 ☐	116 ☐
Gardening	27 ☐	57 ☐	Watching Videos	87 ☐	117 ☐
Grandchildren	28 ☐	58 ☐	Weekend Breaks	88 ☐	118 ☐
Health Foods	29 ☐	59 ☐	Wildlife	89 ☐	119 ☐
Home Baking	30 ☐	60 ☐	Wines	90 ☐	120 ☐

NEWSPAPERS, BOOKS & EDUCATION

2. Which of the following newspapers do you/your partner regularly read?

DAILY	You	Ptnr	SUNDAY	You	Ptnr
Daily Express	01 ☐	15 ☐	Independent on Sunday	29 ☐	43 ☐
Daily Mail	02 ☐	16 ☐	Mail on Sunday	30 ☐	44 ☐
Daily Mirror	03 ☐	17 ☐	News of the World	31 ☐	45 ☐
Daily Record	04 ☐	18 ☐	Sunday Business	32 ☐	46 ☐
Daily Sport	05 ☐	19 ☐	Sunday Express	33 ☐	47 ☐
Daily Star	06 ☐	20 ☐	Sunday Mail	34 ☐	48 ☐
Daily Telegraph	07 ☐	21 ☐	Sunday Mirror	35 ☐	49 ☐
Financial Times	08 ☐	22 ☐	Sunday Post	36 ☐	50 ☐
Independent	09 ☐	23 ☐	Sunday Telegraph	37 ☐	51 ☐
Scotsman	10 ☐	24 ☐	Sunday Times	38 ☐	52 ☐
The European	11 ☐	25 ☐	Sunday Sport	39 ☐	53 ☐
The Guardian	12 ☐	26 ☐	The Observer	40 ☐	54 ☐
The Sun	13 ☐	27 ☐	The People	41 ☐	55 ☐
The Times	14 ☐	28 ☐	Other	42 ☐	56 ☐

3. Typically, who buys your household newspapers? You 01 ☐ Partner 02 ☐

4. Please tick if you have your daily/Sunday newspaper delivered: Daily 01 ☐ Sunday 02 ☐

5. Do you use any of the following to plan your viewing/listening:

	Occasionally	Regularly	Xmas/ Never		Occasionally	Regularly	Xmas/ Never
Daily Newspaper	01 ☐	06 ☐	11 ☐	TV & Satellite Week	16 ☐	21 ☐	26 ☐
Local/Regional paper	02 ☐	07 ☐	12 ☐	TV Times	17 ☐	22 ☐	27 ☐
Radio Times	03 ☐	08 ☐	13 ☐	Weekend Newspaper	18 ☐	23 ☐	28 ☐
Teletext/Ceefax	04 ☐	09 ☐	14 ☐	What's on TV	19 ☐	24 ☐	29 ☐
TV Quick	05 ☐	10 ☐	15 ☐	Other	20 ☐	25 ☐	30 ☐

6. Do you currently subscribe or would you consider subscribing to any of the following magazines?

	Subscribe	Consider		Subscribe	Consider		Subscribe	Consider
Economist	01 ☐	03 ☐	National Geographic	01 ☐	03 ☐	New Scientist	09 ☐	11 ☐
Newsweek	02 ☐	04 ☐	Time	04 ☐	08 ☐	Which?	10 ☐	12 ☐

7. Are you/your partner a member of a book club?

	You	Ptnr		You	Ptnr
Yes	01 ☐	02 ☐	Would Consider	03 ☐	04 ☐

8. What type of music do you/your partner regularly listen to?

	You	Ptnr		You	Ptnr
Classical/Opera	01 ☐	05 ☐	Light Classical	09 ☐	13 ☐
Easy Listening	02 ☐	06 ☐	New Age Music	10 ☐	14 ☐
Folk/Country	03 ☐	07 ☐	Pop/Rock	11 ☐	15 ☐
Jazz	04 ☐	08 ☐	Worship/Gospel Music	12 ☐	16 ☐

9. Do you/your partner have a University Degree? You 01 ☐ Partner 02 ☐

10. If you/your partner are a graduate would you like to receive information on training part-time to become a secondary school teacher? You 01 ☐ Partner 02 ☐

11. Would you/your partner be interested in a government endorsed loan to help you to pay for training or further study to help you in your career?

	You	Ptnr		You	Ptnr
Yes	01 ☐	02 ☐	Possibly	03 ☐	04 ☐

LEISURE, HOLIDAYS AND TRAVEL

12. How many holidays abroad do you normally take each year?
One 01 ☐ Two 02 ☐ Three or more 03 ☐

13. Have you taken (in the last 3 years), or are you considering any of the following types of holiday? *(Please tick all that apply)*

	Taken	May Take		Taken	May Take
Activity	01 ☐	18 ☐	Golfing	35 ☐	52 ☐
Adventure	02 ☐	19 ☐	Hotel Abroad	36 ☐	53 ☐
All Inclusive Abroad	03 ☐	20 ☐	Hotel UK	37 ☐	54 ☐
Apartment/Self Catering Abroad	04 ☐	21 ☐	Lakes & Mountains Abroad	38 ☐	55 ☐
Apartment/Self Catering UK	05 ☐	22 ☐	Motoring Abroad	39 ☐	56 ☐
Boating Abroad	06 ☐	23 ☐	Motoring UK	40 ☐	57 ☐
Camping	07 ☐	24 ☐	Package Holiday Abroad	41 ☐	58 ☐
Canals UK	08 ☐	25 ☐	Package Holiday UK	42 ☐	59 ☐
Caravanning	09 ☐	26 ☐	Religious	43 ☐	60 ☐
Citybreak - UK	10 ☐	27 ☐	Sailing	44 ☐	61 ☐
Citybreak - Europe	11 ☐	28 ☐	Singles	45 ☐	62 ☐
Coach Holiday Abroad	12 ☐	29 ☐	Skiing	46 ☐	63 ☐
Coach Holiday UK	13 ☐	30 ☐	Special Interest	47 ☐	64 ☐
Cottage/Villa Abroad	14 ☐	31 ☐	Summer Sun	48 ☐	65 ☐
Cottage UK	15 ☐	32 ☐	Walking/Trekking	49 ☐	66 ☐
Cruise	16 ☐	33 ☐	Water Sports	50 ☐	67 ☐
Fly Drive	17 ☐	34 ☐	Winter Sun	51 ☐	68 ☐

14. Have you taken (in the last 2 years) or are you considering a holiday (in the next 12 months) to any of the following destinations?

	Taken	Considering		Taken	Considering
Guernsey	01 ☐	04 ☐	Jersey	07 ☐	10 ☐
Ireland	02 ☐	05 ☐	Scotland	08 ☐	11 ☐
Isle of Man	03 ☐	06 ☐	Wales	09 ☐	12 ☐

15. If you have taken a cruise in the last 5 years was it with any of the following?

Carnival Cruise Line	01 ☐	P&O Cruises	05 ☐
CTC Cruise Lines	02 ☐	Orient Cruise Lines	06 ☐
Cunard Line	03 ☐	Royal Caribbean	07 ☐
Princess Cruises	04 ☐	Other cruise line	08 ☐

16. Have you taken (in the last 3 years) or are you considering taking a holiday in:

	Taken	Considering		Taken	Considering
Australia	01 ☐	07 ☐	South America	13 ☐	19 ☐
Canada	02 ☐	08 ☐	Thailand	14 ☐	20 ☐
Caribbean Islands	03 ☐	09 ☐	Other Far East	15 ☐	21 ☐
India	04 ☐	10 ☐	USA – East Coast	16 ☐	22 ☐
Singapore	05 ☐	11 ☐	West Coast	17 ☐	23 ☐
South Africa	06 ☐	12 ☐	Florida	18 ☐	24 ☐

17. Have you taken (in the last 3 years) or would you consider taking (in the next 2 years), a ferry/hovercraft or rail link to any of the following destinations?

	Taken	Considering		Taken	Considering
Belgium	01 ☐	05 ☐	Northern Ireland	09 ☐	13 ☐
Channel Islands	02 ☐	06 ☐	Republic of Ireland	10 ☐	14 ☐
France	03 ☐	07 ☐	Scandinavia	11 ☐	15 ☐
Holland	04 ☐	08 ☐	Spain	12 ☐	16 ☐

18. Which ferry/hovercraft or rail link service(s) have you previously used?

B&I Line	01 ☐	Le Shuttle	05 ☐
Brittany Ferries	02 ☐	P&O European Ferries	06 ☐
Eurostar	03 ☐	Stena Sealink	07 ☐
Hoverspeed/Seacat	04 ☐	Other	08 ☐

19. How many times have you/your partner flown on business in the last 2 years?

	You	Ptnr		You	Ptnr
1-5 times	01 ☐	02 ☐	6-10 times	03 ☐	04 ☐	More than 10 times	05 ☐	06 ☐

20. Have you/your partner travelled to the USA on business in the last 3 years? You 01 ☐ Ptnr 02 ☐

21. Do you own/part own or would you consider buying an overseas property in any of the following?

	Own	Considering		Own	Considering		Own	Considering
Canary Islands	01 ☐	04 ☐	Florida	07 ☐	10 ☐	Malta	13 ☐	16 ☐
Caribbean	02 ☐	05 ☐	France	08 ☐	11 ☐	Portugal	14 ☐	17 ☐
Cyprus	03 ☐	06 ☐	Italy	09 ☐	12 ☐	Spain inc. Balearics	15 ☐	18 ☐

22. Do you currently own or are you considering buying a Timeshare property?

	Currently own	Considering buying		Currently own	Considering buying
In the UK	01 ☐	02 ☐	Abroad	03 ☐	04 ☐

Figure 1. An example of a 'lifestyles' questionnaire.

widest range of purchases (from cars to pet food), leisure pursuits (for example, theatre, fitness), activity patterns (such as propensity to take weekend breaks), health (for example, asthma, backaches), and travel opportunities (for example, details of transport, regular and occasional trip-generating behaviour). These vivid depictions of what is going on in modern Britain stand in increasingly stark contrast to conventional census sources, with their tired socioeconomic classifications (Snowdon, 1998) and all-but-irrelevant surrogate income measures. Lifestyles data are widely used in mail marketing activity and as a basis to 'one-to-one' marketing (Peppers and Rogers, 1997; such activities are arguably more ethnical than area-based targeting, insofar as lifestyles survey respondents often have the option to withhold their address from marketing activities).

In this fast-changing context, lifestyles data offer a number of potential advantages over conventional geodemographic indicators, not only in direct mailings to respondents but also as the basis to area-based generalisations. Yet the negatives associated with such sources are manifold. Some are basic to the design used to mail out questionnaires and are transparent in the analysis of lifestyles data. One such consideration is the common practice of excluding houses in multiple occupation, because residents of such properties tend to be low consumers and in any case are less likely to generate responses because of their high propensities to move. Other problems are more insidious and arise out of the voluntary basis to completion and reliance upon the postal questionnaire survey instrument (characterised by highly variable and unsystematic response rates, incomplete returns, misunderstanding and misinterpretation, etc; Dixon and Leach, 1977). These latter considerations make it doubtful whether lifestyles data are representative of the characteristics and habits of the survey respondents, never mind the broader populace of nonrespondents. Some of the characteristics of lifestyles data vis-à-vis conventional geodemographics are summarised in table 1.

Taking these considerations together, there is a number of respects in which lifestyles data may provide much better digital depictions of human activities and may allow GIS representations to move beyond the static mosaic metaphor of conventional social area analysis. Yet the principles and practice of lifestyles data collection are

Table 1. Some characteristics of geodemographics and lifestyles date in the United Kingdom.

	Geodemographics data	Lifestyles data
Unit of aggregation	Census ED	Household or individual
Population coverage	100% or 10%, depending on variable	10.8%
Sampling	100% or 10% random sample	Self-selecting
Consumption or behaviour indicators	Indirect	More direct
Compatibility with postcode geography	70–80% (using ED-to-postcode directories)	Perfect
Frequency of update	Decennial	Dependent upon data warehouse priorities, but like to be frequent
Bias	The 'missing millions' of the 1991 Census	Part quantifiable if we know where HMOs and non-ER registrants are; response bias, however, is likely to be multivariate and unquantifiable

ED, enumeration district; HMOs, houses in multiple occupation; ER, Electoral Register.

transparently unscientific. Thus the emergent digital data infrastructure provided by lifestyles data may be contemporary and relevant but it is set on far shakier foundations than those of the census and other large-scale public-sector surveys. Lifestyles data sets are data rich, in terms of the number of variables they contain and their relevance to measuring a diversity of lifestyles from hobbies to holidays. But the standards used in the assembly of these data sets fall far short of those required by the 'linear project design' of conventional survey research (Goodchild and Longley, 1999).

If lifestyles data are profoundly unscientific in their collection, is there any contribution that they are likely to make to the geographical analysis of social conditions? And if so, might the richness and diversity of social characteristics that they depict make an important contribution towards new attempts at system-wide urban modelling?[3] Any revitalised systematic analysis of the form and functioning of urban social areas must be founded upon appropriate data models of the spatial form of the city and the constellation of human activities that take place within it. Yet conventional public-sector sources are unlikely to keep pace with the fragmentation and diversity that characterises postindustrial societies—indeed governments sometimes seem to lack the will even to maintain the sedate pace of renewal of conventional public-sector data infrastructure.

We do not concur with those commentators (see Curry, 1995) who have suggested that digital depictions of geographical reality can *never* provide meaningful abstractions. Geodemographic 'cocktails' of census data retain enduring popularity in modelling numerous short-term aspects of consumer behaviour (Birkin, 1995), and repeat purchasing by industry provides ample evidence that they have an important niche role to play. Indeed they are today used in a wider range of business and service planning contexts than ever before (for example, monitoring access to university education; Batey et al, 1999; Utley and Thompson, 1999). But the domain of their application is ultimately limited by the constraints governing content, organisation, and dissemination of national censuses. In contrast, lifestyles databases are up-to-date, relevant, but of dubious scientific validity. Taken together, a best course might require us to identify the degree to which lifestyles data *are* representative of populations at large, in order that we might identify the applications domains within which data-rich models of spatial distributions might be developed as a precursor to further generalised analysis. In the next section we begin to explore the characteristics of a major lifestyles data set to begin to identify whether lifestyles data are sufficiently robust to be considered part of the new digital infrastructure of urban analysis.

4 Lifestyles in Bristol, United Kingdom

Despite the remarkable developments in data capture, warehousing, and application in recent years, there are very few documented examples of linkage of lifestyles data to 'framework' data (Rhind, 1997), such as the census, in any systematic manner (but see Birkin and Clarke, 1995, pages 372–384). We are unaware of any substantial scientific analysis of the content and coverage of lifestyles data—although there has been some informed speculation as to the relative merits of lifestyles and geodemographic analysis (Birkin, 1995; Cosijn and Brown, 1993a; 1993b). This is a glaring omission from the literature, and testimony to the gulf that presently divides data modelling research and practice. Any first attempt to begin to resolve this must necessarily be preliminary and here we will develop a four-point investigation of the feasibility of using lifestyles data

[3] One approach to these questions has been the suggestion that large and complex data sets constitute a fertile application area for a new range of data-mining technologies within the 'geocomputation paradigm' (Openshaw, 1998). However, it seems unlikely that machine intelligence can resolve the biases inherent in the collection of 'nonscientific' quantitative data.

in urban modelling. Given the vagaries inherent in the collection of lifestyles data, our approach will seek to anchor lifestyles analysis to the framework provided by conventional sources—that is, census data and their geodemographic derivatives. As has been suggested above, conventional geodemographics had developed considerably over the last twenty-five years and some of these developments have also entailed additional assumptions. Thus, in addition to establishing the credentials of lifestyles analysis, we will also begin to examine the relative strengths of the assumptions made in contemporary geodemographics versus lifestyles approaches. Thus we will adopt a four-stage preliminary assessment of the scope for urban modelling by using lifestyles data.

(1) We assess the degree to which a lifestyles database might be 'anchored' to 1991 Census data. The census framework provides a potential means of justifying (subject to caveats) the development of household classifications based on lifestyles data—which are likely to be richer and more up-to-date than their existing geodemographic counterparts. Our case study will also seek to establish a basis to generalisation across the wider British system.

(2) We carry out a cluster analysis of lifestyles data and compare the results with a more conventional (SuperProfiles) geodemographic classification.

(3) We investigate the degree of heterogeneity within small (census enumeration district) areas—in order to gain an idea of the fission of consumption activities within small areas and quantify the likely implications for conventional geodemographics.

(4) We compare the detail of the clustered lifestyles data with the 'pen portraits' devised for the 'freshened up' SuperProfiles geodemographic system—as a preliminary assessment of the validity and accuracy of descriptive labels that are not integral to the classification schema and which are vulnerable to scale and aggregation biases.

4.1 Lifestyles: a national and regional snapshot

The case-study data form a subset of a database collated from responses to a national postal questionnaire survey undertaken by a commercial data-warehousing company during September to October 1996. The questionnaire, which took about twenty minutes to complete in full, was mailed to addresses recorded on the February 1996 British Electoral Register (which was based on residence information as at 10 October 1995). One survey was mailed to each address, apart from: addresses with more than three different surnames per register entry (that is, those properties deemed to be houses in multiple occupation, HMOs); and those who opt out from receiving 'mail drops' through the Mail Preference Scheme. Wherever possible the questionnaires were addressed to females or to respondents to previous questionnaires. In total, 20 million questionnaires were mailed across Great Britain, of which approximately 2 million were returned, a response rate of approximately 10%.

The data warehouse anticipates that such surveys will disproportionately enumerate people who are 'mail responsive'. The characteristics of this population, which form the lifestyle database (henceforth, 'lifestyle population'), will therefore likely differ from the 1991 Census-enumerated population (henceforth, 'Census population'). Identification of response bias is difficult to disentangle from sample bias created through exclusion of HMOs and Mail Preference Scheme opt outs, as well as changes in population characteristics between 1991 and 1996. Nevertheless a crude comparison of the age profile of the lifestyles survey with that of the census (table 2) makes the underenumeration of young adults very apparent. Table 3 illustrates the relative underenumeration of individuals living in terraced properties or flats, in contrast to the relative overenumeration of those residing in semidetached property.

The case-study region comprises the 1658 1991 Census enumeration districts (EDs), which have 'BS' (Bristol) postcodes (and shaded grey in figure 3, see over). There were

Table 2. National age profile of lifestyle and census populations (source: authors' calculations based on data supplied by the lifestyles data company).

Age group	Lifestyles	Percentage	Census	Percentage	Difference
18–24	100 531	4	5 614 045	13	−9
25–34	545 143	20	8 361 463	20	0
35–44	574 076	21	7 665 001	18	+3
45–54	519 389	19	6 371 795	15	+4
55–64	420 817	15	5 662 670	13	+2
Over 65	560 146	21	8 811 230	21	0
Total	2 720 102	100	42 486 204	100	

Table 3. Great Britain property types: lifestyles and 1991 Census analysis (source: authors' own calculations based on data supplied by the lifestyles data company).

	Lifestyles	Percentage	Census	Percentage	Difference
Detached	307 704	19	4 383 168	20	−1
Semidetached	613 897	37	6 481 705	30	+7
Terraced	449 203	27	6 343 156	29	−2
Flat or maisonette	284 307	17	4 389 770	20	−3
Total	1 655 111	100	21 597 799	100	

73 310 individual and adult respondents to the lifestyle survey in this region, from 51 882 households. The data were supplied in a form that allows the location of the households to be identified to the level of the unit postcode (for example, BS8 1SS). Unit postcode geography is not coincident with that of the 1991 Census in England and Wales and so it was necessary to use the 1991 and 1995 enumeration district-to-postcode directories to obtain an approximate match (see Martin, 1992). In practice, where a 'postman's walk' (the basis of definition of UK unit postcodes) is identified as crossing the boundaries of one or more EDs, then the postcode was deemed to be located in the 'pseudo-ED' in which the directories deem the majority of the postcode population to lie. Approximately one quarter of unit BS postcodes cross ED boundaries, although this does not create any problems if the two or more EDs share common attributes (such as the same geodemographic category, which is assigned at the ED level). Nevertheless the mismatch between census and postal geographies creates ambiguity in assigning SuperProfile classes to addresses in 15% of all BS unit postcodes.

The adult census population of the study region comprises 623 132 individuals, implying (if all survey respondents were adults) that the survey was completed by 11.8% of the adult population. This is a large proportion—and larger than the absolute number of census returns used to compile 10% census returns at the ward scale (for example, for occupational data). However, 'excluded' members of HMOs (to which no surveys were mailed, amounting to 1.7% of all households in the study region) aside, the lifestyle respondents are entirely self-selecting. Hitherto there has been no research into the problems inherent in generalising from self-selecting lifestyles samples, and any attempt flies in the face of scientific approaches to statistical generalisation. Survey research practice is rightly and avowedly sceptical of the dangers of postsurvey stratification and differential grossing in light of subgroup response rates (see Moser and Kalton, 1993). Response biases are also likely to be compounded by partial completion of survey forms by respondents. This is most evident with respect to measurements of income—20% of survey respondents in the survey used here did not state their incomes,

Table 4. Regional and national enumeration district breakdowns according to SuperProfiles cluster [source: the SuperProfiles (SP) ten-cluster typology, after Brown and Batey (1994) and the 1991 UK Census].

Cluster (economic rank)	Percentage in study region	Percentage in Great Britain	Percentage in England and Wales	Percentage in England and Wales (not London)
Affluent achievers (SP1)	10	10	10	10
Thriving greys (SP2)	13	11	12	13
Settled suburbans (SP3)	12	12	12	12
Nest builders (SP4)	18	16	16	18
Urban venturers (SP5)	11	10	10	6
Country life (SP6)	2	3	3	3
Senior citizens (SP7)	8	7	6	7
Producers (SP8)	16	12	12	13
Hard-pressed families (SP9)	3	8	7	8
The 'have nots' (SP10)	5	12	11	10
Pearson correlation		0.79	0.85	0.82

a figure that is broadly in line with the results of National Office of Survey trials for an income question in the 2001 Census.

Table 4 compares the incidence of the ten SuperProfiles geodemographics categories in the study region with some broader aggregations. Bristol and environs is broadly representative of the broader national picture, as indicated by the Pearson correlation coefficients, although the 'lower status' SuperProfiles (clusters SP9 and SP10) are underrepresented in the study region. The smallest category in the national classification (cluster SP6) is also underrepresented in and around Bristol. This finding is also substantiated by reference to other composite geodemographic sources (CACI Information Services, 1998; see also Harris, 1998). However, the characteristics of the lifestyle respondents differ from those recorded in the 1991 Census, and only part of the observed discrepancy can be attributed to the five-and-a-half year interregnum between the census and completion of the lifestyle survey. We have chosen not to try to estimate the extent to which this is a consequence of excluding households in multiple occupation from the lifestyles survey, as the necessary grossing would risk committing ecological fallacy. Figure 2 shows the percentage of the regional populations of each lifestage group according to the lifestyles data set and according to the census. It is apparent that, as with the national picture, it is the young who do not return lifestyle questionnaires. As Rae (1998, page 6) has commented, referring to the data of figure 2, "groups that are equally common in the population[-at-large] have radically different representation

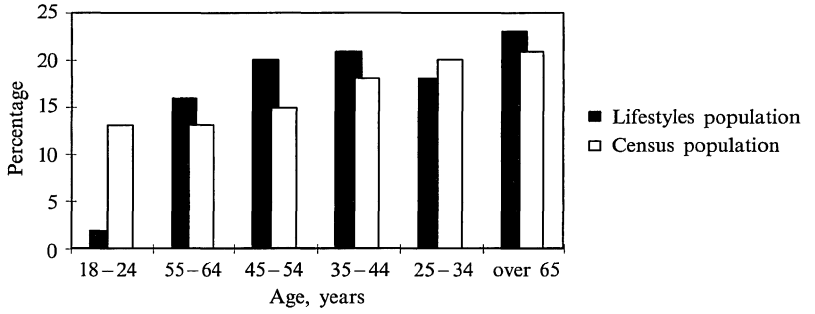

Figure 2. Representation of age groups by using lifestyles and the census.

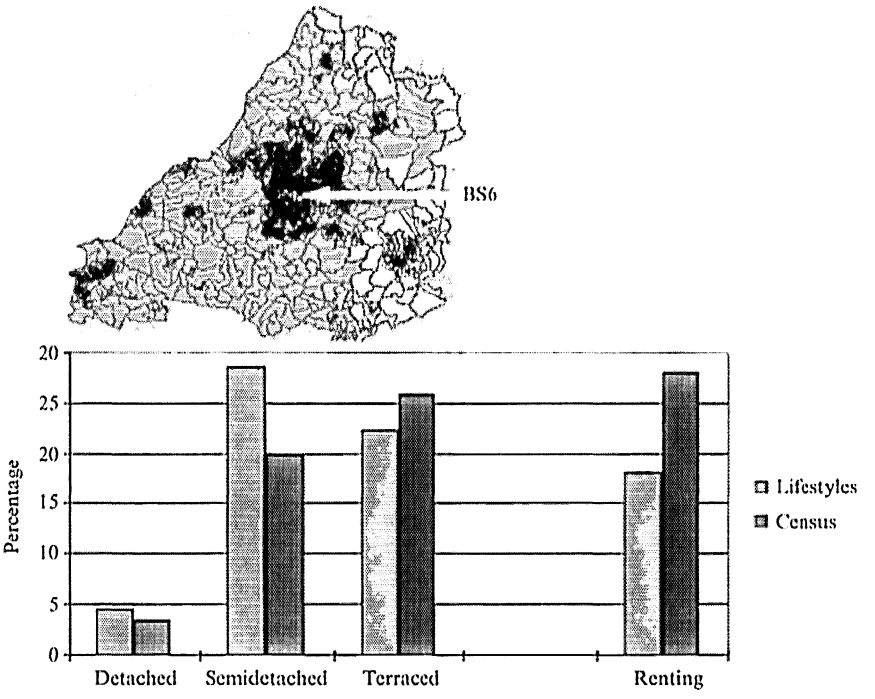

Figure 3. Comparative analysis of numbers of 'BS6' households in particular property types and tenure, estimated by using the lifestyles data and the 1991 Census.

in lifestyle data"—compare the 18–24 and 55–64 age bands, for example. Such differences could lead to misleading area profiles (see figure 3; and for further details see Harris, 1998).

4.2 Classification of a lifestyles data set

An iterative cluster analysis was developed by using the lifestyles data at the household level (for details, see Harris, 1998). The household 'response rate' for the survey was 16%. As with the creation of conventional geodemographic clustering, the outcome of clustering obviously depends upon the range and type of variables included in the analysis (Openshaw, 1996). We used 241 variables, chosen to represent a wide range of socioeconomic characteristics and behavioural information, as the basis for a clustering procedure. The cluster program performed best when the data were divided into sixteen clusters. The characteristics of the sixteen clusters are summarised in tables 5 and 6.

Table A1 in the appendix shows the principal defining characteristics of the clusters. They encompass a far broader range of household and individual characteristics than conventional census-based geodemographic indicators and it is interesting that table A1 reveals the importance of leisure, holiday, and consumption interests—and also other characteristics such as health. Indeed, in a number of instances, groups seem at least as much tied together by consumption as by conventional age, socioeconomic status, and family cycle considerations, if not more so. As with all conventional geodemographic classifications, and as noted above, the nature of the end classification is conditioned foremost by the nature and range of the input variables. The input variables used here are much more suggestive than conventional geodemographic indicators of whether people are sedentary, limited in physical mobility, participants in neighbourhood or city-wide activities, patronise 'traditional' or out-of-centre retailing,

Table 5. Variables used in the formation of the household typology.

Type of variable	Number of variables
Age of household member	6
Alcoholic beverages consumed	9
Children: number in household and age	8
Consumer goods owned	7
Daily newspaper read	10
Household income	7
Financial investments and plans	15
Gender	2
Have credit cards, store cards, etc	7
Hobbies and pastimes	32
Holiday choices	22
Home improvements made	11
Home type, tenure, and value	18
Household size	4
Illnesses	9
Duration of residence	6
Mail order purchases	7
Marital status	4
Charity support	14
Number of cars owned, make, and value	23
Smoking	2
Social-economic group	5
Supermarkets regularly visited	10
Other	3

and so forth. On average, across the sixteen clusters, any one household value would share the same (yes or no) value as its cluster for 84% of the 241 variables. As with any cluster analysis, this does not necessarily represent *the* optimal solution but *an* optimised solution. It is also important that the results of the cluster analysis make sense in substantive terms.

4.3 Lifestyles and geodemographics: competing or complementary classifications?

The results of the lifestyles classification were aggregated into EDs in order to facilitate comparison with the ED-scale SuperProfiles for the study region. With regard to the lifestyles data, EDs were given the lifestyles descriptor which pertained to the largest absolute number of households in the ED. Table 7 (see over) shows how households in each SuperProfile category are spread across the sixteen lifestyle groups; and table 8 (see over) shows the spread of each lifestyle category across the SuperProfile categories. (χ^2 analysis confirms that lifestyle groups are not uniformly distributed across all the SuperProfile categories; Harris, 1998.) These tables suggest strong correspondence between the two classifications, with most SuperProfile categories being spread out between two or three lifestyle categories, and vice versa. This is an important finding, on at least two counts. First, the representation of all SuperProfile groups in some shape or form suggests that there are no gaping holes in the classification (arising particularly because of the underrepresentation of the young in the lifestyles survey). Second, it follows that classifications that are richer than conventional geodemographics can be built at disaggregate scales.

Differences between the two classifications are likely to have arisen from the following, alone or in combination.

Table 6. The sixteen consumer clusters (summarised).

Cluster	Percentage[a]	Summary
A	5	Wealthy older couples in upgraded homes and with diverse financial investments that afford many pastimes and regular holidays
B	4	Computer-friendly couples and families with financial provision for retirement. Church attendees with interest in the arts
C	3	Comfortable couples living in improved homes with PCs and satellite TV. Suffer from aches and pain and holiday in 'the Med'
D	5	Comfortable city-dwelling (older) couples. Long-time residents of improved homes, within which their interests are pursued
E	6	Younger couples and families living in improved semidetached properties. Holiday in the United Kingdom on camping and caravan trips
F	7	Other couples in owner-occupied properties and with no children
G	8	Affluent retired couples living in upgraded (city) homes. Diverse financial investments afford overseas and UK holidays
H	7	Younger and middle-aged couples in improved homes. Financially comfortable with provision for retirement. Have home PCs and gamble upon the Pools or National Lottery
I	5	Young outgoing singles. Have active social lives and often frequent holidays
J	6	Couples and families taking few holidays and gambling upon the Pools or National Lottery. Do not smoke
K	5	Mail-order responsive, lower-income couples residing in the city. Suffer from stress or other aches and pains. Smoke
L	4	Female residents living alone and with home-based interests
M	11	Low-income retired couples
N	5	Elderly female widows
O	11	Other lone-female households
P	6	Low-income, single females and single mothers, living in housing association or local authority properties

[a] Percentage of Bristol population.

(a) Inherent differences in the data—that is, differences arising out of the different constructs measured in the data sets, temporal changes between the two surveys, and the effects of response and sampling bias in the lifestyles data set.
(b) The effects of aggregation—as Birkin (1995) has pointed out, geodemographic classifiers may be misleading if used to suggest that ED labels pertain to every household within each classified ED, because this is patently almost invariably not the case in reality. This is an inherent problem in geographical classification and analysis, which is only ultimately resolvable through recourse to individual or household units of analysis (Openshaw, 1984). Information has been discarded in the compilation of tables 7 and 8, in that a simple 'highest count' rule has been used to label every ED with a lifestyle category.

It is difficult to disentangle these different considerations, although what is evident from figure 4 (see over) is a quite staggering diversity of lifestyles within EDs. This kind of small-area heterogeneity is hidden in conventional geodemographic analysis, yet the clear implication is that the mosaic of small areas used in conventional geodemographic analysis conceals considerable diversity. Moreover, table 9 (see over) shows that the EDs assigned to different SuperProfile categories are characterised by different degrees of diversity—with the most affluent EDs (in SuperProfile terms) characterised by the greatest degree of diversity. This table suggests that almost all census areas are neither ghettos of 'have nots' nor islands of 'affluent achievers'—and that prescriptive urban

Table 7. The spread of SuperProfiles across the lifestyle groups (household analysis).

Percentage	A	B	C	D	E	F	G	H
SP1	25.0	32.7	13.0	20.0	8.5	21.8	21.5	6.3
SP2	15.4	21.2	13.0	11.1	12.2	5.9	34.5	1.0
SP3	23.1	1.9	13.0	28.9	36.6	6.9	14.1	17.7
SP4	15.4	19.2	39.1	24.4	32.9	23.8	11.9	55.2
SP5	3.8	19.2	4.3	2.2	0.0	21.8	0.6	4.2
SP6	5.8	1.9	0.0	0.0	0.0	5.9	0.0	0.0
SP7	5.8	3.8	0.0	0.0	0.0	4.0	5.6	2.1
SP8	3.8	0.0	13.0	11.1	7.3	8.9	11.3	10.4
SP9	0.0	0.0	4.3	2.2	2.4	1.0	0.6	2.1
SP10	1.9	0.0	0.0	0.0	0.0	0.0	0.0	1.0
Sum	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0
	I	J	K	L	M	N	O	P
SP1	3.1	0.0	0.0	0.0	9.3	0.0	3.7	0.0
SP2	12.2	13.3	13.3	13.3	14.6	17,1	7.9	1.9
SP3	5.1	6.7	6.7	0.0	15.3	11.4	7.5	1.9
SP4	19.4	30.0	30.0	20.0	15.3	11.4	13.3	4.8
SP5	35.7	26.7	26.7	40.0	0.7	5.7	7.9	8.7
SP6	0.0	0.0	0.0	0.0	0.0	0.0	0.8	0.0
SP7	16.3	3.3	3.3	20.0	7.8	42.9	14.5	8.7
SP8	6.1	16.7	16.7	6.7	32.0	8.6	27.4	23.1
SP9	2.0	0.0	0.0	0.0	2.8	0.0	7.9	14.4
SP10	0.0	3.3	3.3	0.0	2.1	2.9	9.1	36.5
Sum	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0

Table 8. The spread of lifestyle groups across SuperProfiles (household analysis).

Percentage	A	B	C	D	E	F	G	H	
SP1	8.5	11.1	2.0	5.9	4.6	14.4	24.8	3.9	
SP2	4.2	5.8	1.6	2.6	5.2	3.1	31.9	0.5	
SP3	6.4	0.5	1.6	7.0	16.0	3.7	13.4	9.1	
SP4	2.8	3.5	3.1	3.8	9.3	8.3	7.3	18.3	
SP5	1.6	8.1	0.8	0.8	0.0	17.9	0.8	3.3	
SP6	25.0	8.3	0.0	0.0	0.0	50.0	0.0	0.0	
SP7	2.4	1.6	0.0	0.0	0.0	3.2	8.0	1.6	
SP8	0.8	0.0	1.2	2.0	2.4	3.5	7.8	3.9	
SP9	0.0	0.0	1.8	1.8	3.6	1.8	1.8	3.6	
SP10	1.4	0.0	0.0	0.0	0.0	0.0	0.0	1.4	
	I	J	K	L	M	N	O	P	Sum
SP1	2.0	0.0	0.0	0.0	17.0	0.0	5.9	0.0	100
SP2	6.3	2.1	2.1	1.0	21.5	3.1	9.9	1.0	100
SP3	2.7	1.1	1.1	0.0	23.0	2.1	9.6	1.1	100
SP4	6.6	3.1	3.1	1.0	14.9	1.4	11.1	1.7	100
SP5	28.5	6.5	6.5	4.9	1.6	1.6	15.4	7.3	100
SP6	0.0	0.0	0.0	0.0	0.0	0.0	16.7	0.0	100
SP7	12.8	0.8	0.8	2.4	17.6	12.0	28.0	7.2	100
SP8	2.4	2.0	2.0	0.4	35.3	1.2	25.9	9.4	100
SP9	3.6	0.0	0.0	0.0	14.5	0.0	34.5	27.3	100
SP10	0.0	1.4	1.4	0.0	8.5	1.4	31.0	53.5	100

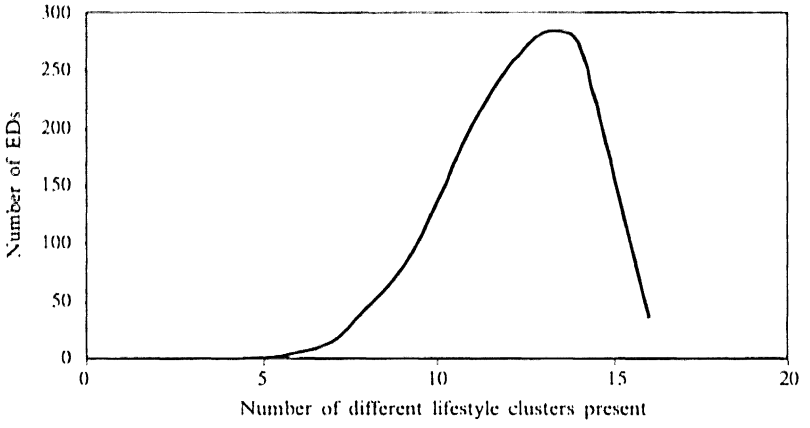

Figure 4. The number of household consumer types per enumeration districts (EDs).

Table 9. The degree of heterogeneity of enumeration district (ED) lifestyles, according to Super-Profile category.

	Number of EDs	1	2	3	4
Affluent achievers	153	1.3	22.2	68.0	8.5
Thriving greys	191	4.2	29.8	54.5	11.5
Settled suburbans	187	1.1	17.1	62.0	19.8
Nest builders	289	0.3	19.0	56.4	24.2
Urban venturers	123	13.0	41.5	43.1	2.4
Country life	12	8.3	66.7	16.7	8.3
Senior citizens	125	9.6	35.2	45.6	9.6
Producers	255	2.0	29.4	58.0	10.6
Hard-pressed families	55	3.6	47.3	43.6	5.5
The 'have-nots'	71	25.4	54.9	18.3	1.4

Note:
1, least heterogeneous (8 or less household types per ED);
2, ⋮ (9 to 11 household types);
3, ⋮ (12 to 14 household types);
4, most heterogeneous (15 or more household types per ED);
5, table includes only the 1461 (of 1568) EDs in which 17 or more households were enumerated by the lifestyles survey.

modelling should move away from such crude conceptions of social patterning and neighbourhood function. The analysis does not present a direct comparison in that we have compared a geodemographic classification based on data that have been aggregated to the ED level with a lifestyles classification that has been based on individual observations. The lifestyles classification has subsequently been aggregated to the ED scale as a convenience to facilitate comparison, and this comparison has the advantage of making use of the most available detail from each of the two data sets. More direct comparisons could be made in either of two ways. First, we might compare the classification of individual lifestyles data with the results of a cluster analysis of individual-level census data (the Sample of Anonymised Records) at the district scale. Or, second, we might aggregate the lifestyles data to the ED level prior to clustering and then compare the results with the cluster analysis of census data. These should each be the focus of further research.

4.4 'Pen portraits': an outline appraisal

The classification typology developed above is grounded in data and, the vagaries of the conduct and response to the lifestyle survey aside, in an analytical rigorous way. The more recent geodemographic systems have sought to 'freshen up' census-based classifications with reference to ancillary data sources that are more recent, relevant, and detailed with respect to consumer behaviour. Such sources have variously included the General Household Survey, Family Expenditure Survey, National Readership Survey, and a range of service industry sources. This potentially brings a wealth of detail to geodemographic classification, although not at the fine spatial scales for which classifiers are used to discriminate behavioural types—indeed they are rarely capable of statistically valid comparison at finer spatial scales than the district level. Thus, in practice, ancillary sources remain *external* to the classification procedure but are used to provide 'thick descriptors' of the classification ex post facto. Of course this procedure is inherently unscientific and potentially introduces a number of scale and aggregation-induced effects into the interpretation of classifications.

Yet words are more seductive than numbers and the resulting 'pen portraits' add intuitive plausibility to classifications. We can see this, for example, in the SuperProfiles category of 'Affluent Achievers':

> "*High income families* with a lifestyle to match. *Detached houses predominate*, reflecting the professional status of their owners. Typically living in the stockbroker belts of the major cities, the Affluent Achiever is likely to own *two* or more *cars*, which are the top of the range, recent purchases, and are needed to pursue an active social and family life.
>
> Affluent Achievers have sophisticated tastes and aspirations. They *eat out regularly, go to the theatre* and opera and take an active interest in sports (such as cricket, rugby union, and golf). They are *able to afford several expensive holidays* every year. *Financially aware*, with a high disposable income, this group *invests in both quoted and privatised companies. They are likely to use credit and charge cards and are likely to have private health insurance.* Investments are followed closely in broadsheets, such as *The Financial Times*, *The Times*, and *The Telegraph.* For more leisurely reading, *Hello*, *Harpers & Queen*, and *Vogue* are likely to be found in the home of the Affluent Achiever" [source: promotional literature (cited by Brown and Batey (1994). Italics added to highlight similarities with table A1].

Hyperbole aside, this pen picture accords with the characteristics of type A consumer group identified in the lifestyles classification shown in table A1. This suggests superficial correspondence between the lifestyles and geodemographic classifications, yet conceals a more heterogeneous reality. Of the 231 high-concentration EDs for consumer type A, 60 (26%) are found in SP1, another 48 (21%) are in SP2, and an absolute majority (53%) of such households are scattered across the remaining eight clusters! Further, a 'high concentration' need mean only that about 10% of the consumer-classified households within the ED are of consumer type A. It would thus be fallacious indeed to characterise the whole ED area as this consumer type when 90% of the households are of a different consumer type. The magnitude of these differences suggests that lifestyles analysis should be used to supplement, even replace, conventional geodemographic typologies.

5 Discussion

Previous urban models have been deficient not because data were 'unscientific' in collection but principally because the data models on which they were founded were outdated, pertained only to coarse zonal aggregations and, perhaps most fundamentally, provided only very imperfect and indirect indicators of human decisions and

activity patterns. These deficiencies have become more apparent over time, as the scale, complexity, and diversity of society has increased, and have resulted in two dominant views of urban modelling in the research community. First, some formal urban analysis has 'carried on regardless', yet today undoubtedly accounts for a much reduced real share of intellectual activity in academic geography and planning. Second, some critics have suggested that digital data and urban models can never be up to the task of generating understanding of real-world problems and that the research effort should be channelled into other (by implication, more idiographic and/or small-scale) approaches. Thus academic discourse has become increasingly polarised between those who cling to lingering but increasingly marginal scientific certainties about data and those who refute the notion of any valid domain for quantitative analysis.

However, a third course is developing, as a result of the explosion in the extent and availability of digital data and the improvement in geographical data-handling technologies. This is based on the view that technology does not just cause consumption to fragment but it also empowers us to provide ever richer depictions of the diversity of population characteristics and behaviour within city systems. This is very much the view advocated in this paper and is also consistent with microsimulation approaches (Clarke, 1996). Indeed, following Johnston (1999), we suggest that data-rich GIS-based model building may be poised to move beyond the 'mosaic metaphor' which has governed almost all applications of GIS to date towards more convincing depictions of variety in space and time that are consistent with new theory. Description coming before theory is the normal pattern in science and the spirit of what we have described here is very much that data models can be made sensitive to context without sacrificing generality. Such an approach is also consistent with a reinvigorated contribution to rational planning policy.

A problem with this view is that the foundations to data-rich analysis, at least those explored here, are clearly unscientific. Yet our tentative empirical investigation has suggested that clear commonalities may be established between 'framework' data, such as the census and geodemographic systems, and new, relevant, and timely lifestyles data. We believe that application of concatenation and conflation procedures (Longley and Goodchild, 1999) to lifestyles data sets offers the prospect of creating vastly enhanced data models of the form and functioning of urban systems. Such models will need to manage error and bias and this may cause some unease in a modelling community that has been more focused on statistical formalism than messy empirical data problems. The social theory fraternity will doubtless be able to identify aspects of anecdotal historiographies that cannot (yet) be represented in digital form. But real-world business and service planning is already embracing the use of such data series and is using lifestyles data as successfully as their geodemographics forbear. The linear project design of conventional social scientific research was never a panacea in practice (Goodchild and Longley, 1999) and the subsequent flight to untested social theory has done much to marginalise academic contribution to rational planning policy. Today's digital data infrastructure is not by any means perfect but it has much to offer a reinvigorated approach to urban modelling.

Acknowledgements. The authors are very grateful to a lifestyles data company for the supply of data, and to Peter Brown for the supply of information pertaining to the SuperProfiles geodemographic system. This paper was prepared for the CUPUM '99 Conference, Venice, Italy.

References

Batey P W J, Brown P J B, 1995, "From human ecology to customer targeting: the evolution of geodemographics", in *GIS for Business and Service Planning* Eds P Longley, G Clarke (GeoInformation International, Cambridge) pp 77 – 103

Batey P W J, Brown P J B, Corver M, 1999, "Expanding participation in higher education: a geodemographic analysis of recent experience in the English regions" *Journal of Geographical Systems* forthcoming

Batty M, 1981, "Urban models", in *Quantitative Geography: A British View* Eds N Wrigley, R J Bennett (Routledge and Kegan Paul, London) pp 181 – 191

Batty M, Longley P, 1994 *Fractal Cities: A Geometry of Form and Function* (Academic Press, London)

Batty M, Longley P, 1997, "The fractal city" *Architectural Design Profile* **129** 74 – 83

Beaumont J R, 1991 *An Introduction to Market Analysis*, in the *Concepts and Techniques in Modern Geography* series, number 53 (Environmental Publications, Norwich)

Birkin M, 1995, "Customer targeting, geodemographics and lifestyles approaches", in *GIS for Business and Service Planning* Eds P Longley, G Clarke (GeoInformation International, Cambridge) pp 104 – 149

Birkin M, Clarke G P 1995, "Using microsimulation methods to synthesize census data", in *Census Users' Handbook* Ed. S Openshaw (GeoInformation International, Cambridge) pp 363 – 387

Brown P J B, 1991, "Exploring geodemographics", in *Handling Geographical Information* Eds I Masser, M J Blakemore (Longman, Harlow, Essex) pp 221 – 258

Brown P J B, Batey P W J, 1994, "Characteristics of Super Profile lifestyles and target markets: index tables, pen pictures and geographical distributions", Super Profile Technical Note 2, Urban Research and Policy Evaluation Regional Research Laboratory, University of Liverpool, Liverpool

CACI Information Services, 1998 *The Geodemographic Pocket Book* (NTC Publications, Henley-on-Thames)

Clark B D, Gleave M B (Eds), 1973 *Social Patterns in Cities* Institute of British Geographers Special Publication 5 (Institute of British Geographers, London)

Clarke G P (Ed.), 1996 *Microsimulation for Urban and Regional Policy Analysis* (Pion, London)

Cole K, 1993, "The 1991 local base and small area statistics", in *The 1991 Census User's Guide* Eds A Dale, C March (The Stationery Office, London) pp 201 – 247

Cosijn P N C, Brown P J B, 1993a, "Comparison of an individual level and an aggregate level market analysis system: performance of the lifestyles selector and Super Profiles classification. Part 1: Sources and classification of data" *Journal of Targeting, Measurement and Analysis for Marketing* **1** 257 – 272

Cosijn P N C, Brown P J B, 1993b, "Comparison of an individual level and an aggregate level market analysis system: performance of the lifestyles selector and Super Profiles classification. Part 2: Discriminatory and penetrative power" *Journal of Targeting, Measurement and Analysis for Marketing* **1** 355 – 365

Curry M R, 1995, "GIS and the inevitability of ethical inconsistency", in *Ground Truth: The Social Implications of Geographical Information Systems* Ed. J Pickles (Guilford Press, New York) pp 68 – 87

Dixon C, Leach B, 1977 *Sampling Methods for Geographical Research* Concepts and Techniques in Modern Geography series, number 17 (Geo Books, Norwich)

Flowerdew R, Leventhal B, 1998, "Under the microscope" *New Perspectives* **18** 36 – 38

Goodchild M F, Longley P A, 1999, "The future of GIS and spatial analysis", in *Geographical Information Systems: Principles, Techniques, Management and Applications* Eds P A Longley, M F Goodchild, D J Maguire, D W Rhind (John Wiley, New York), volume 1, pp 567 – 580

Goss J, 1995, "Marketing the new marketing: the strategic discourse of geodemographic information systems", in *Ground Truth: The Social Implications of Geographic Information Systems* Ed. J Pickles (Guilford Press, New York) pp 130 – 170

Harris R, 1998, "A comparative analysis of a lifestyle and geodemographic typology", WP-2, School of Geographical Sciences, University of Bristol, Bristol; http://www.bitsmart.com/geodemographics

Harvey D, 1973 *Social Justice and the City* (Edward Arnold, Sevenoaks, Kent)

Johnston R J, 1999, "Geography and GIS", in *Geographical Information Systems: Principles, Techniques, Management and Applications* Eds P A Longley, M F Goodchild, D J Maguire, D W Rhind (John Wiley, New York), volume 1, pp 39 – 47

Longley P, 1998, "GIS and the development of digital urban infrastructure" *Environment and Planning B: Planning and Design* **25** 53 – 56

Longley P A, Goodchild M F, 1999, "Modern geographic information systems and model linking"; copy available from the author

Marsh C, 1993, "The validation of census data: general issues", in *The 1991 Census User's Guide* Eds A Dale, C March (The Stationery Office, London) pp 155 – 167

Martin D, 1992, "Postcodes and the 1991 Census of Population: issues, problems and prospects" *Transactions of the Institute of British Geographers, New Series* **17** 350 – 357

Moser C A, Kalton G, 1993 *Survey Methods in Social Investigation* second edition (Heinemann, Oxford)

Openshaw S, 1984 *The Modifiable Areal Unit Problem* Concepts and Techniques in Modern Geography series, number 38 (Geo Books, Norwich)

Openshaw S, 1996, "Developing GIS-relevant spatial analysis methods", in *Spatial Analysis: Modelling in a GIS Environment* Eds P Longley, M Batty (GeoInformation International, Cambridge) pp 55 – 73

Openshaw S, 1998, "Building automated geographical analysis and explanation machines", in *Geocomputation: A Primer* Eds P A Longley, S M Brooks, R McDonnell, W D Macmillan (John Wiley, Chichester, Sussex) pp 95 – 115

Pacione M, 1999 *Applied Geography* (Routledge, London)

Peppers D, Rogers M, 1997 *The One-to-one Future: Building Business Relationships One Customer at a Time* (Piatkus Books, London)

Rae J, 1998, "Why do you have to mix lifestyle and census data?" *Marketing Systems Today* **13** 6 – 7

Rhind D, 1997 *Framework for the World* (GeoInformation International, Cambridge)

Sayer R A, 1979, "Understanding urban models versus understanding cities" *Environment and Planning A* **11** 853 – 862

Snowdon R, 1998, "New system could spell end for ABCs" *The Times* 15 November, page 31

Stapleton C M, 1980, "Reformulation of the family life-cycle concept: implications for residential mobility" *Environment and Planning A* **12** 1103 – 1118

Timms D, 1969 *The Urban Mosaic: Towards a Theory of Residential Differentiation* (Cambridge University Press, Cambridge)

Utley A, Thompson A, 1999, "Wraps of v-cs' postcode database" *Times Higher Education Supplement* 29 January, page 3

APPENDIX

How to read the table

Taking the 18–24 age group as an example, a global mean of 2% of households per cluster are shown (by the lifestyles data) to have an adult aged 18–24 years resident. [Here 100 is an index value assigned to the global mean (GM) across all the clusters for a given variable.] These young adults are concentrated within cluster I: 'young outgoing singles'. This cluster has 6.5 times (650/100) the average proportion of young adult households, so 13% (650/100×2) of households in cluster I have a person aged 18–24 resident. By comparison, cluster N ('elderly female widows') has no young adult households. Instead, there is an above-average proportion of households with at least one member aged 65 or above.

Table A1

	GM	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P
Age																	
18–24 years	2	0	0	50	50	50	100	0	100	650	150	150	100	0	0	150	450
25–34 years	19	52	100	89	110	152	131	0	210	236	163	157	110	5	0	57	152
35–44 years	22	127	186	145	136	186	131	9	218	95	140	150	118	9	0	54	109
45–54 years	21	180	171	209	176	142	128	42	109	61	123	142	95	33	9	61	80
55–64 years	17	182	88	117	117	76	100	211	35	23	100	117	88	135	52	76	70
65 or above	24	58	16	8	16	4	12	283	0	4	16	29	37	312	350	100	29
Gender																	
Female in household	80	108	107	110	106	105	90	106	107	75	101	112	111	98	98	92	95
Male in household	66	131	119	113	113	122	112	122	128	80	127	122	21	121	31	62	54
Marital status																	
Couple	59	140	137	130	132	137	122	138	138	11	130	123	15	135	5	50	25
Divorced	11	54	63	81	54	63	63	36	63	118	63	90	345	27	81	145	318
Single	16	43	50	50	75	43	93	18	43	468	75	81	212	25	75	131	237
Widowed	8	25	0	12	12	0	12	112	0	0	0	12	125	87	875	200	75
Household size																	
1 adult	20	35	25	35	40	20	25	55	30	215	30	40	350	45	400	100	290
2 adults	54	125	137	116	114	131	124	133	144	48	133	127	25	133	11	328	40
3 adults	10	140	110	160	150	130	110	90	80	120	100	140	60	70	20	80	70
4 adults	4	175	100	225	200	125	150	50	75	200	100	125	50	50	0	75	75
Children																	
No children	62	117	69	108	108	66	101	148	46	130	88	95	108	140	145	54	83
1 child in household	9	77	188	122	100	188	100	11	222	66	133	133	122	11	0	88	177
2 children in household	12	108	208	125	133	216	116	0	291	25	141	116	91	8	0	58	116
3 children in household	4	75	200	75	125	200	100	0	250	25	150	175	75	0	0	50	175
4 children in household	1	100	200	100	100	100	100	0	100	0	200	300	100	0	0	100	300
Age																	
0–4 years	9	77	166	88	88	200	100	0	288	33	177	155	55	0	0	66	166
5–10 years	13	76	207	92	123	184	115	7	261	30	138	153	107	7	0	76	184
11–15 years	11	136	218	136	127	200	100	9	245	36	127	154	118	18	0	72	145
Economic status																	
Employed	47	138	159	163	161	163	155	23	172	157	153	140	80	17	4	38	55
Self-employed	8	150	212	175	150	175	150	50	175	87	137	137	112	37	0	62	62
Unemployed	4	25	25	50	50	50	75	0	25	50	125	175	100	50	0	175	475
Student	1	0	200	0	100	100	100	0	100	600	100	100	300	0	0	200	300
Retired	32	93	25	21	28	18	18	271	6	9	28	50	50	268	287	100	43
Family income per annum																	
Under £5000	10	10	0	10	10	10	10	70	10	30	30	80	140	150	360	180	390
£5000–£9000	14	28	14	28	35	28	50	178	28	50	71	128	150	214	200	114	150
£10000–£14900	15	80	53	66	86	80	93	173	93	133	120	153	140	106	60	60	80
£15000–£19900	12	133	91	108	141	141	133	100	175	150	150	150	108	58	25	33	41
£20000–£29900	18	177	161	166	172	177	127	61	200	138	127	100	72	27	11	22	11
£30000–£39900	6	266	316	233	200	216	166	33	150	133	83	66	50	0	0	16	0
£40000 and above	4	350	550	325	250	175	175	25	100	125	75	25	25	0	0	25	0

	GM	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P
Property																	
1 bedroom	5	20	40	40	40	0	80	20	20	240	80	80	200	60	280	180	300
2 bedrooms	19	42	57	78	73	26	89	89	47	157	131	100	152	89	168	105	173
3 bedrooms	49	112	69	83	93	144	87	124	144	75	89	134	83	128	67	61	69
4 bedrooms	13	207	261	238	207	115	153	107	100	69	107	46	76	53	46	38	23
5 bedrooms	3	200	466	166	200	100	166	100	66	100	100	33	100	33	33	66	33
City residence	44	97	102	77	134	86	100	127	140	77	95	125	90	93	90	90	79
Conservatory	10	260	90	200	170	150	40	260	140	40	50	70	50	90	60	20	10
Double glazing	42	200	57	176	176	183	38	197	176	57	61	76	57	76	64	28	38
Fitted bathroom	38	228	71	200	205	192	36	202	186	57	44	68	52	52	55	23	26
Fitted bedroom	21	319	47	200	185	185	28	219	166	42	47	71	47	47	47	14	19
Fitted carpets	26	257	76	180	173	176	50	169	180	73	61	100	65	57	42	26	46
Fitted kitchen	44	209	75	190	190	186	40	195	179	63	59	75	61	59	59	25	31
Fitted shower	43	204	74	188	186	181	41	190	174	65	55	81	60	62	55	25	34
Loft conversion	5	300	140	180	220	160	40	160	140	40	60	60	40	40	20	20	20
New driveway	23	282	82	182	165	195	43	200	165	52	60	86	52	65	43	26	30
Occupancy																	
under 2 years	4	25	75	75	25	175	75	25	100	300	175	100	150	25	75	125	250
2–3 years	9	33	100	122	44	144	100	33	144	233	166	122	155	22	44	88	222
4–5 years	9	66	133	133	77	133	122	33	155	177	144	144	144	22	44	100	155
6–7 years	9	77	155	111	100	111	133	33	177	111	133	122	133	22	44	77	111
8–10 years	12	116	166	141	108	133	141	66	166	91	116	116	141	41	66	75	91
11 years or more	48	135	77	89	127	81	81	164	60	50	70	89	66	170	150	77	56
Roof renovation	7	271	114	171	185	171	28	185	128	57	42	71	71	57	71	14	28
Security system present	19	257	94	221	178	184	42	200	152	63	52	68	52	57	52	21	26
Tenure																	
housing association/council	14	14	14	21	14	14	35	28	21	35	92	185	85	92	150	185	471
owner-occupied	73	131	124	126	128	127	109	126	128	100	104	87	101	108	91	50	16
rented	5	20	60	40	20	20	120	20	20	340	100	120	180	60	100	180	260
Type																	
bungalow	7	100	42	71	57	28	71	242	42	28	57	57	71	185	200	100	28
detached	15	200	213	200	166	93	146	133	106	66	93	40	73	73	53	53	26
flat	12	16	50	41	33	0	83	33	16	225	100	91	166	58	233	166	333
semidetached	35	114	94	94	102	182	88	111	111	68	97	120	80	120	68	71	77
terraced	25	76	92	96	112	68	96	72	148	132	124	140	132	92	80	76	80
Value																	
Up to £40 000	10	20	40	50	50	30	50	40	60	160	120	170	140	100	160	100	230
£40 000–£60 000	28	82	53	96	89	128	64	100	171	121	128	153	114	103	100	50	67
£60 000–£80 000	15	166	100	120	106	173	73	166	140	86	93	86	86	120	93	33	20
£80 000–£120 000	12	225	183	166	141	141	91	191	100	66	91	41	83	91	66	25	8
£120 000–£200 000	5	280	380	260	160	100	140	180	60	80	80	20	80	60	60	20	0
Financial holdings																	
Amex card	3	266	300	233	166	133	66	33	100	133	66	66	66	0	0	0	0
Children's savings	9	222	233	133	155	155	66	44	233	22	77	144	100	22	11	33	77
Company health insurance	9	233	222	200	177	177	133	33	166	155	100	111	55	11	11	22	22
Credit card	62	153	143	140	133	125	108	137	124	122	108	100	112	72	90	32	29
Credit card debt	5	100	80	120	80	160	40	60	200	140	160	180	120	40	40	20	60
Debit card	13	330	146	138	161	123	53	153	176	100	76	100	92	38	46	7	15
Equity plan	15	293	180	146	140	106	73	193	86	93	60	60	100	80	80	20	13
Funeral plan	3	200	66	66	66	33	33	200	33	33	33	166	100	166	266	66	133
Gold card	5	400	300	240	160	120	100	120	80	80	60	40	20	20	20	0	0
Investment trust	6	400	200	133	166	83	83	266	66	83	50	66	100	83	83	16	16
Life assurance	49	175	163	151	157	140	81	114	155	71	75	132	118	59	55	38	46
Lump sum investment	15	373	140	126	146	86	66	233	73	53	46	80	73	93	86	20	20
Mortgage	34	208	200	135	202	188	91	52	217	102	102	111	102	29	20	23	17
Pay credit card bill in full	7	214	142	200	100	185	42	85	214	171	142	157	128	28	57	14	28
Private health insurance	13	215	153	146	123	92	92	153	84	92	76	115	107	84	100	38	30
Private pension	31	238	206	141	219	119	93	80	206	103	80	122	87	35	25	25	29
Regular savings	15	360	186	153	200	106	80	100	133	113	60	93	93	33	33	20	20
Stocks or shares	24	295	175	133	158	91	95	187	104	87	62	79	75	87	70	25	16
Store card	24	254	141	162	116	145	58	137	141	129	100	87	120	50	70	25	33
Telephone banking	7	257	171	185	128	128	71	71	157	185	100	114	114	14	28	28	57
TESSA	17	223	147	123	129	100	94	211	70	76	64	47	100	117	117	35	17
Unit trust	9	344	188	122	144	77	77	255	66	66	44	44	88	100	77	22	0
Will	30	226	196	196	110	96	66	213	100	50	60	73	103	83	103	30	30

	GM	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P
Car details																	
1 car in household	49	73	55	73	87	42	71	151	140	138	138	132	157	144	61	57	51
2 cars in household	24	208	258	175	150	287	166	66	95	54	79	75	25	41	12	37	16
3 cars in household	5	200	100	300	280	120	180	60	80	120	80	80	20	20	0	40	20
Bought new	10	230	110	150	100	140	50	240	100	90	80	60	80	120	70	20	10
Bought used	26	150	96	130	84	180	57	130	176	126	130	142	103	88	26	26	53
Company car	7	214	314	242	214	228	157	14	157	142	100	71	42	0	0	14	14
Insurance (per annum)																	
under £300	41	170	85	97	70	175	48	170	168	100	109	141	109	92	41	26	39
over £300	7	228	171	228	142	142	71	85	128	157	157	100	100	57	28	28	28
Make																	
BMW	1	200	200	300	100	200	200	100	100	100	100	100	0	0	0	0	0
Citroen	2	200	300	150	150	150	150	150	100	100	100	50	100	50	0	0	0
Fiat	1	200	200	200	200	200	200	100	200	100	100	200	200	200	0	0	0
Ford	23	134	104	117	130	160	108	100	152	108	130	139	95	86	30	39	47
Honda	1	200	100	100	100	100	100	200	0	0	100	0	0	100	0	0	0
Nissan	5	160	120	120	120	120	80	180	80	80	80	100	100	100	40	40	20
Peugeot	4	175	200	150	150	175	125	100	150	150	100	75	125	50	25	25	0
Renault	3	166	233	166	166	166	133	100	133	133	100	100	100	66	33	0	0
Rover	12	175	150	133	141	150	108	133	108	91	108	125	91	108	41	41	33
Toyota	2	200	150	200	100	100	100	200	100	50	100	100	100	150	0	0	0
Vauxhall	14	142	135	135	128	171	114	100	164	100	114	121	71	78	21	35	28
Volkswagen	3	166	266	100	133	133	166	66	100	133	100	66	133	66	33	33	0
Volvo	2	250	300	150	150	150	150	150	150	50	100	100	100	100	50	0	0
Privately owned	68	135	125	119	122	129	108	127	129	107	116	117	110	104	39	35	30
Under 3 years old	9	244	144	188	122	177	77	155	122	111	111	66	66	66	22	11	11
Consumer goods owned																	
Camcorder	13	238	138	215	169	200	69	130	192	69	130	123	38	46	15	23	30
Fax machine	5	300	300	280	180	140	100	60	140	140	100	80	60	20	0	20	20
Hi-Fi	49	167	130	163	130	157	48	118	167	142	134	138	73	38	26	32	81
Home PC	28	217	232	235	128	117	96	57	196	121	96	100	89	28	17	39	53
Internet connection	2	300	300	300	100	150	50	0	150	250	100	50	100	0	0	50	50
Mobile phone	16	212	175	200	125	181	81	75	156	143	112	106	75	31	37	31	50
Satellite TV	24	150	70	254	108	141	75	79	166	95	133	166	58	50	25	50	95
Holidays																	
1 overseas holiday per annum	42	140	152	109	104	95	102	104	164	116	85	102	161	73	76	54	80
2 overseas holidays per annum	22	263	131	154	168	104	131	163	40	186	72	68	63	72	68	36	36
Activity holidays	9	266	255	122	188	100	122	77	111	222	55	77	155	22	33	22	55
Camp or caravan	37	159	162	86	167	156	94	64	175	110	91	129	102	51	35	43	100
Coach holidays	20	185	70	85	110	55	60	220	60	70	50	90	90	140	195	50	55
Cruises	8	250	100	112	125	62	87	175	62	87	62	87	75	87	112	37	37
Destination																	
Australia/ New Zealand	5	180	140	80	80	60	100	180	40	120	60	40	120	100	120	60	40
Canada	7	328	157	114	142	71	85	200	57	114	42	42	100	85	142	42	28
Caribbean	4	275	150	150	175	100	150	125	100	175	75	75	100	50	25	50	50
Europe	35	220	202	108	194	85	108	131	91	128	62	68	105	68	74	37	40
Mediterranean	43	179	144	151	155	95	151	141	102	158	58	79	81	67	65	37	44
rest of world	13	230	184	138	153	84	123	138	76	176	76	61	115	69	76	46	46
United Kingdom	72	131	125	109	122	112	102	119	119	106	93	113	108	97	97	41	84
USA	15	240	160	153	146	100	120	140	93	166	73	53	100	66	86	40	40
Lakes/mountain holidays	10	310	180	110	150	110	60	180	100	150	60	80	110	50	80	20	40
Motoring holidays	17	317	182	100	152	76	88	176	70	105	47	82	94	76	64	23	29
Safari holidays	2	250	150	150	100	100	50	100	100	200	50	50	100	0	50	0	50
Self-catered holidays	46	176	176	147	160	86	141	117	156	130	52	130	126	50	41	28	54
Skiing/snowboarding holidays	6	200	283	150	200	116	150	33	133	333	66	66	133	0	16	16	33
Take short breaks	51	172	147	125	145	105	131	129	117	125	45	113	117	66	74	35	50
Visit holiday camps	22	122	72	104	86	145	68	77	186	100	109	172	100	63	54	45	140
Visit theme parks	22	163	145	136	131	140	100	40	177	172	109	150	118	22	27	45	118
Interests																	
Antiques	19	200	168	105	110	78	73	147	68	94	78	126	126	89	115	47	63
Betting	8	175	62	100	62	87	62	100	125	125	125	225	50	100	50	62	150
Bingo	10	80	10	70	30	70	40	100	90	40	90	240	60	120	130	100	230
Church	25	152	220	88	104	72	72	144	72	52	64	100	96	116	216	64	60
Collecting	32	206	100	112	112	87	71	134	93	59	81	196	81	93	96	50	84
Competitions	31	193	74	90	90	70	54	116	116	83	87	222	93	77	96	48	129
Cooking	45	166	155	84	144	71	55	133	135	82	126	144	144	66	73	44	84

	GM	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P
Interests (continued)																	
Current affairs	22	250	209	104	122	77	77	145	86	113	68	95	104	81	104	31	45
DIY	39	182	141	74	143	141	51	148	158	64	141	161	71	76	30	30	66
Eating-out																	
frequently	14	171	107	135	107	92	107	150	50	128	71	92	92	92	107	50	64
occasionally	40	130	110	100	97	97	95	110	175	67	102	107	110	90	100	57	87
rarely	31	77	103	90	109	87	103	80	83	45	106	135	93	112	100	103	148
regularly	26	207	146	130	138	115	111	100	88	230	80	107	92	57	46	38	46
Gardening	56	151	128	112	121	114	57	148	116	42	76	126	116	130	87	48	51
Music																	
classical/opera	24	195	275	120	120	70	79	158	62	100	70	87	120	91	129	45	50
easy-listening	54	144	98	127	125	112	62	144	74	59	72	140	72	124	116	61	79
folk	28	153	92	110	96	89	67	132	75	60	82	189	85	107	100	67	96
jazz	14	221	207	114	128	85	85	135	78	128	78	114	114	78	64	42	78
light classical	40	180	167	145	145	67	60	182	55	62	67	92	82	107	175	47	42
pop/rock	54	131	124	138	133	142	124	31	162	157	137	151	122	18	14	48	142
National Lottery	64	128	60	118	62	128	43	117	134	117	129	132	100	90	90	53	126
Photography	24	262	162	112	141	95	70	145	104	87	83	133	87	79	62	33	54
Playing the pools	43	153	51	137	74	88	51	141	155	69	144	165	58	100	79	53	88
Pubs	38	152	94	78	92	144	92	63	165	210	97	163	84	50	34	39	150
Reading	63	141	136	117	122	95	85	117	104	100	92	119	115	92	112	42	93
Sewing	26	161	138	115	126	92	57	142	107	53	88	142	119	92	115	50	73
Theatre	27	229	266	122	129	77	81	122	70	185	62	74	125	62	107	33	40
TV																	
1 hour per day	8	125	275	87	150	100	125	50	75	175	75	37	175	50	62	75	62
2 hours per day	20	135	170	120	130	115	125	85	110	160	95	60	135	70	80	60	55
3 hours per day	25	124	104	124	116	124	108	112	124	112	112	104	100	88	88	68	76
4 hours per day	21	85	47	95	90	95	85	123	114	66	109	119	76	119	109	80	104
5 or more hours	19	57	21	73	47	68	57	110	73	31	94	152	57	142	131	121	189
Wine	29	248	234	117	134	96	89	127	120	124	89	110	89	62	44	24	48
Daily newspaper																	
Express	10	180	80	110	110	80	90	160	80	80	70	90	90	120	130	60	40
Financial Times	3	333	300	133	133	100	100	133	33	133	66	66	66	66	33	33	33
Guardian	5	160	320	80	120	60	120	60	80	240	80	40	220	40	60	40	80
Independent	5	200	320	100	120	60	120	60	60	220	80	60	180	40	40	40	40
Mail	21	176	119	133	119	100	90	142	95	119	90	95	109	100	123	57	52
Mirror	17	88	29	82	82	82	76	94	105	76	100	158	70	129	100	100	135
Star	3	100	33	66	100	100	66	66	133	66	133	300	33	100	66	100	233
Sun	27	74	22	77	70	96	77	66	118	74	114	233	59	92	74	96	233
Telegraph	13	223	176	130	138	76	100	169	61	115	69	46	84	92	107	38	23
Times	7	200	242	128	100	85	100	114	85	157	71	57	114	71	71	42	42
Alcoholic beverages																	
Beer																	
heavy drinker	5	180	80	100	100	120	60	100	120	120	80	160	40	80	40	40	100
light drinker	23	204	134	121	104	134	78	147	139	108	104	121	60	95	39	34	56
medium drinker	11	218	127	118	109	145	90	127	172	118	100	154	45	81	18	36	63
Brandy	10	270	140	130	140	80	90	190	90	70	80	130	70	100	70	30	50
Gin	9	333	244	122	166	77	111	188	77	100	55	88	77	88	66	11	22
Lager																	
heavy drinker	5	180	100	120	140	140	120	40	180	200	120	240	60	20	0	40	160
light drinker	29	186	117	120	100	131	79	93	151	134	113	137	96	51	44	37	117
medium drinker	13	169	115	138	123	161	92	53	192	184	123	169	61	30	7	30	130
Whiskey	27	211	133	114	111	96	81	174	85	81	81	107	62	122	74	37	51
Health																	
Arthritis	24	150	62	104	75	50	45	200	50	29	62	133	70	170	187	75	79
Asthma	24	141	125	129	112	116	83	83	137	120	112	145	87	66	58	50	112
Backaches	34	176	102	185	97	94	70	114	105	61	91	194	82	82	82	55	100
Diabetes	3	133	33	66	66	66	33	233	33	33	100	133	33	233	166	66	66
Earaches	13	161	76	92	69	61	53	207	61	30	61	130	53	184	153	76	69
Headaches	30	143	110	206	106	103	80	76	120	90	96	213	96	60	50	56	126
Industrial accident	9	177	100	133	77	144	33	100	155	122	111	188	100	55	77	55	155
Stomach problems	24	166	83	141	91	95	54	145	112	62	100	183	66	104	100	50	104
Stress	17	170	111	152	100	100	58	76	111	100	94	194	117	58	64	58	152
Smoke?																	
No	73	106	120	110	112	108	106	116	104	98	104	50	102	110	113	104	21
Yes	33	96	48	81	75	93	81	66	109	87	103	230	72	78	48	78	257

	GM	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P
Mail order purchases																	
Have made	78	125	119	120	120	94	98	106	115	96	94	123	114	85	94	67	93
Alcohol	5	480	300	140	160	40	80	140	80	100	40	120	80	40	60	20	20
Books	38	213	186	176	184	73	78	102	110	81	65	176	121	52	76	36	68
Bulbs	21	290	142	128	157	66	66	171	80	33	52	147	104	104	104	38	33
CDs/tapes	28	232	150	146	142	75	82	82	121	117	75	214	114	39	57	35	85
Fashion	48	154	131	150	150	75	81	108	143	85	70	156	147	60	87	52	85
Gifts	14	228	114	171	100	92	28	85	164	107	71	228	135	35	71	28	114
Supermarkets regularly visited																	
Asda	42	138	83	147	145	102	85	97	107	145	95	114	85	90	73	69	100
Co-op	19	147	73	89	94	78	68	147	78	52	73	136	78	131	126	89	115
J Sainsbury	44	159	163	152	97	90	88	143	88	152	79	79	143	81	120	52	45
Kwik Save	23	104	56	65	78	78	65	117	100	56	86	160	82	113	100	95	173
Marks and Spencer	19	215	136	136	121	73	89	173	73	115	63	89	115	100	131	47	47
Safeway	24	154	120	108	120	108	104	125	125	87	95	116	104	100	83	58	62
Sommerfield	17	182	105	76	117	64	88	152	70	82	64	141	94	105	117	64	88
Tesco	61	126	116	114	113	109	109	109	116	113	111	98	113	91	70	52	88
Waitrose	7	200	314	128	114	71	100	142	42	157	57	42	171	71	128	42	28
Other	42	140	83	76	73	90	59	150	95	73	80	157	90	92	111	71	159