

Key Findings from the 1991 SARs

Yaojun Li, Angela Dale, Ed Fieldhouse

30 October, 2001

Introduction

The SARs publications list is updated annually following the publications trawl. We have reviewed and report below some of the most important research findings from the 1991 SARs. These 'key findings' are very important in providing potential users with information about the research use of the data. They also demonstrate that the SARs are excellent value for money!

Research using the SARs has covered a wide range of areas in a number of different disciplines including sociology, human geography, economics and social statistics. Many users of the 2% Individual SAR have exploited the large sample size and relatively detailed geography to look at social differences between sub-populations (especially ethnic groups) and between geographical areas. Users of the 1% Household file have exploited the hierarchical nature of the data to look, in particular, at various aspects of household and family structure and inter-relationships within households. Users of both data sets have undertaken various types of multivariate analysis, taking advantage of the large amount of individual level information on the SARs in comparison to other census outputs.

Main areas of research

Ten main areas of research are covered here as follows:

- Ethnic differences
- Employment and labour markets
- Household and family composition
- Migration
- Health
- Methodological developments
- Use of the SARs in Northern Ireland, Scotland and Wales
- International comparative research
- Policy use of the SARs
- Marketing and commercial use of the SARs

Details are discussed below.

Ethnic differences

Research on ethnic differences has been undertaken using the SARs in a variety of areas including differences in unemployment, educational attainment, participation in full time education, class and occupational attainment, and economic activity. Most of the research

highlights the considerable disadvantages suffered by ethnic minority populations as compared with the majority white population.

- *Ethnic differences in unemployment* In contrast to the small sample size of survey data which forces many researchers to over-collapse the categories and call all minority ethnic groups 'black', the SARs allows for detailed analysis of ethnic groups. Differences both between and within major ethnic groups can be explored in various aspects of their socio-economic lives while at the same time controlling for other important characteristics. For instance, using the 2% SAR, Blackburn, Dale and Jarman (1997) show striking differences between ethnic groups in the vulnerability to unemployment, even among people with the same level of educational qualifications. One in five (20 per cent) of UK-born Black-African men and women with higher qualifications were unemployed, but the rate for similarly qualified UK-born Whites was only one fifth as many (3 to 4 per cent). This is a case of what might be called 'ethnic penalty' (See also Fieldhouse and Gould, 1998).
- *Ethnic differences in educational attainment* Equally significantly, the large sample size of the SARs allows the exploration of nativity effects whilst controlling for other important variables. Significant differences between ethnic groups who are otherwise similar to one another would then allow factors such as cultural differences or traditions to be explored. For example, Blackburn, Dale and Jarman (1997) show that, even among people born in the UK, Chinese, Other-Asian, and Black-African groups were highly qualified but the Black-Other, Black-Caribbean, Pakistani and Bangladeshi groups were far behind. Cheng (1996) suggests that, for some of the groups such as the Chinese, there might be a strong effect of traditional attachment to academic learning (See also Drew, Gray and Sporton, 1997).
- *Ethnic differences in participation rates* Similarly, one can use the SARs to make differentiations for detailed analysis of the subpopulations which is not possible even with special-purpose surveys like Youth Cohort Studies where ethnic groups are usually collapsed into a two or three way variable, especially when other variables are built into the multivariate models. In this regard, Drew, Gray and Sporton (1997) study the participation rates of the 16-19 year olds in full-time education using the SAR. The results show that Chinese, Other-Asian, Indian and Black-African groups are much more likely to be in full time education than Black-Caribbean, Black-Other and White groups. And this holds true even when the class position of the family head is taken into account.
- *Ethnic differences in class attainment* Ethnic differences in class attainment were pronounced. Heath and McMahon (1997) show that while around a third of British-born Whites were in professional and managerial (service-class) positions, only 1 in 10 of the first-generation Bangladeshi, 12% of Black-Caribbean and 14% of Pakistani groups were found in similar positions. Moreover, the educational qualifications do not translate into occupational positions equally well for all ethnic groups. Among the first-generation and for men and women respondents alike, each of the ethnic groups (except the Other-Asian men) were

significantly less likely than the Irish-born Whites to reach service-class positions rather than avoiding working-class positions; and among the second-generation and for both sexes alike, each of the ethnic minority groups were significantly less likely to be in service- rather than working class positions than British-born Whites except Black-Other group. And all this is shown even when educational qualifications and the age-groups of the respondents are taken into account.

Employment/labour markets

In addition to the research referred to above on ethnic differences in the labour market, there has been a great deal of work on the SARs relating to the labour market more widely. These include analysis of occupation, and of geographical variations and contextual influences on employment, unemployment and activity.

- *The impact of geographical context* The ability to identify geographical areas at a relatively detailed level allows 'place' to be included in the analysis by the use of multilevel modelling methods. Using the 2% SAR in combination with data from the Small Area Statistics, Fieldhouse and Gould (1998) shows that the type of local labour market does affect the probability of unemployment over and above the effect of individual characteristics such as age and educational qualifications. They are also able to show the impact of particular characteristics of an area. Black Caribbeans were especially hard hit in areas with high levels of unemployment whereas the Asian groups were much less affected (see also Gould and Fieldhouse, 1997). These analyses can be extended to include the role of neighbourhood context (see 'methodological developments' below).
- *The use of SARs linked with the LS for analysis of women's employment patterns* Holdsworth and Dale (1997) use the 1% SAR and the ONS Longitudinal Study (LS) to explore variations in patterns of employment and occupational attainment among women from different ethnic groups. They show that women from different ethnic groups have different employment patterns. The response to child-bearing is much stronger for white women than for most other groups - except Pakistani and Bangladeshi women. Generally, it is only white women who have high levels of part-time working. Logistic regression analysis shows that whilst women under 35 with no dependent children were unlikely to be economically inactive, the likelihood increases sharply over the other life-stages, especially for white women with unemployed partners or for unpartnered women with dependent children. Among the ethnic minority women, Pakistani-Bangladeshi women were generally found to have the highest odds of economic inactivity across the different life-stages. Educational qualifications also played a significant role. Even with life-stages controlled for, Black-Caribbean & Black Other, and Pakistani & Bangladeshi women were about 13 times more likely to be economically inactive than the reference group (the group least likely to be constrained from employment, namely, those who were under 35, UK-born, unpartnered and had degrees with no dependent children, of whom over 95% were economically active). White women were markedly more likely than the

other ethnic groups to work part-time rather than full-time if they had dependent children or if they were over the age of 35 or they were single-mothers. The results also showed that cultural norms were being modified by socio-economic factors. Thus, even though Pakistani-Bangladeshi women were, as a group, the most disadvantaged, 97% of the reference group were economically active. The authors then used the LS data to show that minority ethnic women had higher levels of full-time working over child-bearing, and one may expect that to be advantageous to their career progression. Detailed analysis shows, however, that this was not the case. Even though more women from minority ethnic groups who had children between these dates were working full-time at both time points, they did not benefit from working full-time. Ethnic minority groups were doubly disadvantaged because, as the authors say, 'they are more likely than White women to retain a full-time profile and be in manual jobs' (p. 453).

- *To use the SARs for studying integrated occupations* Sociological studies have long noticed the existence of occupational segregation, also called 'ghettoisation' occasionally, with men's jobs and women's jobs fairly clearly demarcated. For instance, most of senior managerial positions were occupied by men but most of junior secretarial jobs were done by women. Is there gender inequality within the same occupations where men and women incumbents have similar characteristics with regards to education, marital status, dependent children, or even social class? Survey data usually do not allow detailed analysis for specific occupations because of the sample n problems. To use the SAR can overcome this problem. For example, Hakim (1998, ch.9) used the 1% SAR to study pharmacy, which she termed 'an integrated occupation', namely, one 'with roughly equal numbers of men and women holding a pharmacy qualification and/or working as pharmacists' (p. 221). This was also an occupation 'that has been expanding and has escaped the constraints of recession' (ditto). The results show that male pharmacists were around 8 times as likely to work in small firms as female pharmacists (32% and 4% respectively), and that male pharmacists were twice as likely to be employers/managers as female pharmacists (13% and 6%), but the latter were around 10 times to work part time (30% and 3%). Although the two groups were little different in terms of educational qualifications, social class status, age, marital status, ethnicity, having dependent children, housing tenure, the male pharmacists spent more hours on paid work than the female pharmacists: only 4% of the former but 44% of the latter worked less than 36 hours per week. The patterns shown could thus put the thesis of occupational segregation to a serious test. Hakim says that 'Differences in work orientations lead to faster promotions for men and thus vertical job segregation *within* the profession, which again remains invisible in most occupational classifications' (p. 233). Hakim concludes that an 'insistence on paid work being fitted around familial responsibilities and a preference for convenience factors over high pay mean that women will generally be concentrated in the lower grades of professional and management occupations, even in the absence of sex discrimination, while men will continue to take the lion's share of higher grade jobs' (p. 234).

Household/family composition

The Household SAR in particular has been of great value in understanding household and family composition. The hierarchical structure of the 1% SAR allows inter-relationship to be made between individuals within the same household or family, thus facilitating the detailed analysis of family and household composition including the ethnic mix of family structure and inter-ethnic unions. Research is also available on differences in patterns of leaving the parental home and in household/family formation amongst black, white and Asian men and women, particularly with regard to the role of marriage and cohabitation.

- *Analysis of the ethnic mix of family structure* The hierarchical structure of the 1% SAR allows inter-relationship to be made between individuals within the same household or family, thus facilitating the detailed analysis of family and household composition (Holdsworth and Dale, 1997). For example, only a small minority (15%) of Black Other children have both parents of the same ethnic group whereas an overwhelming majority (over 96%) of White, Indian and Chinese children have both parents of the same ethnic group. Over two thirds of Black Other children (69%) have both parents of different ethnicity. Among Black children of mixed origin, nearly two thirds (65%) have a White mother and Black father, reflecting the highest number of inter-ethnic unions of this type.
- *Analysis of family formation, leaving home and cohabitation among ethnic groups* Using the 1% SAR for people aged 16 to 35, Heath and Dale (1994) found very significant differences in patterns of leaving the parental home and household/family formation amongst black, white and Asian men and women, particularly with regard to the role of marriage and cohabitation. For example, between ages 22 and 28, Asian men, and to a lesser degree Asian women, were much more likely to live with parents than white and black men and women. As compared with black women, Asian women were also much more likely to be married than cohabiting. Moreover, young Asian men and women with families of their own were also more likely to live with their parents or parents-in-law. However, among those who are either born in the UK or who have higher educational qualifications, white, Asian and black women are not much different, although this does not extend to cohabitation.
- *Analysis of inter-ethnic unions* Berrington (1996), based on the 1% SAR, shows that cohabiting unions are relatively common among young White, Black-Caribbean, Black-African and Black-Other men and women, but remain unusual among the South Asian populations. In terms of inter-ethnic unions, 'Black Other' men were more likely to have white (52%) than 'Black Other' (42%) women as partners [Table 7.9]; for both men and women alike and for most of the ethnic minority groups, the younger cohort (aged 16-34) were more likely than the older groups (aged 35-59) to have white partners, with the 'Black Other' men and women having the greatest tendency [Table 7.10]. Controlling for social class and educational qualifications, 'Black' men were the most likely to have a white partner and, for women, the 'Other-Asian' group were most likely to have a white partner.

Migration

The SARs, particularly the 2% SAR, have allowed research into the characteristics and mobility patterns of migrants, and how in-migrants differ from non-migrants. Related work has examined how migration relates to other social and economic outcomes such as economic activity and health. This has complemented work using the Special Migration Statistics.

- *Characteristics of migrants* Because of the large sample size and the individual level information on the SARs, including migrant status, the characteristics of migrants can be analysed. For example, this information has helped to understand differences in the movement of elderly people depending on their marital status and other characteristics (Al-Hamad et al, 1997, see also Al-Hamad et al, 1995).
- *Mobility* The data also allow analysis of inward migration in the last year which can provide evidence to the debate over the role of social housing in restricting residential mobility and employment opportunities (Boyle, 1995a). This is a question with important policy implications to which survey data can hardly answer. Boyle (1995a) uses the SAR to assess the impacts of tenure on long-distance migration compared with short-distance migration. Logits models are used to control for the effects of a wide range of socio-economic variables, allowing the independent impact of tenure to be identified. The results show that long-distance migrants are less likely to move into council housing than other tenures, but this does not support the assumption that the sale of council housing will necessarily increase labour mobility. Rather, increasing the supply of public housing, combined with more advertising of working class jobs more widely, may be more effective ways of increasing the migration propensities of the working classes.

Health

Other analyses of health have similarly used information on long term limiting illness, examining how health varied geographically and by different types of living arrangement amongst elderly people. Migration, and particularly transitions to institutional residence, have also been linked to long term illness amongst the elderly.

- *Geographic variations in morbidity* Gould and Jones (1996) uses the 2% SAR to explore individual and geographical variations in morbidity through a multilevel modelling. They modelled the probability of being long-term ill, controlling for age, housing tenure, ethnicity, car ownership and class. They find that (1) for the youngest cohort, there is no difference between whites and non-whites, but differences become increasingly marked for older groups, which cannot be accounted for by socio-economic characteristics; (2) variations in morbidity are strongly related to individual socio-economic characteristics; (3) there are marked

geographic differences in patterns of reported morbidity; and (4) the variations between SAR areas remain substantial even when socio-demographic variables at the individual level are taken into account.

- *Health of elderly people* Glaser et al (1997) examine the frequency of a limiting long-term illness among older people according to different types of living arrangements, including the population in institutions. Their results show that previous studies based only on the private household population have underestimated the prevalence of illness among older people. Long-term illness rates vary across family and household types, with higher frequencies found for those not living in families or in lone parent families, compared with those living as part of a couple. There is also a clustering of long-term illness among households. Those aged over 45 and suffering from long-term illness were more likely to live in households with other members also suffering from long-term illness. Thus health needs to be considered from a household, rather than just an individual, perspective; and institutional populations cannot be ignored. Sample surveys which are confined to private households may present misleading pictures. For instance, elderly people with poor health are more likely to be living in institutions than in their own home, particularly if they are single. Analysis of social surveys based on the non-institutional population therefore find the health of single elderly people better than elderly married people. However, if the institutional population is included this relationship is reversed (see also Grundy and Glaser, 1997; Murphy et al, 1997; Glaser and Grundy, 1998).

Methodological Developments

Much of the work on the SARs has exploited the novelty of the data (particularly the sample size and relatively detailed geography on the 2% SAR) to make methodological developments in a number of areas. The increasing use of multilevel modelling techniques by social scientists along with the availability of the SARs has led to developments in the understanding of the ecological fallacy and spatial variations in unemployment, deprivation and in health. The SARs have also contributed to improvements in population projections, in small area estimation and in micro-simulation of whole populations. Many other methodological developments have been made outside the academic arena.

- *The ecological fallacy.* Recent years have seen an increase in the analysis of deprivation in Britain. In most studies the unit of analysis has been geographical units such as local government wards or districts. This reflects, in part, a reliance on small area statistics and local base statistics from the censuses of population. Although useful in identifying specific problem areas, this type of approach may be subject to ecological fallacy which ‘arises when results from an analysis based on area-level aggregate statistics are incorrectly assumed to apply at the individual level’, and this happens in spite of ‘within-area homogeneity’ whereby individuals in the same area tend to have similar characteristics’ (Tranmer and Steel, 1998, 817). Because local areas are homogeneous, a correlation calculated from aggregated data is likely to be highly inflated representation of the true

correlation between individuals. Tranmer and Steel (1998) investigate four regions – three in England and one in Australia – and find that the ‘grouping variables’ most associated with homogeneity of enumeration districts are similar in each region: age structure, housing type and ethnicity. This shows that knowledge of the behaviour of the grouping variables allows an adjustment of how for the correlation between other variables based on aggregated data. This adjustment greatly reduces the aggregation bias. Thus a way is found for avoiding the ecological fallacy.

- *Using the SARs for assessing the relative significance of deprived place and deprived people* Areas of high levels of deprivation may be home to high proportions of particular social or demographic groups but it cannot be automatically assumed that these groups are themselves deprived. Although some studies have been based on purpose designed individual level survey data, these often lack sufficient sample sizes to effectively analyse small subgroups of the population or allow geographical disaggregation. Fieldhouse and Tye (1996) use the SAR data to investigate the social, demographic and geographical dimensions of deprivation. The distribution of individual level deprivation (deprived people) is compared with an equivalent area level index constructed from standard census output using conventional techniques.
- *Using area classifications in analysis of neighbourhood effects* A classification of census enumeration districts, has been added to both SARs. Through multilevel modelling techniques, Fieldhouse and Tranmer (forthcoming) use the area classification information on the SAR to investigate geographical differences in unemployment. Other research has indicated that where a person lives can effect their propensity to unemployment (see above). However, the understanding of these relationships are confounded by the reciprocal nature of the relationship between unemployment, housing and geographical location. Fieldhouse and Tranmer examine the relative importance of the individual, the type of neighbourhood of residence, and the local labour market in which one lives in explaining variations in unemployment risk. The paper concludes that most neighbourhood level variation in unemployment is due to housing market effects, particularly through neighbourhood selection.
- *To use the SARs for population projections* Unlike sample surveys which omit the institutional population, the SARs have data for people living in residential homes, hospitals, prisons or army quarters. This has proved very important for research. For example, Murphy and Wang (1996) use the SARs to make marital status population projections for England and Wales using a two-sex life table multi-state model called LIPRO. Whereas the conventional ‘atomistic’ approach tends not to take account of the wider socio-demographic context, the model used by the authors emphasises the role of multiple transitions in producing the observed numbers in a particular state. For instance, the model produces macro-level estimates of population parameters such as the annual numbers of births, and population (1981-2040) by sex, age and marital status, based on the application of transition rates. The model emphasises the interconnectedness of stocks and flows

across time and social space and produces easily comprehensible summary measures. The model provides a valuable sight into the processes at work and is the appropriate method for forecasting such systems.

- *Small area estimation using census microdata* The focus in the use of small area statistics is on the characteristics of a local area, usually in order to provide the information needed to plan services or meet customer demand. In Britain, standard census output provides tabulations for local areas using the 100% of the census records. Simpson (2000) shows that census microdata can help fill in the gaps in the small area tabulations, and estimate for local area information which cannot be obtained from the census. For example, the number of young people living alone in a local authority district is not available from the 100 per cent census tabulation. But an estimate of the correct value can be extracted from sample microdata. The SARs can be used for small area estimation in a variety of ways. The methods improve the reliability of the direct estimate from microdata by combining it with more reliable national data and with relevant data from the 100 per cent census tabulations. Characteristics not recorded by the census can be inferred for small areas by knowledge of their relationship with census variables. In addition, aggregate data, including non-census data, can be adjusted to avoid the ecological fallacy by using microdata to establish the basis for correlation between individuals or households. Finally, sub-samples of microdata can be used to simulate areas smaller than those to which the microdata are coded. The following are some of the methods discussed.

Supplementing a small sample The finest geographic definition of the SARs is 120k of population. For labour force forecasts in multi-ethnic areas, estimation of economic activity rates by age and sex for each ethnic group is not feasible via 100 per cent tabulation, and hardly so via the SARs. To improve the reliability of the estimates, Bradford Council (1996) supplemented the SAR data for Bradford with SAR data from larger regions containing Bradford – West Yorkshire, England and Wales – until the sample reached a minimum of 100 for a given cell.

Consistency with 100 per cent tabulations using iterative proportional fitting (IPF) Owing to the sampling nature and to the fact that no imputed records for households missed by the census are included, the SARs may give inconsistent data to the 100 per cent tabulation. The inconsistency is likely to be largest with smallest samples. To gain consistency and greater accuracy, the microdata results can be scaled to be consistent with the 100 per cent tabulation using the IPF method as Bradford Council did in developing labour force forecasts. (See below on SARs for policy use).

Micro-simulation of whole populations Micro-simulation attempts to assemble complete individual census records for the smallest areas – enumeration districts in England and Wales. For an enumeration district of n households, the method seeks a sub-sample of n records from the household microdata that best matches the 100 per cent and 10 per cent tabulations for the enumeration district. Williamson et al (1998) describe and compare algorithms for searching the SARs for sub-samples that match chosen local tabulations. They aim to reduce the discrepancy between the local tabulations and a random sub-sample of the

microdata. The best methods reduce the discrepancy by around 70 per cent, thus recreating a considerable amount of the diversity between enumeration districts – although not all of it. The authors validate the simulation by recreating tabulations that were not used in the simulation itself and find good results.

Use of the SARs in Northern Ireland, Scotland and Wales

Although many data sets are available with ‘boosted’ sample sizes for Wales, Scotland and Northern Ireland, detailed analysis is often not possible using such data. The SARs overcome these problems due to their large sizes and rich details in various aspects. Research has been conducted on the relative labour market disadvantage amongst Catholics in Northern Ireland, advantages for Welsh speakers in Wales, and the relationship between migration and class position, and class and health in Scotland.

- *Northern Ireland:*

‘Is there a penalty to being a Catholic in Northern Ireland?’ was a research question that Borooah (1999) addresses. Using the SAR for Northern Ireland and multinomial logistic regression, the author investigates the possibility of religious discrimination in occupational attainment and finds a religious penalty for both male and female Catholics as compared with Protestants in occupational attainment, although Catholic men faced much greater penalty than Catholic women. Both Catholic men and women were over-represented in the jobless and under-represented in professional and managerial positions, controlling for other personal characteristics. In other words, the Catholics were penalised for *who* they were holding constant *what* they were. Occupational attainment pertains to two aspects: different treatment of persons employed in the same occupation and different chances of access to an occupational status. Previous studies have shown that Catholics were about 20% worse off than average in structural disadvantages. How much penalty did they suffer in terms of identity disadvantage? Controlling for age, educational qualifications and area, 17% of Catholic men were jobless as opposed to only 9.2% of Protestants; for women, 10.2% of Catholics and 5.7% of Protestants were unemployed (Tables 8 and 9). For men with low qualifications, 18.9% of Catholics but only 10.4% of Protestants were jobless (10.6% and 5.9% for women: Tables 11 and 12). And having the various characteristics adjusted for, Catholic men were underrepresented in professional and managerial positions by 4 percentage points (32.4% for Protestants and 28.4% for Catholics). There was thus clear evidence against Catholics in Northern Ireland.

‘Why are Deprivation Levels in Northern Ireland Higher for Catholics than for Protestants?’ is a question that Borooah (2000) asks and endeavours to answer based on the Northern Ireland SAR. This question has two aspects: the deprivation level between Catholics and Protestants in Northern Ireland; and two, the penalty dealt to Catholics given their personal characteristics and circumstances. The basis upon which the deprivation index is created is: having no car, no central heating, no WC, no running water, no connection to public

sewage, non-permanent accommodation, more than one person per room, limiting long-term illness and no earner in household. [Note that this index is different from that by Dale et al, 2000: 215]. The incidence, for every DIC (deprivation-inducing condition, namely the variables named above), was greater for Catholics than it was for Protestants and this was true for both retired and non-retired parts of the sample (Table 1). The author then trichotomised the index: non-deprived (index = 0), mildly-deprived (> 0 but <= mean) and strongly-deprived (> mean), and used an ordered logit model to compare Catholics and Protestants while controlling for age, educational qualifications, gender, retiral status, employment status, number of people in household, single parenthood and residential area. And the results show that Catholics pay a higher price of deprivation than Protestants even with the same set of personal attributes/circumstances.

- *Scotland*

While Scottish Highlands and Islands experienced a century (mid-19th to mid-20th century) of depopulation, a re-population was witnessed in the recent decades. What are the main social-economic characteristics of the in-migrants? Boyle (1995b) studies the flows of migrants into the Scottish Highlands and Islands from the remainder of Britain using the SAR. The findings suggest that immigration from middle-class districts of southern England was a notable element of the population change in this remote rural destination. Boyle (1997), again using the SAR, tries to identify and contrast the characteristics of the English and Scottish residents in this destination area; he then compares the characteristics of the English incomers in this area with those in other areas of Scotland. The results show that the socio-economic characteristics of the English and Scottish residents in the Scottish Highlands and Islands differ considerably. The family circumstances of English and Scottish residents were distinct with English being more likely to be in 'non-traditional' households. English-born residents in the Highlands and Islands are also significantly different to the English living elsewhere in Scotland. The English in Scottish Highlands and Islands were more likely to be in service class or petty bourgeois occupations, to own their own homes, to be remarried or part of a family with cohabiting parents, in a word, to be 'urban refugees'. They were also more likely to be in service-class jobs in other parts of Scotland than the Scottish.

'Does migration exaggerate the relationship between deprivation and limiting long-term illness?' Boyle and co-authors (2001) ask in a recently published paper using the sub-sample of the SAR for Scotland. In the paper, they investigate whether the migration patterns of ill individuals influences the relationship between limiting long-term illness and material deprivation. More specifically, they seek to determine whether individuals who are well are more likely to migrate away from deprived areas and whether ill individuals are more likely to migrate towards deprived areas. If true, this would suggest that the apparent relationship between deprivation and limiting long-term illness is exaggerated by the effects of migration. They then examine the issue controlling for individual-

level characteristics expected to influence limiting long-term illness and pay special attention to the role of public housing in these relationships. From the 1991 Census Small Area Statistics, they calculated the mean deprivation scores (mean Carstairs scores) for each of the 13 groups in the ONSCLASS). They then model the probability of limiting long term illness (LLTI) on a variety of variables expected to affect LLTI. They find that migrants tend to be healthier than non-migrants; that once age is controlled for, short-distance migrants are more likely to report LLTI than long-distance and non-migrants; that controlling for other variables, both short- and long- distance migrants are less likely to suffer LLTI in all areas, except 'rural areas'; that controlling for individual characteristics did not alter the conclusion that migration does not appear to influence the relationship between health and deprivation in a consistent way; and finally, migrants in public housing are more likely to be ill than non-migrants.

- *Wales:*

The use of SARs permits fine-tuned analysis of issues of interest at regional or 'country' level. Some 'surprising' results may turn up. For instance, there was a rather high rate of unemployment at the time of the 1991 census and marginal groups suffered more disadvantages. Minority groups were generally more likely to be unemployed than the majority groups and this was not merely a matter of skin colour (Irish immigrants were as vulnerable to unemployment as other minority ethnic groups). But Welsh speakers (less than 20% of people in Wales speak Welsh) were found to be less likely than non-Welsh speakers to be unemployed in the Welsh labour market (Drinkwater and O'Leary, 1997). For those proficient in speaking, reading and writing Welsh, unemployment rates were 4.5 percentage points lower for males and 2.2 points lower for females than non-Welsh speakers. Why did the Welsh speakers enjoy the economic advantage? One reason may be that they enjoyed verbal and non-verbal, cultural advantages (social capital) as compared with the non-Welsh speakers in the job-seeking process; another reason might be due to the positive discrimination policy, especially for certain types of jobs in the public sector. These possibilities can, however, not be explored using the census data. But it is possible to compare Welsh and non-Welsh speakers in a number of aspects. The authors show that the two groups were similar in a range of profiles such as age structure, marital status, but Welsh speakers were on the whole better qualified. But even with these characteristics held constant, the ability to read, write and speak Welsh was found to have a significant effect on unemployment (Table 5). (See also Blackaby and Drinkwater, 1997.)

International comparative analysis

Many countries conduct a census and a growing number of them permit census microdata to be sent and used overseas, including the USA, Canada, Australia and Spain. This opens up the opportunity for cross-national comparative research and, even where microdata files are not allowed to go outside the country, comparative research can be

profitably pursued by collaboration between researchers in different countries. This section highlight some of the work conducted by resort of the microdata. As shown below, international comparisons of ethnic variations in social and economic outcomes yield highly interesting results using the SARs in conjunction with comparative data from USA and Canada.

- *The Chinese in GB, USA and Canada* Many contributions in Ethnicity in the 1991 Census (Karn, 1997) show that the Chinese and the Indians ‘do rather well’ in comparison with the Whites and the other ethnicity groups in a whole range of aspects, some of which are discussed above. Discussion of the Chinese at an international level is also fairly frequent (Cheng (1996; Model 1997). Model (2000) used the GB SARs, the US PUMS and the Canadian PUMFs to compare the Chinese socio-economic status by gender and nation, controlling for nativity. The results show that, first, Chinese occupational attainment varies not only be where Chinese settle but also by where they were born; second, cross-national differences in the propensity to hold ethnically typed jobs contribute to cross-national differences in Chinese occupational attainment; and third, controlling for nativity and job type, the few significant cross-national differences that arise tend to favour the USA.
- *Ethnic composition of families in Great Britain and the USA* Holdsworth (2000) uses the 1% SAR and the US PUMS to compare the extent of racial/ethnic family homogeneity in the two counties. The analysis shows that there are far more inter-ethnic unions in GB than inter-racial unions in New York. Just under one-third of married Black men and one-fifth of married Black women were married to a partner of a different ethnic group, which is particular the case for the Black-Other group. In contrast, under 10% of Black New Yorkers are married to someone from a different racial group. Comparing the race/ethnicity of children with that of their parents, the results show that a greater proportion of children in Britain live with parents of different ethnic groups, particularly among Black and ‘any other ethnic group’ children. Among all Black children, the largest number with parents from different ethnic groups live with a White mother and black father, reflecting the higher number of inter-ethnic unions of this type. Two thirds of Black-Other children have both parents from different ethnic groups. In contrast, a much smaller proportion of New York children live with mixed-race parents, even in the ‘other race’ group. On the whole, the research suggests a greater degree of ethnic heterogeneity in Britain than racial heterogeneity in the USA.

Policy uses of the SARs

There have been a number of specific policy related uses of the SARs in such areas as labour force forecasts, improving the precision of small area estimates, the use of SARs with survey data for synthetic estimation, and in household projections.

- *The use of SARs for greater precision in policy-making* The SARs can deliver more than the conventional approach by the use of techniques to avoid the 'ecological fallacy' and related problems. For example, the SARs can be used to predict the probability of long-term illness for each age of the 0-9 and 80-89 year olds, controlling for gender, class and tenure etc. which cannot be done via the SAS data (Gardiner, 1996). The results of this analysis show that people living in public housing were consistently more likely to have long-term illness than people in private sector housing; and that age and tenure were highly significant, but not gender.
- *The wide range of issues on which the use of SARs can provide enhanced value* In Manchester City, the SARs were used to show that, for single households in the city, tenure patterns vary significantly with age, with younger age groups concentrated in private rented property and the older ones in Council renting or owner occupations. The SARs can provide information for policy making which cannot be obtained from LBS/SAS. The Children's Services Division in the City Council wished to have information on the family type and economic circumstances for the 0-8 year olds, but what SAS can provide is the 5-year age bands where children aged 0-8 cannot be grouped. Another example concerned a comparison between cyclists and others. The SAR data showed that the distances to work travelled by cyclists were the same as by non-cyclists; that the cyclists were better educated, and that although half of them had at least one car at home, twice as many cyclists as non-cyclists did not have a car. Such data proved very good background information to the Sustainability Group (Butler, 1995).
- *The use of SARs for labour force forecasts* To conduct labour force forecasts in multi-ethnic areas, detailed economic activity rates by age and sex for each ethnic group are required, and the information can only be obtained from the SARs (Bradford City Council, 1996). The imprecision of estimates, due to the problem of small cell size, can be improved by using the univariate or bivariate distributions of key variables from the 100% Small Area Statistics (SAS). Simpson (1998) describes how the marginal numbers in each age category obtained from the sample data can be scaled to give numbers consistent with the numbers obtained from the 100% count available for age using Iterative Proportional Fitting (IPF).
- *The use of SARs with survey data for synthetic estimation* The fact that the SARs cover a wide range of socio-demographic variables enables the use of synthetic estimation whereby data from the SARs and from social surveys can be used in combination to obtain estimates at local authority level. Charlton (1998) uses regression coefficients derived from a General Practitioner (GP) morbidity survey to project the uptake of GP services at a local authority level. The predictions were applied to each individual in the 2% SAR and the probability was obtained of their having a serious illness in each of the SAR areas.
- *The use of SARs for household projections* Given the importance of household projections with regard to future housing requirements in England, King and Bolson (1998) argues that the SARs offer actual and potential opportunities to

add value to traditional projections. The availability of the SARs enables testing definitional sensitivity of projection outcomes, assisting further detailed disaggregation of projected components, assisting the matching of household projections to dwelling supply, and offering scope to explore via data linkage the relationships between household projections and 'backlog' housing needs, affordability, dwelling size, and tenure. The SARs can also assist in less conventional approaches to household projection including gross formation modelling, dynamic modelling, and microsimulation.

Marketing/commercial

The SARs can also be put to great use for marketing and commercial activities (Dale et al, 1995; Leventhal, 1994). For example:

- *Area profiling* In order to carry out a quota-based survey, it is first necessary to obtain a profile of the area to be used. With the SARs, one can produce independent or interlocking quotas which can be filtered on any other demographic area.
- *Demographic Profiling* Although ethnic minority groups make up only a small proportion in the whole population, the size of the SARs makes it possible to conduct a detailed analysis of the ethnic groups. For example, it is possible to find from the SAR the number of people born in Australia and New Zealand and resident in the GB, where they live, and what jobs they are doing. A similar analysis could be done for a particular ethnic group.
- *Penetration Rates* The SARs enables market researchers to set realistic field targets on certain penetration rates such as the proportions of households with young children in different geographic areas.
- *Weighting surveys* The SARs can be used as a benchmark to check the reliability of the survey results with regard to geographical spread and different demographic variables.
- *Market planning* If the target market for a product can be defined demographically, census microdata may be employed to obtain the market size and understand its characteristics and regional dispersion.
- *Retail site planning* Knowledge of the population within a district is essential for planning locations of retail developments. The SARs permit in-depth analysis of population structure and dynamics, and so could form a benchmark for site planning.
- *Other* Apart from the above, the SARs can be used for market share evaluation, customer segment evaluation, demographic modelling, etc.

Conclusion

In summary, the SARs have opened up whole new areas of research which would not have been possible with conventional Census or survey data. Not only have the SARs attracted a wide range of users from a wide range of disciplines and sectors, but these users have been highly prolific in generating high quality research published in leading academic journals.

Reference:

- Al-Hamad, A., R. Flowerdew, L. Hayes, R. Davies, and A. Geddes, eds. (1995) *Migration of Elderly People to Join Existing Households: Some Evidence from the Household SARs*. Lancaster: Lancaster University.
- Al-Hamad, A., L. Hayes, and R. Flowerdew (1997) 'Migration of the Elderly to Join Existing Households: Evidence from the Household SAR,' *Environment & Planning A* **29** (7): 1243-55.
- Berrington, A. (1996) 'Marriage patterns and inter-ethnic unions,' in *Demography of Ethnic Minority Populations*, OPCS ed., D. Coleman and J. Salt, eds. London: HMSO, 178-213.
- Blackaby, D. and Drinkwater, S. (1997) 'Welsh speakers and the labour market,' *Contemporary Wales* **9**: 158-70.
- Blackburn, R. M., A. Dale, and J. Jarman (1997) 'Ethnic Differences in Attainment in Education, Occupation and Lifestyle,' in *Employment, Education and Housing among Ethnic Minorities in Britain*, V. Karn, ed. London: HMSO.
- Borooah, V. (1999) 'Is there a penalty to being a Catholic in Northern Ireland? An Econometric Analysis of the Relationship between religious belief and occupational success,' *European Journal of Political Economy* **15** (2):163-92.
- Borooah, V. (2000) 'Targeting Social Need: Why are Deprivation Levels in Northern Ireland higher for Catholics than for Protestants?,' *Journal of Social Policy* **29**(2): 281-301.
- Boyle PJ (1995a) 'Public housing as a barrier to long-distance migration,' *International Journal of Population Geography* **1**: 147-64.
- Boyle, P. J. (1995b) 'Modelling population movement into the Scottish Highlands and Islands from the remainder of Britain, 1990-91,' *Scottish Geographical Magazine* **111** (1): 5-12.
- Boyle, PJ J. (1997) 'Contrasting English and Scottish residents in the Scottish highlands and islands,' *Scottish Geographical Magazine* **113** (2): 98-104.
- Boyle, P., Norman, P. and Rees, P. (2001) 'Does migration exaggerate the relationship between deprivation and limiting long-term illness? A Scottish analysis,' *Social Science & Medicine* **0**: 1-11.
- Bradford Council (1996) *Forecasts of the labour force: technical report*. Corporate Services, City Hall, Bradford BDI IHY.
- Butler, D. (1995) 'Use of the SARs by Manchester City Planning Department,' *SARs Newsletter* No. 5.

- Charlton J (1998) 'Use of Census sample of Anonymised Records (SARs) and survey data in combination to obtain estimates at local authority level.' *Environment and Planning Series A*, **30**: 775-784.
- Cheng, Y. (1996) 'The Profile of the Chinese in Great Britain,' in *Ethnicity in the 1991 Census*, Vol. 2, *The ethnic minority populations of Great Britain*, ONS ed., Vol. 2, C. Peach, ed. London: HMSO.
- Dale, A., Leventhal, B. and Moy, C. (1995) 'Applications of census microdata to research and marketing', *Journal of the Market Research Society*, **37** (2): 103-123.
- Dale, A., Fieldhouse, E. and Holdsworth, C. (2000) *Analysing Census Microdata*, London: Arnold.
- Drew, D., Gray, J. and Sporton, D. (1997) 'Ethnic differences in the educational participation of 16-19 year olds', in V. Karn (ed) *Ethnicity in the 1991 Census: Employment, education and housing among the ethnic minority populations of Britain*, London: HMSO.
- Drinkwater, S. and O. O'Leary (1997) 'Unemployment in Wales: Does language matter?,' *Regional Studies* **31** (6): 583-91.
- Fieldhouse, E. and R. Tye (1996) 'Deprived People or Deprived Places,' *Environment & Planning A* **28**: 237-59.
- Fieldhouse, E. and Gould, M. I. (1998) 'Ethnic minority unemployment and local labour market conditions in Great Britain,' *Environment & Planning A* **30** (5): 833-853.
- Fieldhouse, E. and Tranmer, M. (2001) 'Concentration effects, spatial mismatch or neighbourhood selection? Exploring labour market and neighbourhood variations in male unemployment risk using census microdata from Great Britain.' *Geographical Analysis* **33** (4): 353-369.
- Gardiner, C. (1996) 'Applications of the SARs in Local Government Policy Making,' *BURISA* **126**: 10-2.
- Glaser, K., Murphy, M. and Grundy, E. (1997) 'Limiting long-term illness and household structure among people aged 45 and over, Great Britain 1991,' *Ageing and Society* **17**: 3-19.
- Glaser, K. and Grundy, E. (1998) 'Migration and household change in the population aged 65 and over, 1971-1991,' *International Journal of Population Geography* **4**: 1-17.
- Gould, MI and Fieldhouse, E (1997) 'Using the 1991 Census SAR in a multilevel analysis of male unemployment,' *Environment and Planning A* **29**: 611-628.
- Gould, MI and Jones, K (1996) 'Analysing perceived limiting long-term illness using UK Census microdata,' *Social Science and Medicine* **42**: 857-869.
- Grundy, E. and Glaser, K. (1997) 'Trends in, and transitions to, institutional residence among older people in England and Wales, 1971 to 1991,' *Journal of epidemiology*

- and Community Health*, **51**(5): 531-540.
- Hahim, C. (1998), *Social Change and Innovation in the Labour Market: Evidence from the Census SARs on Occupational Segregation & Labour Mobility, Part-time Work & Students' Jobs, Homework & Self-Employment*. Oxford: Oxford University Press.
- Heath, A. and McMahon, D. (1997) 'Education and occupational attainments: the impact of ethnic origins,' in V. Karn (ed) *Ethnicity in the 1991 Census: Employment, education and housing among the ethnic minority populations of Britain*, London: HMSO, pp.91-113.
- Heath, S. and A. Dale (1994) 'Household & Family Formation in GB: The Ethnic Dimension,' *Population Trends* **77**: 5-13.
- Holdsworth, C. (2000) 'Ethnic composition of families in Great Britain and the USA,' in Dale et al. (2000), pp. 200-205.
- Holdsworth, C. and Dale, A. (1995) 'Ethnic Homogeneity and Family Formation: Evidence from the 1991 Household SAR', *CCSR Occasional Paper*, No 7. Manchester, CCSR.
- Holdsworth, C. and Dale, A. (1997) 'Ethnic group difference in women's employment,' *Work, Employment and Society*, **11** (3): 435-457.
- Karn, V. (1997) (ed.) *Ethnicity in the 1991 Census*, Vol. Four: *Employment, education and housing among the ethnic minority populations of Britain*, London: the Stationery Office.
- King, D. and D. Boldsdon (1998) 'Using the SARs to add policy value to household projections,' *Environment & Planning A* **30** (5): 867-80.
- Leventhal, B. (1994) 'Case Study Examples to Demonstrate the use of Samples of Anonymised Records in Marketing Analysis,' *CMU Occasional Paper* No. 5. Manchester: CMU, University of Manchester.
- Model, S. (1997) 'An occupational tale of two cities: minorities in London and New York,' *Demography* **34**: 539-550.
- Model, S. (2000) 'A cross-national look at Chinese immigrant occupations,' in Dale et al. (2000), pp. 195-200.
- Murphy, M. and D. Wang (1996) 'A dynamic multi-state projection model for making marital status population projections in England and Wales,' in *Exploiting National Survey & Census Data*, CCSR Occasional Paper 10, A. Dale, (ed.) Manchester: CCSR, University of Manchester.
- Murphy, M., Glaser, K. and Grundy, E. (1997) 'Marital status and long-term illness in Great Britain,' *Journal of Marriage and the Family*, **59**: 156-164.
- Simpson, S. (1988) (ed) *Making local population statistics, a guide for practitioners* LARIA.

- Simpson, S. (2000) 'Small area estimation using census microdata in the UK' in *Analyzing Census Microdata*, A. Dale, E. Fieldhouse and C. Holdsworth. London: Edward Arnold, pp: 217-221.
- Tranmer, M. and D. G. Steel (1998) 'Using census data to investigate the causes of the ecological fallacy,' *Environment & Planning A* **30** (5): 817-32.
- Williamson, P., M. Birkin, and P. Rees (1998) 'The estimation of population microdata by using data from small area statistics and samples of anonymised records,' *environment & Planning A* **30** (5): 785-816.