
James Harris, Danny Dorling, David Owen, Mike Coombes, and Tom Wilson

SYNOPSIS

This chapter reviews the principal lookup tables that have been assembled to aid the comparison of data between the different zonal systems used for different censuses and also between census zones and alternative geographical systems. The construction and structure of the tables are reviewed, and extensive resources, including many of the lookup tables themselves, are provided in the CDS Resources that accompany this book.

5.1 INTRODUCTION

This chapter discusses the use of lookup tables to re-aggregate census statistics to different spatial units. Lookup tables are poorly understood and their value to research is not widely appreciated. However, these data sets provide the crucial link between the fixed geography of the census and the spatial units for which many research applications require data.

It is helpful, at the outset, to explicitly state what is meant by a lookup table. A lookup table can be defined simply as a data structure that links one set of data to another. The focus of this chapter is specifically on geographical lookup tables and the 1991 Census geography. There are, of course, many other types of lookup tables in common use (e.g., telephone directories linking names and addresses to telephone numbers, cinema listings giving cinema names and films showing, and book content pages listing chapters and page numbers). There are many ways in which the two data sets in a lookup table can be arranged, though in a lookup table consisting of data sets A and B, a simple distinction can be made between one in which data set A is in some kind of order and one in which B is in order (e.g., 1981 Census areas listed in census Office area code order in terms of 1991 Census areas or 1991 Census areas in terms of 1981 Census areas).

The plan of this chapter is as follows. In Section 5.2 the official lookup tables produced by the Census Offices are discussed. The next section describes the creation of a lookup table linking the 1971, 1981, and 1991 Censuses on the basis of the 1981 geography,
followed by Section 5.4 discussing the production of a lookup table linking the 1981 and 1991 Censuses on the 1991 geography. Lookup tables linking census areas to functional regions form the basis of Section 5.6, while the next section examines the late 1990s local government reorganization in Great Britain and describes a lookup table that links the 1991 Census to the 1998 geography. The final section looks forward to the 2001 Census and includes some lookup table suggestions.

5.2 CENSUS OFFICES' LOOKUP TABLES

With the exception of the Samples of Anonymized Records (Chapters 14 and 15) and the Longitudinal Study (LS) (Chapter 16), all of the statistical outputs from the 1991 Census of population consist of areally aggregated data. That is to say, rather than providing information about individuals, the census data sets comprise counts of persons and households by area. Aside from the issue of confidentiality, it would be impractical to analyse the voluminous individual-level data, and therefore one of the objectives of aggregation is to achieve a level of statistical parsimony to reduce the large number of observations to a manageable set of descriptive statistics (Barr, 1993b). The smallest areas to which census data are aggregated are enumeration districts (EDs) in England and Wales (typically 200 households) and (Output Areas (OAs) in Scotland (50 households)). These areas form the building blocks of the census geography and nest into higher level administrative units, such as electoral wards and local authority (LA) districts. In 1991, this meant that the local authorities’ statutory obligation to report statistics at ward level was met without need for further manipulation of the data. In addition to the postal address, two geographical codes were recorded on the census returns: the ED/OA code and the full postcode. This enabled the data to be aggregated to postcode sectors, districts, and areas, allowing comparison with other postcode-based survey data. Over time, both the administrative and postal geographies have been subject to continual change in response to demographic, political, and operational factors. This constant redefinition of geographical units presents a significant hurdle to researchers who wish to use the 1991 Census statistics in conjunction with many other data resources, including previous censuses.

The ideal solution to this problem would be to create new aggregations from the individual-level census database to reflect each new alteration to the boundaries, be they postcode sectors, wards, or any other administrative or user-defined zones. Indeed, with the advent of Ordnance Survey’s ADDRESS-POINT product it is now technically feasible to aggregate the data to any arbitrary geography using the address as the atomic unit. Unfortunately this could potentially compromise the anonymity of the data, a risk that arises from the potential of identifying individual records from marginally different aggregations using a technique called differencing, an issue further addressed in Chapter 24. As a result, the Census Offices do not supply new aggregations and the only method of re-zoning the data is to re-aggregate the existing published outputs. Re-aggregation is the process of assigning data values from one set of zones to another using geographical intelligence about the intersection of the two geographies. This geographical intelligence may come from detailed boundary data or a grid reference for the centroid of each zone or it may simply be descriptive information about the composition of the zones.

In the context of the census, the objective of re-aggregation is to define units composed of census zones that approximate as closely as possible to the target areas for which we
require data. In the most straightforward scenario the target geography may consist of areas into which the census units nest perfectly, but in reality a degree of fuzziness will be introduced by the inevitable mismatch between the two geographies. Various methods can be applied to resolve the best approximate match between the two sets of areas, often involving an areal or population-based weighting factor. In most cases this will be achieved by overlaying the two geographies in a geographical information system (GIS). It would be highly inefficient, however, to repeat such a resource-intensive process each time we want to perform a new re-aggregation task. Instead, it is preferable to create a link file recording the allocation of census zones to the target zones. These files are variously called lookup tables (Cole, 1993; Australian Bureau of Statistics, 1996), constituencies (OPCS, 1993a), geographic equivalency tables, or correlation lists (Blodgett and Meij, 1997). They are produced by a number of organizations in the United Kingdom including the Census Offices, Ordnance Survey, and Royal Mail, as well as local authorities, research groups, and private sector organizations. Lookup tables are essential for census data re-aggregation as well as a range of other data manipulation and geographic analysis tasks, and yet they are poorly understood and inadequately documented.

The Economic and Social Research Council (ESRC)/JISC Census Programme has acquired a number of lookup tables that relate the 1991 Census geography to other spatial units, which, due in part to lack of suitable software for accessing and manipulating the data, are presently underexploited by academics. With the release of the 1991 area statistics, the Area Master File (AMF) was made available by Office of Population Censuses and Surveys (OPCS) and provided a ‘universal’ census geography-to-administrative geography lookup table for England and Wales (the Scottish equivalent was the Output Area/Higher Area Index (OHA) produced by General Register Office Scotland (GROS)). A sample of records from the top of the AMF is illustrated as Figure 5.1. These tables are available from the census site (http://census.ac.uk), and are linked from the enclosed CDS Resources CD. This made possible re-aggregation of the area-based data to a variety of spatial units including Regional and District Health Authorities, FHSAs, Urban Development Corporation zones, and Parliamentary and European constituencies. As each of these geographies has changed over time, the usefulness of the AMF has diminished, and it has largely been replaced by one-off lookup tables for individual geographies. An example of this is the new local government geography lookup table described later in this chapter.

Perhaps the most important lookup tables for census research are the postcode link files produced jointly by the Census Offices and Royal Mail: the Postcode/Enumeration District Directory (PCED) for England and Wales and the Scottish postcode index, again linked from CDS Resources and the census Web site. Annual releases of these files have enabled users to re-aggregate census counts to reflect the most up-to-date reorganization of postcodes, thereby allowing comparability with other survey data and the large resources of data collected by health authorities and local and central government for operational purposes. A welcome development in this area has been the All Fields Postcode Directory (AFTP) that relates unit postcodes to a range of administrative and statutory units and incorporates the PCED link to 1991 EDs for England and Wales. However, this file does not replace the AMF because it requires further processing to produce a usable link from the census areas directly to geographical units other than the postcode. Software has been developed by Yu and Simpson (2000) and Simpson (2001) and is accessible via http://convert.mimas.ac.uk/aftp/.
### Field | Variable
---|---
1 | Area type (1=county; 2=district; 3=ward; 4=ED)
2-3 | County code
4-5 | District code
6-7 | Ward code
8-9 | ED code
10-15 | ED code in 1981
16-25 | Centroid grid reference
26 | ED type (1=standard; 2=special; 3=shipping)
27-35 | Number of private households in ED
36-44 | Number of persons in ED
45-46 | New town
47 | Regional health authority
48-50 | District health authority
51-55 | Parish
56-57 | Central statistical area
58-59 | National park
60-64 | County electoral ward
65-67 | Parliamentary constituency
68-69 | Urban development corporation
70-71 | Housing action trust
72-73 | European parliamentary constituency
74-80 | Urban area

**Figure 5.1** The first few records of the AMF.
As each successive census adds to the corpus of digital demographic data, an increasingly important research activity involves comparing data from different censuses. Incompatibilities between population bases and cross-tabulations are only one of the problems that present themselves with change over time. The lack of a common geography between the 1981 and 1991 Censuses meant that comparative studies were all but impossible at a fine spatial resolution, and this is another area where lookup tables have made an important contribution to census research. The construction of 1981 to 1991 intercensal lookup tables is discussed next in this chapter.

5.3 THE 1991 CENSUS IN TERMS OF THE 1981 CENSUS GEOGRAPHY

This section describes the work undertaken to produce a lookup table of 1981 Census wards in terms of 1991 EDs and OAs covering Great Britain. It includes a description of what has been done, together with details of, and justification for, the methodology employed. Dorling and Atkins (1995) and Atkins et al. (1993) give full reports—including statistical and graphical analysis of the resulting data, pseudo-code of some of the algorithms used, and tables highlighting the more interesting spatial changes.

A number of factors, including the availability of boundary data for 1981, but not at the time for 1991, and the existence of a lookup table of the 1971 Census in terms of 1981 Census wards, led to the selection of 1981 Census wards as the consistent areal base to be adopted for this study.

It became apparent before work had progressed very far that the initial geographical referencing of the 1991 Census data was not as reliable as had been hoped, and consequently a large amount of manual checking was needed to facilitate the production of an allocation table linking each 1991 ED to its best fitting 1981 ward. Digitized boundaries did not become available for academic research until after the work had been completed. Four independent allocations of a 1981 ward were made for each 1991 ED. This was done despite knowing that a large number of wards had unchanging boundaries so that the method as a whole could be checked. The data sources used include the 1981 and 1991 ED level Small Area Statistics (SAS) files, the 1981 and 1991 AMFs, the 1981 digitized ward boundaries and the 1991 PCED file. The four allocations made were then combined into one robust lookup of a best-fitting 1981 ward for every 1991 ED. This process was largely automated, although the programs used highlighted a large number of cases in which further manual investigation was necessary. When the available data had been combined into a single lookup table, there still existed an unacceptable number of misallocations of EDs to wards. To remedy this, a number of additional checks were performed, cross-checking the data in the lookup file with various other sources, paper-based maps of census geography and documentation on boundary changes.

The last stage of the work reported here was the production of 1991 SAS files aggregated to the 1981 ward level. The SAS files produced, along with the final version of the lookup file, have been deposited at the ESRC Data Archive at Essex University and at the Census Dissemination Unit at Manchester so that they can be of benefit to other researchers. The lookup table is also available as part of CDS Resources, which provides a linked set of 1971–1981–1991 SAS counts aggregated to common 1981 ward boundaries together with a program for the user to extract data selections.
Attention should be drawn to the following possible sources of error in the data sets:

- The process of aggregating blurred counts has the potential to introduce significant errors (especially where the total counts are low, which is likely to be the case when performing analysis at ward level). This does not affect the population totals used in this study, as they were never blurred.
- Where 1991 EDs are split between 1981 wards, misallocation is inherent to the process of assigning them solely to one or other of the wards they span. The misallocations will 'cancel out' locally, however, and so should have almost no effect on, for instance, district level analysis of ward-based statistics.
- Census data for residents in special EDs that contain very small communal establishments are suppressed in the ED level SAS. Consequently these people cannot be allocated to a 1981 ward using the methods outline here. This affected approximately 0.06% of the 1991 population. The influence of this effect should have been insignificant for this study, and at higher geographical levels these people are included in census output.
- Where the census variables for an ED were imported or exported by OPCS to other 1991 EDs (to maintain confidentiality) that were across 1981 ward boundaries, there will be local misallocation of the population. Again, this does not affect the population counts used in this study but would be important in particular areas for studies based on statistics other than population counts.
- Visitors who did not provide an accurate address of their usual residence were not included in the 1991 ED SAS, and hence could not be allocated. This affects approximately 0.4% of the population, and would have been a similar problem in 1981 and 1971 so is unlikely to bias the results shown here.
- All the errors associated with the undercount of 2.2% of the population in 1991 remain. This is by far the most significant source of error and eclipses the aforementioned problems. The remedy used for this study was the choice of a changing population definition that behaved similarly to the national level changes.

This work resulted in the production of a lookup file of 1991 EDs to 1981 wards, covering all of England, Wales, and Scotland, and the creation of the full set of 1991 10 and 100% SAS tables aggregated to 1981 ward level. The lookup file allocates a best fitting 1981 ward to each 1991 ED, and the grid reference of the ED centroid is given along with an indication of the source from which the 1981 ward code was derived. The 10 and 100% SAS files are in a similar format to the OPCS ward level SAS tables.

Following the 1981 Census, a great deal of work was done to link together the 1971 and 1981 Censuses, in terms of both spatial equivalence and comparability of the variables. For a criticism of the census tract method of linking censuses see McKee (1989). The 1971 census data were aggregated to 1981 ward level at Newcastle University (Dorling 1991). As a consequence of this, the vast majority of change-over-time analysis at Newcastle, covering the 1971 to 1981 period, makes use of the 1981 ward level (or aggregates of 1981 wards) as its areal base. Because 1981 is also the mid-point in the three census periods from 1971 to 1991, the 1981 areal units are a sensible base for studies covering this period. The alternative approach, aggregation of 1981 EDs to 1991 wards, would not facilitate comparisons to the 1971 Census (unless a similar exhausting task was undertaken with the 1971 Census).
The development of the lookup file was a three-stage process, following which the transformed SAS files were derived from the refined version of the lookup file. The first stage made use of OPCS ED level data, in particular, the ED centroids (for both 1981 and 1991 EDs), together with the digitized 1981 ward boundaries. These data were used to produce three alternative lookup files from 1991 EDs to 1981 wards. The grid references of the 1991 EDs and 1981 ward boundary data were then incorporated within a GIS database. With this visual aid a large number of EDs, suspected to be in the wrong positions, were checked interactively using on-screen maps (see Atkins et al., 1993). The second stage made use of the PCED file, which contains details of individual unit postcode centroids and population counts and the 1991 EDs within which they fall, for all of England and Wales. Data from this file were used to produce population weighted ED centroids, which were allocated to 1981 wards by the point-in-polygon algorithm. This produced a further independent lookup file of 1991 EDs to 1981 wards. This was not necessary in Scotland where the OA naming scheme contains an intrinsic link to 1981 Census postcode geography. In the third stage, the lookup files created in the previous two stages were combined to produce what became, after further scrutiny and revision, the final lookup file. This was achieved by choosing the 1981 ward indicated by one or other of the lookup files, the choice of which to use being determined by the matching between the 1981 wards suggested by the various lookup files for each 1991 ED.

Numerous checks were then carried out on this penultimate lookup file and corrections made accordingly. Ensuring one to one matching of wards in districts that have not had ward boundary changes between the censuses, checking the distances between EDs within each ward, and drawing ‘ward polygon’ diagrams (see Dorling and Atkins, 1995: Figure 19) were among the checks carried out. When all the checks had been carried out, the lookup table was used to produce 1991 SAS data files aggregated to the 1981 ward level, from 1991 ED level SAS data.

5.4 THE 1981 CENSUS IN TERMS OF THE 1991 CENSUS GEOGRAPHY

One of the most serious omissions of census planning in England and Wales is the failure to provide a definitive link between the geographical units used by successive censuses, in sharp contrast to Scotland and Northern Ireland. After the 1981 Census, an attempt to address this problem was made by the definition of a set of census ‘tracts’ forming spatial common denominators between 1971 and 1981 EDs. The availability of these spatial units plus OPCS/GROS definitions of variables in terms of the 1971 and 1981 Censuses greatly facilitated the analysis of change over the decade 1971 to 1981 (Champion et al., 1987). Despite the dramatic improvement in computer power and data-handling techniques, the ability to analyse change over the decade 1981 to 1991 was made harder by the simultaneous changes in census geography, census base and variables. However, the one advance was in the availability of digitized boundaries for census areas.

The work of Atkins et al. (1993) in using a variety of techniques to identify the way in which 1991 EDs were related to 1981 Census wards, provided a resource for spatially linking data from the 1971, 1981, and 1991 Censuses in the form of a lookup table (which they made available to the academic community on the Manchester Information and Associated Services (MIMAS) computer service). This data set has been used to create two files that describe the spatial relationship between the wards used in 1981 and 1991. There is a large amount of socioeconomic data for which the ward is the smallest

Spatial unit (e.g., monthly unemployment data and annual vital statistics data). Some of these data sets are collected using 'current' wards (i.e., the electoral wards that exist at the time the data are collected), whereas others use 'frozen' wards, usually those defined by the most recent census. Through the creation of estimates for one set of ward definitions by converting data collected for a different set of wards, it becomes possible to create longer spatial time-series of data or estimate rates using denominators collected for a different set of wards than the numerator.

Two converters were created: one for translating data collected on a 1991 ward base to a 1981 ward base; and the other to convert data collected for 1981 wards to a 1991 ward base. An example application of the former would be the creation of a 1991 economic activity data set for 1981 wards, enabling ward-level unemployment rates to be calculated from JUVOS unemployment data. The latter converter was used in a project to estimate ward-level population change by ethnic group, converting 1981 Census data on country of birth to a 1991 ward basis, in order to enable comparison with 1991 Census data on population by ethnic group (Owen and Ratcliffe, 1996). This form of conversion is useful where 1991 data are richer and more complex than 1981 data, and thus the possible loss of information incurred through the approximation involved in the conversion is less great for 1981 Census data.

Each converter takes the form of a set of 'factors' that distribute the population resident in a ward as defined by one census to the wards defined by another census that cover the same territory. The method relies on having population data for smaller spatial units that can be located within the each set of wards, in this case 1991 EDs with 1981 ward codes.

The first step is to read the lookup file and create two linked lists, one for 1981 wards and one for 1991 wards (identified from the first six characters of the 1991 ED code), which identify for each 1991 ED the sequence number of the next ED within the same ward. The sequence number of the first ED in the ward is also noted. During this first pass, 1991 population totals for each 1981 and each 1991 Census ward are calculated.

In the second step, the program works though each ward in turn, identifying all combinations of wards from the two censuses, and summing the 1991 population for each combination. In the third step, the program works through each ward in turn, converting population sums into proportions of the 1991 population of the ward. The same basic methodology works in both directions, yielding factors that split 1981 wards into 1991 wards and vice versa. The 1981 to 1991 lookup table can be found in CDS Resources.

5.5 LINKING THE 1991 CENSUS TO FUNCTIONAL REGIONS

This section of the chapter outlines the reasons—and the methods—that led to the creation of lookup tables linking 1991 Census data to:

- the CURDS functional regionalization (Coombes et al., 1995), and
- the localities and city regions (Coombes et al., 1996).

These two sets of areas illustrate how lookup tables are essential to sets of boundaries that have been created purely for research purposes.

Analyzing data from the census almost always provokes the question of which areas to use for the analysis (Openshaw, 1984). This question arises because census data are collected, assembled, modified, and disseminated within a hierarchy of areas (Coombes,
1995). The fact that there is a choice of areas for which data will be available allows researchers to decide which set of areas is in principle the most relevant to the research questions being examined. If the research is not explicitly dealing with local politics or administration, then there is no reason to assume that local authority areas—which figure so largely among the census published data sets—are indeed the appropriate areas to study. For example, the local authority area of Nottingham includes only a part of that urban area, whereas the nearby city of Sheffield includes not only the urban core but also some of the surrounding rural areas. As a result, data on the two cities’ local authority areas cannot be reliably compared in studies of the many social and economic factors—such as ethnicity and unemployment—that tend to vary between any city’s inner areas and its more suburban and rural surroundings (Coombes, 1997). As recognized in many countries (e.g., Dahmann and Fitzsimmons, 1995), the need is to group together each city with those nearby rural areas that are closely linked with it: when this grouping is undertaken on a consistent and meaningful basis, the sets of boundaries that are created provide meaningful and comparable definitions for all cities. Areas defined in this way will then be much more appropriate for many research purposes than are local authority areas that are the standard areas for most published census outputs.

No doubt, one reason for lookup tables to be considered ‘essential or highly desirable’ by 64% of the researchers surveyed by Rees (1998a) was that without them analysis of change over time is scarcely possible. The reason is that each census is tied to the sets of local authority areas that are in existence at the time. Thus, a lookup table becomes essential to provide a link between the areas used in one census and those used for another. Researchers who choose to use non-standard areas will in fact require a separate lookup table to group data to their areas of interest from each census and its specific set of building block areas. One consequence here is that a slightly different pattern of errors results each time building blocks from different censuses are ‘best fitted’ to non-standard areas. In practice, these errors in lookup tables are only likely to be significant for areas with small total populations—or to be more precise, sets of areas with a low number of building block areas per ‘output’ area (Coombes 1995)—but this is not the case for either the functional regions or the localities that are of interest here. In addition, both these sets of areas tend to group together adjacent urban and rural areas, so their boundaries rarely run through the urban fringe areas that are the most dynamic in terms of new development and which, as a result, are also where building block areas are most likely to be changed from one census to the next. In short, both the sets of non-standard areas of particular interest here are largely immune to the two main sources of error to which lookup tables can be prone.

Researchers wishing to analyse change through time need to choose which time-point to use as the basis for the definitions of the units within which the change will be traced. For example, a study of the growth of new towns will almost certainly need to use boundaries that are defined in terms of the built-up areas at the latest time-point so that the data can show the new towns ‘growing into’ their current state. In contrast, an analysis of counter-urbanization might well use boundaries that reflect the limits of the built-up areas at the start of the period so as to monitor the growth taking place in those areas that were initially rural. The two sets of boundaries examined here provide for both perspectives:

- the functional regions reflect British towns and cities as they were in 1971, and
- the localities’ definitions are primarily based on 1991 Census data.
With their emphasis on grouping cities with their hinterlands, both these sets of boundary definitions aim to identify areas that are 'functional' in the sense that most patterns of flows and interactions tend not to cross their boundaries. The earlier—functional region—definitions relied exclusively on commuting patterns for evidence on flow patterns, with the method of definition an advance upon those that had evolved over several decades in the United States (Dahmann and Fitzsimmons, 1995). By contrast, the locality definitions were far more ambitious because they drew upon migration and other flow data sets in addition to commuting; an entirely new form of analysis had then to be devised to collate the evidence on patterns of interaction that are present in a variety of very different sources of information (Coombes and Openshaw, 2001).

In terms of lookup tables, the Functional Regions, which were defined specifically to support 1971 to 1981 Census change-over-time analysis, benefited from the innovation in 1981 of widely accessible computerized ward-level data. The original boundary definitions had been derived from the 1930 areas in the 1971 Census commuting data set. The remaining process was a straightforward but time-consuming inspection of maps of the 9267 1981 wards to create a lookup table to the Functional Region Zone boundaries. The subsequent allocation, of the 1991 Census wards to the 1981-based boundaries, used the Atkins et al. (1993) lookup table of 1991 EDs to 1981 wards, following the computerized approach described earlier in this chapter. Figure 5.2 shows how the 1991 wards were allocated to Functional Region Zones according to the assignment of the majority of their 1991 population, as indicated by the allocation of their constituent EDs (which was derived via the Atkins et al. link to 1981 wards).

In the case of the localities that were defined more recently, lookup tables were inherent to the process of definition itself. The main building blocks for the definitions were to be 1991 wards, but the aim was to draw upon a very wide range of information, which in turn led to relying upon source material based on many different building block areas.

![Diagram](image)

**Figure 5.2** Deriving a lookup table for functional region zones.
Some of the information was in the form of boundaries that had been linked to 1981 wards by Owen et al. (1986), so these could be linked to 1991 wards through the same method as was used for Functional Region Zones (Figure 5.2). In other cases, a point-in-polygon operation was carried out to identify into which of the areas of interest each 1991 ward’s centroid fell. Coombes and Openshaw (2001) describe how the innovative next step was to create ‘synthetic data’, which, in effect, answer the question ‘within how many of the sets of areas examined was this pair of wards grouped together within the same area?’ The matrix of these values was then input to the Coombes et al. (1986) regionalization algorithm and the localities were defined accordingly. Thus, the building blocks for defining localities were 1991 wards, but the definitions were based on ‘synthetic data’ whose very creation depended fundamentally upon lookup tables. The functional regions lookup table, localities and city regions lookup table, list of the localities and city regions, and list of functional regions are all provided in the accompanying CDS Resources.

5.6 THE 1991 CENSUS IN TERMS OF THE NEW LOCAL GOVERNMENT GEOGRAPHY

The local government structure and geography of the United Kingdom changed substantially between the time of the 1991 Census and April 1998. Many two-tier local authorities were replaced by a single tier, and a substantial number of boundaries were revised. In order for researchers to obtain 1991 Census data, and other data based on the census geography, for the 1995 set of local authorities, and to facilitate a straightforward comparison with the results of the 2001 Census (assuming no more boundary changes occur), a set of lookup tables were constructed. This section begins with a brief description of the need to create the lookup tables—the introduction of a new geography—and then moves on to discuss the lookup tables themselves, and the SASPAC91 programs and data sets based on them that are available to download from the Web.

The change in geography resulted primarily from the wholesale reorganization of local government that was implemented over the period April 1995 to April 1998. A very brief description of the changes follows. More details can be found in Chisholm (1995), Jackson and Lewis (1996), Johnston and Pattie (1996a, 1996b), Leach (1994), and Wilson and Rees (1998, 1999). Maps showing the local government geography of Great Britain in both 1991 and 1998 are available as part of CDS Resources (Figures 5.3, 5.4, and 5.5). The idea of reorganization originated with the Conservative government of the early 1990s. It believed that there was a strong argument for replacing the two-tier structure of local government that existed outside the English metropolitan areas with single-tier authorities. This, it argued, would reduce the duplication of services between upper and lower tier councils, and would eliminate the public’s confusion as to which council was responsible for which particular service, thus increasing accountability. In Scotland and Wales, proposals were drawn up; little public consultation took place; minor amendments were made, and the altered proposals became law in the Local Government etc. (Scotland) Act 1994 and the Local Government (Wales) Act 1994. The new local government organization of Scotland and Wales came into effect on 1st April 1996. In Scotland, the nine regions and 53 districts and island areas were replaced by 32 single-tier council areas; in Wales the eight counties and 53 districts were replaced by 22 unitary authorities. In England, however, matters were not so simple. The government established a Local Government Commission to review local authority boundaries in England and discuss options with
Figure 5.3  Local government authorities, 1998: England.
councils, local people, and other interest groups. Not surprisingly, conflicts of interest emerged between different parties, and the review process took longer than originally envisaged. The eventual outcome was considerably different to the all-unitary authority option the government had hoped for. Eighteen two-tier counties were left unchanged; a further 16 underwent changes in their boundaries, but not their structure, four were abolished, and one (the Isle of Wight) underwent a change to its structure (but obviously not its boundaries!). Unitary authorities were created only for the main towns and cities. Thus, pre-reorganization non-metropolitan England consisted of 39 counties, and 296
districts within those counties, while the new reorganized local government geography comprised 46 unitary authorities and 34 counties, these counties being made up of 239 districts. The new organization for England was introduced in four phases on 1st April 1995, 1996, 1997, and 1998. Local government boundaries were also reviewed in Northern Ireland in the early 1990s. Some small changes were made in the Local Government (Boundaries) Order (Northern Ireland) 1992.
However, the 1998 local authority geography is not only the result of this reorganization. A significant number of small boundary changes took place in the period between the 1991 Census and reorganization. These minor boundary changes may seem rather insignificant compared to the wholesale redrawing of boundaries, but it is still important to account for them. This is especially the case when local level time-series population data are being analysed: changes in population size that appear to be fairly large may be due less to demographic change and more to boundary change. For example, the London borough of Brent’s mid-1991 population on 1991 boundaries was 248,600 while on 1998 boundaries it was 242,300 thousand. The estimated population of the borough in mid-1997, on the new boundaries, was 249,600. The real increase in population between 1991 and 1997 was 3.0%, but if the change was incorrectly calculated using the different boundaries of 1991 and 1998 then the increase would appear to be 0.4%.

Perhaps the most useful lookup table linking the 1991 Census with the new local government geography is one that describes the 1998 local authorities in terms of 1991 Census areas (rather than the other way round). Such a lookup table has been compiled using information obtained from Office for National Statistics (ONS) for England and Wales, the General Register Office for Scotland, and the Northern Ireland Statistics and Research Agency, and is available in CDS Resources. This lookup table has also been embedded in a number of simple SASPAC91 programs that can be used to extract user-defined 1991 Census data for the new geography, and are included in CDS Resources. Details of how to use SASPAC91 can be found in the SASPAC User Manual (London Research Centre, 1992), and on the MIMAS Web site, http://www.mimas.ac.uk/. These SASPAC91 programs comprise two parts: the first part reads in the original 1991 Census system files for the 1991 geography and then creates a new set of system files for the 1998 local authorities; the second part then reads in the new system files as would any ordinary SASPAC91 program and extracts whatever census data has been requested by the user. The programs are initially set up to extract just the total resident population. One set of these programs extracts the SAS, the other set extracts the local base statistics (LBS).

The lookup tables mentioned in this section have been used to produce a variety of data, and two data sets are available to users of CDS Resources. The first is a re-worked version of the set of tables included in the 1991 Census monitors and at the beginning of the published reports, produced using the lookup tables in the SASPAC91 programs for the LBS. The second data set is a file of mid-1991 population estimates for the 1998 local authorities (Great Britain only). As described in detail in Wilson and Rees (1998) this was produced using the ED and OA mid-1991 estimates produced by the Estimating with Confidence project. Details of this project are given in Simpson et al. (1995), and Simpson et al. (1997a), and the new geography population estimates are also available in CDS Resources. Further, a number of the lookup tables described here have been used in the creation of a new Web interface for linking censuses through time (LCT), which provides access to 1971, 1981, and 1991 data referenced to 1981 wards and re-aggregations to many other geographies. This project has also included extensive recalculation of 1991 area statistics to ‘correct’ for the 1991 undercount, discussed in Chapter 12.

5.7 CONCLUSION: LOOKUP TABLES FOR THE 2001 CENSUS AND BEYOND

Plans for the 2001 Census outputs include a number of important developments for lookup tables, including the availability of such tables under common licensing arrangements, and
the production of a new table identifying the relationship between unit postcodes and 2001 OAs (see Chapter 3). Cole (1998) has suggested a number of further features these lookup tables should possess, and progress is being made towards these objectives:

- they should link 2001 Census areas to previous Censuses;
- they should be updated as boundary changes occur;
- lookup tables should cover the whole United Kingdom; and
- lookup tables should be integrated with digitized boundary data.

**Box 5.1** Key web links for Chapter 5.

| http://census.ac.uk/cdu/let/ | Interface to the Linking Censuses through Time data sets
| http://census.ac.uk/cdu/Datasets/Lookup_tables/Area_Master_File.htm | Area Master File. A list of EDs and the larger areas in which they nest (England and Wales)
| http://www.census.ac.uk/cdu/Datasets/Lookup_tables/Output_Area_Higher_Area_Index.htm | Output Area Higher Area Index. A list of OAs and the larger areas in which they nest (Scotland)
| http://census.ac.uk/cdu/Datasets/Lookup_tables/Postal/Central_Postcode_Directory.htm | Postcode-Enumeration District directories (annual). A list of unit postcodes and the EDs in which they nest (England and Wales)
| http://census.ac.uk/cdu/Datasets/Lookup_tables/Postal/Scottish_Postcode_Index.htm | Scottish postcode indexes (annual). A list of unit postcodes and the OAs in which they nest (Scotland)

**Box 5.2** Additional CDS resources for Chapter 5.

| Census monitor data | Includes boundary maps and re-aggregated data published in census monitors
| Functional and city regions lookup table | A list of 1991 wards and the localities in which they nest