

## SOCIO-ECONOMIC INDICATORS

### Introduction

The ONS Longitudinal Study is a 1% sample of the population of England and Wales. It contains linked census and life events data, beginning with a sample from the 1971 Census.

Social inequalities is a popular topic area for survey data: many studies based on the LS have focussed explicitly on socio-economic variations in demographic events. The LS has, for example, been the basis of much influential research on inequalities in mortality, and in health status, and has been used to analyse socio-economic variations in fertility and in migration as well. The LS is also a unique source of data for measuring social differentials in cancer incidence.

After completing this guidance, you will be able to:

- Choose appropriate socio-economic indicators from the LS for your research question
- Combine individual, household and area level indicators to their best effect
- Formulate a request for such data

The socio-economic indicators that are most commonly used to study health inequalities are:

- Occupation
- Income or wealth
- Education

The LS has very good information on occupation from census records. Social class is derived from these occupations and employment status. It also contains occupational data from certain event registrations. The large sample size of the LS means that information on occupation and industry can also be used to proxy occupational exposures (e.g. noise, toxic substances, dangerous equipment) in order to investigate their effect on health and mortality<sup>1</sup>.

Although there is no information on income in the LS, various indicators of household resources can act as proxy measures, such as housing tenure, access to a car and other household amenities. However, it is important to note that when using household-level socio-economic indicators, a household measure might not apply to all of the individuals within that household. In addition, household level information is not available for those LS members living in an institution, and changes in the distribution of certain household indicators over time means that their socio-economic meaning at one time may not apply at a later time point.

The LS also contains information on educational qualifications but for the 1971, 1981 and 1991 Censuses only qualifications obtained after the age of 18 were recorded on the census form, so it is not possible to differentiate among those with lower-level qualifications. In contrast, the 2001 and

---

<sup>1</sup> See: Goldblatt, P. and Fox, J. Mortality of men by occupation. In: Goldblatt, P. (ed.) *Mortality and social organisation*. Series LS no. 6: HMSO 1990

2011 Census education question included qualifications from GCSE/O-level onwards, therefore for those LS members who were also present at previous censuses, it is possible to use these data retrospectively.

We will now consider the occupation-based social class measures used in the LS. Since occupation based socio-economic indicators can only be derived for people who are in paid employment, we will introduce a range of techniques that can be used to classify those LS members who are not in employment (33% in 1991), e.g. children, the retired, the long-term sick, the unemployed and people who provide unwaged care work in the household (i.e. not volunteering). In addition to this problem with occupation based social class measures, the questions on occupation in the census have changed over time, meaning that more information is available in some years than in others. Furthermore, occupational classification schemes have changed over time, which affects the social class that LS members are allocated to.

Following the sections examining occupational based indicators, we will also explore the range of alternative social-economic indicators that are available in the LS. These alternative social indicators are at the individual and ecological (area) levels. Examples of alternative individual level social indicators include household resources (tenure, access to a car and other household resources) and educational level. Ecological level (area-based) socio-economic indicators include the relative deprivation of the area in which an LS member lives. These ecological level socio-economic variables are based on the aggregate circumstances of the individuals or households in that area.

In the final section of this guidance, we pose some example research questions which show how the socio-economic information in the LS can be used in practice.

## Occupation in the census

It is common to use occupation to group individuals with conceptually similar socio-economic circumstances for the purposes of analysis. It is used as the basis of *social class* measures.

Occupational data in the LS comes from the occupation questions in the census forms from 1971 to 2011 (see [census forms](#) on the CALLS-HUB), but there are some important differences between them. Four main themes are covered by the census questions:

- Economic activity: whether in employment (full- or part-time), seeking work, retired, studying, etc.
- Employment status: for those in employment, whether an employee, self-employed (and not employing others), or self-employed (and employing others).
- Type of business.
- Actual occupation.

Although these four themes are addressed in all the censuses, there is some variation within them.

### *Economic activity*

Economic activity tells us whether an individual is in paid employment, and if not, whether they are actively seeking employment.

Census questions on economic activity are only completed by those of working age. In 1971, the lower age limit for those in employment was 15, but in 1981 this increased to 16 in line with the increase in school leaving age.

The categories for the economic activity question have changed with each census. All the censuses distinguish between individuals who were employed in the previous week, those that were seeking work, and those who were retired. The other categories have changed, as follows:

- The 'housewife' or 'homemaker' and 'permanently sick' categories were introduced in 1981 and have remained since then. In 1971, there was space in question B7 for the form-filler to write in other categories such as these.
- Students have always been identified; if not in the question on economic activity, then elsewhere, in a separate question.
- The 'seeking work' category has also been present at all censuses from 1971-2011, but in 1991 and 2011 this group included those who were prevented from seeking work by temporary sickness, whereas these groups were identified separately in 1971 and 1981.
- People 'waiting to take up a job' were identified separately from 1981 onwards.

The distinction between part-time and full-time work is only included in 1981 and 1991, and only for the job that the person was doing last week. In 1971, 2001 and 2011, there is a question on hours worked which allows this to be determined, but this was only for a job that the person was doing last

week. In 1981 and 1991, 'part-time' meant 30 hours or less per week but there was no question on the specific number of hours worked per week.

Economic activity is available in the LS as the variables ECONP7, ECONACT8 and ECONPO9 (search for these variables using CeLSIUS [Data Dictionary](#), for information on how they are coded). Under the ECONPO9 coding scheme, students who were also in paid employment were coded as economically active, whereas for ECONP7 and ECONACT8 they would have been coded inactive. To help with continuity, an additional variable has been derived for economic activity in 1991 (ECONPO89), which codes *all* students in 1991 as economically inactive. The variables for 2001 (ECOP0 and ECOP80) and 2011 (ECOP11 and ECOP81) identify students separately regardless of whether they were economically active. This is consistent with the 1981 (ECONACT8) and 1991 (ECONPO89) variables (see the CeLSIUS [Data Dictionary](#)).

### *Employment status*

*Employment status* tells us whether an individual is an employee, self-employed or self-employed and employing others. At another level of detail, it also identifies apprentices and trainees among the employees, and can indicate whether an employee has responsibility for supervising others.

In 1971, 1981, 2001 and 2011 this information was given either for the job that the person was doing in last week, or, if they were retired or out of work, for their most recent job. However, in 1991 this information was only included if the person was working last week.

Each census has a question that distinguishes between employees, the self-employed who employ others, and the self-employed who do not employ others. In 1981, 2001 and 2011, those employees who have a role in supervising other people were also identified in questions 13 (1981 Census), 29 (2001 Census) and 36 (2011 Census) (see the [census forms](#) on the CALLS-HUB).

In 1971, apprentices and employed trainees were identified separately in question B18. In 1981 and 1991, they were included with the 'employed' categories, except trainees on government schemes in 1991 who were separately identified in question 13. In 2001 and 2011 all apprentices and trainees were again included in the 'employed' category (question 18 and question 26 respectively).

For the 1971, 1981 and 1991 Censuses, employment status is given in the LS by the variables STATC7 (1971), TEMST8 (1981) and EMPST9 (1991) (search the CeLSIUS [Data Dictionary](#) for more information). Each of these variables identifies categories of employee, apprentice/trainee, manager, foreman, self-employed employing others and self-employed not employing others. For the most part, these variables are coded based on the form-filler's description of their occupation (see next section). So, categories such as 'apprentice', 'manager' and 'foreman' exist for TEMST8 even though there are no such categories explicitly specified on the 1981 Census form. Extra information is also used where available: in 1971, question B18 is used to improve coding of apprentices, and in 1981 the information

on supervisory status in question 13 can be used to help code managers and foremen. For the 2001 and 2011 Censuses, the employment type variables EMTP0 and EMTP11 respectively only contain the categories of employee, self-employed with employees and self-employed without employees. There is separate variable identifying supervisor or foreman status (SUMP0 in 2001, and SUMP11 in 2011).

### *Type of business and occupation*

On each census form, there are questions that ask for:

- Name and type of business at which each household member individual works, and
- Name and description of their occupation at that business.

[1971 Census: questions B15 and B16; 1981 Census: questions 11 and 12; 1991 Census: questions 15 and 16; 2001 Census: questions 27, 28, 30 and 32; 2011 Census: questions 34, 35, 37 and 38 (see the census forms on the CALLS HUB)].

If an individual is not working at the time of the census (or in the previous week), then they are asked to record their most recent occupation (in 1991 though, individuals who are not working are asked to record their most recent job thinking of the last 10 years).

*NB: This means that the LS contains occupational data for many more people than are actually in paid employment at the date of the census.*

Occupation is coded in the LS according to the occupational classification that was current at the time of each census. Therefore, the 1971 occupation data are coded to the 1970 OPCS Classification of Occupations (CO70)<sup>2</sup>, in 1981 they are coded to CO80<sup>3</sup>, and in 1991 the 1990 Standard Occupational Classification (SOC90)<sup>4</sup> was used. For 2001, the SOC2000<sup>5</sup> scheme is used, but occupations have also been back-coded to SOC90 for the purposes of comparison. For 2011, the SOC2010<sup>6</sup> scheme was used but was back-coded to neither SOC90 nor SOC2000 (the ONS website has information on the relationship between SOC2000 and SOC2010).

The information about the type of business that the individual works in is also available in the LS. For example, at each census there is a variable giving the industry sector worked in IND7, IND8, IND9, INDP0 and INDP11. Industry is coded to the *Standard Industrial Classification* (SIC) that was current

---

<sup>2</sup> Classification of Occupations 1970, London: Her Majesty's Stationery Office  
Hattersley, Lin and Rosemary Creeser (1995) 'Appendix IV 1971 Census: definitions and concepts (reproduced from Census 1971, England and Wales, General Report part 1, Definitions)' in Longitudinal Study 1971-1991 History, organisation and quality of data, pp 256-259. London: OPCS

<sup>3</sup> Classification of Occupations 1980, London: Her Majesty's Stationery Office  
Hattersley, Lin and Rosemary Creeser (1995) Longitudinal Study 1971-1991 History, organisation and quality of data. London: OPCS 'Appendix V 1981 Census: definitions and concepts (reproduced from Census 1981, Definitions, Great Britain)' pp 295-296, 39, 62-63

<sup>4</sup> Standard Occupational Classifications, Volumes 1 and 2. OPCS

<sup>5</sup> Office for National Statistics (2000) [Standard Occupational classification 2000](#) Volumes 1 and 2. London: The Stationery Office.

<sup>6</sup> [SOC 2010](#)

in that census year: so IND7 is coded to SIC68; IND8 and IND9 are coded to SIC80; INDP0 is coded to SIC92; and INDP11 is coded to SIC2007<sup>7</sup>.

In addition, there are a few extra questions on occupation that have only appeared on individual censuses:

- In 1971, the occupation one year ago was recorded, but only for those who were employed in the previous week. This is given by the variable OCC17.
- The 2001 and 2011 censuses ask whether the person has been available for work in the last 2 weeks (AVAILW0 and AVAILW11), whether they have ever worked (EVERWK0 and EVERWK11), and if so, when (LWYP0 and LWYP11).
- The 2001 Census also asked for the size of the organisation that an individual works for (or last worked for), in terms of number of employees (SIZP0).

---

<sup>7</sup> See the ONS' [UK Standard Industrial Classification of Economic Activities](#) website for further information on the previous and current versions of the SIC (current and previous versions)

## Occupation information from events

Information about occupation can also be obtained from three types of events registration records in the LS:

- Cancer registration.
- Birth registration<sup>8</sup>.
- Death registration.

(see the Events guidance for information on these events).

### Cancer registration (CANC dataset)

This records the occupation and industry of a patient (or their husband, if the patient is a married woman, or their father if the patient is a child). From 1993 onwards, this information was voluntary and so data coverage may be poor. The key LS variables for the occupation of a cancer patient are OCCC (for 1971-1980), OCC8CC (for 1981-1990) and SOC90PCC (for 1991 onwards).

Note that, if the LS member were present at a previous census, we could use their occupation as recorded at that census (or, if appropriate, the occupation of their father or husband, using the non-members' file); these are likely to be more complete and reliable since they are derived from census records.

### Birth registration (LBSM or LBSF datasets for live births to sample mothers and fathers respectively)

Records the occupation of the *father* only until 1986; subsequently the occupation of both parents was collected. However, there is still no obligation to record the mother's occupation, and so these data are less complete. Non-recording of occupation is a particular problem amongst young women in low status jobs.

If we are looking at the occupation of parents of a new birth into the LS (NBIR dataset), we could instead look at the non-members' file and use the census-recorded parents' occupation, which is likely to be better completed and more reliable. Using these data, we could do analyses by parents, by spouse, or, in 1971, by the chief economic supporter<sup>9</sup> for the household. This person is identified by the derived variable CESIND7 in table ME71 (1971 Census - LS members), or NCESIND7 in table NM71 (1971 Census - non-LS members).

---

<sup>8</sup> However, up to 1986 occupation was only collected for the 'working parent', usually the father, and recording mother's occupation was not obligatory. Although all cases are coded from 1986, it is thought that many mothers in regular employment are coded as having no occupation, which affects the reliability of the social class coding for mothers in the NBIR dataset.

<sup>9</sup> In the 1971 Census, a household member aged 15 or over who is the head of household, or is related to the head, and who is the highest ranked in the household according to the following criteria, in this order: 1) economic activity (those in full-time employment ranked highest, followed by those out of employment, part-time, retired, and others); 2) family position (married ranked highest, followed by widowed or divorced in a family, and others); 3) sex (male ranked above female); and 4) age (older ranked above younger).

### Death registration (DETH dataset)

Records the occupation of the deceased. For 1971-1981, if the deceased was a married or widowed female then her husband's occupation was recorded. For 1971-1979, if the deceased was under-15 then the occupation of a working parent was recorded. For 1979-1982, if the deceased was *under-16* then the occupation of a working parent was recorded. From 1983, the mother's occupation is used if the deceased is under 16, and the father's occupation is also recorded.

Occupation from death records can be problematic because the information is necessarily second-hand. Informants may be inclined to elevate the status of the deceased's last occupation, or report their most prestigious lifetime occupation rather than their most recent one. Alternatively, the informant may not understand exactly what the deceased's job was. Individuals from institutions are a particular problem because the informant is often not a close relative of the deceased and may not know them very well.

Again, if the LS member dies, we could instead go back to the last census at which they were present and use their self-reported occupation, which is likely to be more accurate than the occupation reported by the informant upon the LS member's death.

## Social classifications based on occupation

Social classifications based on occupation are used by:

- Central government to analyse social and health variation, and thus direct policy and resource allocation.
- The private sector in market analysis.
- Academics in scientific analyses in health and demographic research.

Until 2001, the standard classification of social class was the *Social Class based on Occupation* (SC for short). Since 2001, however, this has been superseded by the *National Statistics Socio-economic Classification* (NS-SEC).

The following sections examine the two classification systems, see how they compare with each other, and how they are manifested in the LS. Alternative socio-economic classifications based on occupation that are available to LS users will also be examined.

### *Social Class (SC)*

The *Registrar-General's Social Classes* (RGSC) were introduced in 1913 and were renamed in 1990 as *Social Class based on Occupation*. The classes are:

- |       |                                       |
|-------|---------------------------------------|
| I     | Professional occupations.             |
| II    | Managerial and technical occupations. |
| IIINM | Skilled non-manual occupations.       |
| IIIM  | Skilled manual occupations.           |
| IV    | Partly skilled occupations.           |
| V     | Unskilled occupations.                |

Individuals are assigned to the classes first by being allocated to an occupational group according to the kind of work they do, then each occupational group is assigned as a whole to one of the Social Classes. Finally, individuals are reassigned to different Social Classes if their employment status suggests that this is necessary (e.g., foremen in Social Class IV may be moved to Class IIIM).

Note that a few individuals will change social class over time because of changes in the classification of occupations. In 1971, social classes were assigned based on the 1970 OPCS Classification of Occupations (CO70), in 1981 the 1980 Classification was used (CO80), in 1991 it was the 1990 Standard Occupational Classification (SOC90), in 2001 this was replaced by SOC2000 and in 2011 by SOC2010, which is the latest classification.

### *Social Class variables in the LS*

Social Class based on an individual's occupation as recorded in the census is represented in the LS by the variables SOCC7, SOC8 and SCLAS9 representing the years 1971, 1981 and 1991 respectively. For 2001, there are two: SCLAS0 (coded to the Occupation Classification scheme

SOC90) and ROSERG0 (coded to SOC2000). For 2011 there is one variable, RGSC11, which is coded to SOC2010.

Looking at the Social Class variables SOCC7 and SOC8 in the Data Dictionary, it can be seen that the coding is not consistent (SOC8 has more categories than SOCC7). To ensure continuity between 1971 and 1981 when using Social Class, you can use SOCS7 instead of SOCC7 since it uses the 'statistics' coding similarly to SOC8.

There are also Social Class variables based on any of the events where occupation is recorded, such as cancer registrations (e.g. OCSC9CC in the CANC dataset), deaths (e.g. OCSCDFDE in the DETH dataset), etc. However, the occupational data obtained from events is often poor quality, and lacks consistency over time. It is therefore usually preferable to derive Social Class from the person's occupation at the last census.

### *NS-SEC*

From 2001 onwards, the Social Class based on Occupation is replaced by a new classification called the National Statistics Socio-economic Classification, or NS-SEC.

Information on NS-SEC, and its history and derivation is available from ONS' [The National Statistics Socio-economic classification \(NS-SEC\)](#) website.

NS-SEC distinguishes individuals based on their 'employment relations', ranging from people who 'provide services in return for compensation (such as salary, job security, promotion prospects)' to those who 'provide discrete amounts of labour in return for a wage representing the amount or time of work done'. It moves away from the concept of 'skill' and the manual/non-manual divide that characterised the old Social Class scheme, since many manual occupations command higher salaries and status than non-manual ones.

NS-SEC has eight main categories (known as the 'analytical' scale), and a further 17 minor categories (the 'operational' scale) for more detailed work. Individuals are coded to NS-SEC by first allocating them to occupational categories using the SOC2000 (or SOC2010) classification. Then the occupational category, the size of the establishment worked at, and employment status (employer, self-employed or employee) are cross-referenced on a lookup table to get the NS-SEC category. It is important to note that NSSEC is **not** an ordinal (ordered) classification, so it makes no sense to talk about moving 'up' and 'down' the classification.

Analytic scale	Operational scale	Description
1.1	L1	Employers in large organisations.
	L2	Higher managerial occupations.
1.2	L3	Higher professional occupations.
2	L4	Lower professional and higher technical occupations.
	L5	Lower managerial occupations.
	L6	Higher supervisory occupations.
3	L7	Intermediate occupations.
4	L8	Employers in small organisations.
	L9	Own account workers.
5	L10	Lower supervisory occupations.
	L11	Lower technical occupations.
6	L12	Semi-routine occupations.
7	L13	Routine occupations.
8	L14	Never worked and long-term unemployed.
N/C	L15	Full-time students.
	L16	Occupation not stated or inadequately described.
	L17	Not classifiable for other reasons.

Note that the last three operational categories are in a separate 'not classified' group (N/C) in the analytical scheme.

Looking at continuity between 1991 and 2001, individuals in 2001 have been coded using both SOC90 (SOC900) and SOC2000 (OCCP0), and the NSSEC has been derived using both the SOC90 and SOC2000, which allows comparisons between 1991 and 2001. The NS-SEC categories can also be mapped back onto Social Class:

Social Class	NS-SEC Operational Categories
I: Professional occupations	3.1, 3.3
II: Managerial and Technical occupations	1, 2, 3.2, 3.4, 4.1, 4.3, 5, 7.3, 8.1, 8.2, 9.2
IIIN: Skilled occupations – non-manual	4.2, 4.4, 6, 7.1, 7.2, 12.1, 12.6
IIIM: Skilled occupations – manual	7.4, 9.1, 10, 11.1, 12.3, 13.3
IV: Partly skilled occupations	11.2, 12.2, 12.4, 12.5, 12.7, 13.1, 13.2, 13.5
V: Unskilled occupations	13.4

The Institute for Social and Economic Research, University of Essex, has derived an estimation of Social Class, [NSSEC](#). For 2001, it has been derived using SOC2000 (ROSERG0), and for 2011 it has been derived using SOC2010 (RGSC11).

#### *Other occupational classifications*

The LS contains a number of other socio-economic measures that are based on occupation and employment status

### Socio-economic group (SEG)

*Socio-economic group*, or SEG, is a classification that aims to bring together people with similar social and economic status.

For many sociologists, SEG is seen as a better measure than Social Class for social scientific purposes. It is a measure of employment status rather than 'skill' or 'social standing'. It comprises 17 groups, three of which are subdivided, as follows:

1.1	Employers in industry, commerce etc. - large establishments.
1.2	Managers in central and local government, industry, commerce etc. - large establishments.
2.1	Employers in industry, commerce, etc. - small establishments.
2.2	Managers in industry, commerce, etc. - small establishments.
3	Professional workers - self-employed.
4	Professional workers - employees.
5.1	Ancillary workers and artists.
5.2	Foremen and supervisors non-manual.
6	Junior non-manual workers.
7	Personal service workers.
8	Foremen and supervisors - manual.
9	Skilled manual workers.
10	Semi-skilled manual workers.
11	Unskilled manual workers.
12	Own account workers (other than professional).
13	Farmers - employers and managers.
14	Farmers - own account.
15	Agricultural workers.
16	Member of armed forces.
17	Inadequately described and not stated occupations.

SEG is derived from occupational group, employment status and size of establishment. Occupations are coded to groups based on whichever classification scheme is in use at that time. The SEG for an individual is calculated using a lookup table, which tabulates occupational groups against a combined variable for employment status and size of establishment.

It is important to note that SEG is **not** an ordered ranking, so it makes no sense to talk about moving 'up' and 'down' the classification.

In the LS, SEGs have been coded at each census (except 2001) based on the information on occupation, employment status and establishment size recorded on the census form. The variables are SEG7 for 1971, SEG8 for 1981, SEG909 for 1991 and SEG11 for 2011.

### Hope-Goldthorpe scheme

The Hope-Goldthorpe scale has 36 categories ranked in order of 'social desirability' of male occupations. The categories are assumed to provide a high degree of differentiation in terms

of both occupational function and employment status. It is important to highlight that this class schema was devised for men, but the scores are commonly used for both men and women. At the time of writing the variable is only available in the LS for 1971 Census data coded to CO70. Where LS data for either occupation or employment status was not available the value of 0 has been allocated. The variable is HGSCORE from the ME71 file.

### Goldthorpe classes

Goldthorpe Classes are produced by combining the 36 categories of the Hope-Goldthorpe scale to create a class schema whose members broadly share similar market and work situations. Individuals in the same Goldthorpe Class are theoretically comparable in terms of income, conditions of employment, economic security, prospects of advancement and in the amount of control they exert over the processes they are involved in. Again, this system was devised for men, but is widely used for both genders.

In 1971, the classes were assigned using CO70, and there were 7 classes, ranging from Class I (higher grade professionals, self-employed or salaried) down to Class VII (manual wage-workers in industry in semi and unskilled grades; and agricultural workers). In 1981, a revised system is used which subdivides three of the original categories to give an 11-point schema, assigned using CO80.

The variables GC70, GC80 and GC90 in the LS give Goldthorpe Classes for an individual at each census. Where LS data for either occupation or employment status was not available the value of 0 has been allocated. The variables are found in tables ME71, OCC81 and OCC91 respectively (see the CeLSIUS [Data Dictionary](#)).

### Erikson-Goldthorpe scheme

The Erikson-Goldthorpe scheme is a four-dimensional measure, discriminating between different values of employment status, between manual and non-manual occupations, between agricultural and non-agricultural work, and between 'service' and 'labour'. 'Service' refers to occupations where the individual is entrusted with responsibility for their own work, is usually paid by a monthly salary, and is rewarded by career progression and job security. 'Labour' refers to occupations in which the worker provides a set amount of labour, usually for a weekly wage, and whose primary motivation comes from their supervision by superiors rather than through incentives.

The scheme has 11 non-ordinal categories, as follows:

1	Higher-grade professionals, self-employed or salaried.
	Higher-grade administrators and officials.
	Proprietors and managers in large firms.
2	Lower-grade professionals.
	Higher-grade technicians.
	Lower-grade administrators and officials.
	Managers of small businesses.

3a	Higher-grade workers in routine non-manual jobs.
3b	Lower-grade workers in routine non-manual jobs.
4a	Small proprietors, self-employed with employees.
4b	Small proprietors, self-employed with no employees.
4c	Farmers and smallholders.
5	Lower-grade technicians, manual supervisors.
6	Skilled manual workers.
7a	Semi-skilled and unskilled manual workers.
7b	Farm workers.

Individuals in the LS can be assigned to each of these groups based on their occupation and employment status using a look-up table. The look-up tables are available in the LS database and individuals can be coded by CeLSIUS staff on request.

#### Cambridge scores

The Cambridge Scale of Occupations is a continuous measure of social and material advantage, based on the scaling of survey respondents' occupational friendship and marriage scores.

The scores take the values 00.00 to 99.99. Low values represent disadvantage, and vice versa.

The Cambridge score is a measure of similarity of lifestyle, and therefore of generalised advantage or disadvantage. There are two (comparable) scores for each occupational group, one for men and one for women, since men and women in the same occupational category do not necessarily occupy the same position in the structure of social advantage/disadvantage.

In assigning scores to combinations of occupation and employment status, a look-up table is used. Employment status is graded on a 7-point scale much the same as for SEG.

The Cambridge scores for 1971, 1981 and 1991 are in variables CAM70 (table ME71), CAM80 (table OCC81) and CAM90 (table OCC91).

Where there is no code available for the particular combination of occupation and employment status, the value on the Cambridge score is 00.00. Where LS data for either occupation or employment status was not available, the field is left blank.

**[NB: None of these classifications is currently available for 2001 or 2011 in the LS.]**

## **Problems with Occupational Social Class**

Most difficulties with Social Class arise because it is only applicable to those people who are in paid employment. Therefore, it does not apply directly to that part of the population who are not in paid employment (about 33% in 1991). Without an occupation, they cannot be assigned to a Social Class, although in most censuses those who have been in employment will record their *most recent* occupation on the form. For this reason, the proportion who cannot be assigned to a Social Class in 1991 is only 16%.

If a person cannot be allocated to a Social Class, then it is up to the researcher to decide how to deal with him/her. In the LS it is possible to reduce the size of the unclassified population by looking at *previous* censuses, or by using the Social Class of other household members. This problem is worse for those who have never been in employment, as there is not even a previous occupation to use. Such people are assigned to a 'not stated' group in the social class scheme. Similarly to those individuals who are not working at a particular census, Social Class for individuals who have never been employed could be approximated based on the Social Class of other household members.

There are also particular issues relating to specific groups in the population, such as women, children, retired people and people who have been sick for a long time.

### Women

In 1971, 48% of women aged 16-64 could not be allocated to one of the Social Classes I to V. In the 1981 the figure was 44%, and in 1991, 22%. Where such women are married, it has been usual practice to assign the husband's Social Class to the woman, but this practice has reduced as the number of women in paid employment has increased. One procedure for dealing with married or cohabiting couples is to apply the highest of the two Social Classes to both members of the couple, although differences between male and female occupational structures mean that this might not be the best way to proceed.

Even where women are in paid employment, the social class scheme and its underlying occupational classification are based on traditional male roles and are less appropriate for differentiating between groups of women. Furthermore, the social class scale has no place for the unwaged work done by women in the household.

### Children

Children do not have occupations, so they cannot be coded directly for Social Class. However, since the vast majority of children are living with, and economically dependent on one or more parents, it is possible to apply the parents' Social Class to the child.

Children who join the LS at birth can be assigned a Social Class by using the Social Class of the parents, either that of the parent whose occupation is recorded on the birth certificate (usually the

father if he is included in the registration), but more usually by the occupation of either parent at the time of the subsequent census enumeration, using the non-members' file.

#### Retired people

Socio-economic studies of older people can be difficult, since the retired do not have a current occupation. However, in the LS, this problem can be addressed using the census records of each person's most recent occupation (although in 1991, there is no information for those who have not worked in the last 10 years, such as men aged 75+ and women aged 70+). Additionally, in the LS it is possible to look back to occupation recorded in a previous census. This may be necessary for 2001, because occupational questions were not asked of those aged 75 years or over. The LS includes information on occupational Social Class at several time points and so provides a kind of occupational history. Therefore, even if there is no data for the current census, it might still be possible to use the data from a previous census. Alternatively, it is possible to use household measures (see later).

#### Long-term illness

A particular group of people who are not in employment are the long-term sick. Since these people are not in employment (and may never have been in work) they are not assigned to a Social Class and may need to be reconceptualised in order to be included in studies of social mobility or social inequalities.

## Other socio-economic measures

We have seen the problems that arise when using various Social Class measures to study social explanations for health and other differentials. Fortunately, there are a range of alternative measures in the LS that can be used instead of Social Class as indicators of social standing. Some of these are ecological variables, which are looked at later, but firstly this section concentrates on household measures.

**[NB:** household measures are not available for certain groups, such as those in communal establishments and others who have completed 'individual' rather than the usual 'household' census forms, because such people do not provide any household data. Also, visitors will usually be linked to the household that they are visiting rather than their home address, so the household data linked to their LS record does not apply to them. However, household data for the usual address of these people might be available in the 'absent members'<sup>10</sup> file (tables AM81 and AM91). In 2001 and 2011 individuals were enumerated at their usual address whether they were present on census night or not.]

There can also be problems using household measures for older people who have moved in with their adult children, because the household circumstances are more a reflection of the child's social position than their own.

### *Housing tenure*

Housing tenure can be used as a measure of social status because it is associated with both income and wealth.

As it is a household variable, it applies indiscriminately to all members of the household. This can be an advantage if social status is considered to be related more to the household circumstances than individual circumstances. For example, consider a 20-year old woman who works part-time in a menial job, but lives with her affluent parents who own their own home. Her individual circumstances suggest a low social status but in fact factors such as her health may be more closely associated with her household circumstances.

Housing tenure is represented in the LS by the variables TENURE7, TENURE8, TENURE9, TENH0 and TENH11, for 1971, 1981, 1991, 2001 and 2011 respectively. Each variable has a slightly different coding scheme, reflecting the differing categories on the corresponding census form for that year (search the CeLSIUS [Data Dictionary](#) to see the coding schemes used).

If housing tenure is to be used across censuses, it is necessary to find a common categorisation that will give continuity between the different coding schemes. For example, in 1971 there is no separate category for people renting from housing associations or renting with their job. These people would be

---

<sup>10</sup> An LS member who is enumerated at his/her usual address but whose whereabouts on census night is given as 'elsewhere' (because they are on holiday, at school, or away for any other reason). LS records for such people are kept in a separate 'absent, usually resident' members' file. Data on absent members is only available in the LS for the 1981 Census onwards.

grouped in with the private rental categories. It is a matter of deciding which categories to aggregate, depending on the years that are being studied.

Another difficulty is that certain variables change their distribution over time to such an extent that their meaning changes. An example is housing tenure: as owner-occupiers became more prevalent during the 1980s, the owner-occupier category became less affluent as a whole, and its use as an indicator of high social status has diminished, although this trend is reversing in the recent decades due to faster growths in house prices outstripping wages in less than buoyant economic conditions.

### *Household amenities*

The LS also includes information on household amenities, which give another indicator of socio-economic status or deprivation, as well as being of interest in their own right.

The census questions on household amenities have changed with each census, so the coding schemes for the corresponding LS variables are also different.

In 1971, the census form included questions on a variety of different household amenities, while in 1981, the range of questions was much reduced. In 1991, the same questions were used with the addition of one on central heating. In 2001, the questions on bath/shower and toilet were combined, so the question distinguishes those who have *both* amenities from those who do not. The table below summarises the questions on household amenities:

<b>Amenity</b>	<b>1971</b>	<b>1981</b>	<b>1991</b>	<b>2001</b>	<b>2011</b>
Car	•	•	•	•	•
Cooker	•				
Sink	•				
Bath/shower	•	•	•	•	
Toilet	•	•	•		
Central heating			•	•	•

The LS variables corresponding to these questions can easily be found by searching the CeLSIUS [Data Dictionary](#) on the description 'household'. The variable names generally correspond to the name of the amenity that they represent, e.g. SINK7, BATH8, WC9 and BSTH0.

The car access variable, represented by CARS7, CARS8, CARS9, CAVH0 and CAVH11 in the LS, has been used in conjunction with housing tenure to provide a useful alternative socio-economic indicator (Goldblatt, 1990<sup>11</sup>).

<sup>11</sup> Goldblatt, P. Mortality and alternative social classifications. In: Goldblatt, P. (ed.) *Mortality and social organisation*. Series LS no. 6: HMSO 1990

## *Education*

Information on education can be obtained from the LS because the census contains questions on qualifications. The problem is that from 1971 to 1991, the question only requires qualifications obtained after the age of 18 to be listed, so only degrees, professional and vocational qualifications are included. This means that you cannot distinguish between the educational qualifications of those people whose education stopped prior to age 18. There is, however, a separate question in 1971 that records A-level, Scottish Higher and equivalent qualifications obtained by those aged 15 or over (though retired persons over 70 were not expected to answer either this question or the higher qualifications question in 1971).

This situation changed in 2001, when the question required all qualifications from GCSE (or equivalent) onwards to be indicated (Q16 in the 2001 Census, and Q25 in the 2011 Census). This raises the prospect of using the 2001 or 2011 data retrospectively, for individuals who were present at one or more of the previous censuses, to relate education to individual circumstances at those earlier censuses.

The proportion of the population in post-18 education has greatly increased over the last few decades in both genders. As such, education is less valuable as a social indicator for very old people, because many of them will have left school at age 15 and not have any formal qualifications anyway.

For each census from 1971 to 1991, the most useful LS variables are those giving the level of the highest qualification obtained (EDUC7, QMLVHIQ8 and QMLVHQT9). These can be used to simply classify people according to whether or not they have a degree or other post-18 qualification.

In 2001 and 2011, there were questions on academic and professional qualifications including questions on all qualifications from GCSE and beyond. This is a real advance as it enables a much more detailed classification than is possible using the qualification data collected in earlier censuses, which tended to ask about more advanced qualifications which large proportions of the population did not have. However, a drawback of the question in 2001, but not 2011, was that it was not asked of people aged 75 and over. The variables in 2001 and 2011 are QUP0 and QUP11 (Qualifications), PQUP0<sup>12</sup> (Professional Qualifications), and HLQP0 and HLQP11 (Highest Qualification, Academic and Professional) respectively. Code is available, which combines these variables in 2001 and in 2011 to distinguish graduates (those who have a degree) from people holding similar qualifications.

In addition, in 1971 there are LS variables giving the level, subject and awarding institution for each of up to six qualifications obtained after the age of 18 (QAL1Q7, QAL1S7 and QAL1I7 for the first qualification; QAL2Q7, QAL2S7 and QAL2I7 for the second qualification, and so on to the sixth qualification obtained after age 18. There is also GCE7, which gives the combination of A-level and equivalent qualifications obtained. In 1991, there are equivalent variables for level and subject only

---

<sup>12</sup> No equivalent variable in 2011.

(QMQUAL19 and QMSUB19 for the level and subject of the first qualification obtained by 1991, QMQUAL29 and QMSUB29 for the second qualification obtained, etc.).

### *Area-based measures*

Area-based measurements can be useful for several reasons. They can be used to inform the distribution of health service resources, and they can be used to analyse the effects of an area's characteristics on health independently of, or in conjunction with, any individual-level effects.

The LS contains the area-based socio-economic indicator, the Carstairs score<sup>13</sup> at census ward level, which has been mapped to individuals for 1981 and 1991 (CARSCO8, CARSCO9). However, it may be necessary to use more recently-calculated scores which have not currently been adopted by the ONS but are available for all census years pending on an acknowledgement to the user who generated them. Your CeLSIUS User Support Officer can advise you further. It is also possible to add external area-based measures to the LS, such as Townsend scores<sup>14</sup> or Breadline Britain indices<sup>15</sup> through the use of a lookup table. For links to deprivation indices, see the Geography guidance.

The mapping of values from lookup tables of external aggregate data to individuals in the LS can be done by researchers and CeLSIUS staff down to the local authority (or equivalent) level<sup>16</sup>, or by the ONS Data Custodian down to finer geographical levels, subject to confidentiality constraints<sup>17</sup>.

There are some problems with the use of such area-based measures:

- The ecological fallacy is an important potential limitation. It results from the false assumption that inferences can be made about individual phenomena based on observations of groups. For example, by no means are all people living in deprived areas themselves deprived.
- Areas are not internally homogenous: a population containing a mixture of deprived and less deprived households is likely to have a middle ranking score. Such mixed populations are more likely to occur in rural areas and large urban centres (e.g. London). It is therefore important to ensure that the geographical scale is appropriate for the analysis being undertaken.
- Areas with large institutional populations: the location of an institution in a particular area may have little effect on the circumstances of its inhabitants, since they are effectively insulated from the surrounding area, particularly so in the case of prisons.

---

<sup>13</sup> A measure of area-level deprivation, similar to the Townsend score. It is calculated as the sum of z-scores for 4 different proportions: 1) the proportion of households in the area that are overcrowded; 2) the proportion of heads of household in the area that are in social classes IV and V; 3) the proportion of male heads of household in the area that are unemployed; and 4) the proportion of households in the area that are in non-owner-occupied properties. Values range from about -7 to 16. The distribution is a bell curve skewed towards the positive.

<sup>14</sup> A measure of area-level deprivation, similar to the Carstairs score. It is calculated as the sum of z-scores for 4 different proportions: 1) unemployed residents over 16, as a percentage of all economically active residents aged over 16; 2) the proportion of households in the area with 1 person per room and over; 3) the proportion of households with no car; and 4) the proportion of households not owning their own home.

<sup>15</sup> An estimate of the percentage of 'poor' households in an area, obtained by a weighted sum of the following six variables: unemployment, lack of owner occupied accommodation, lack of car ownership, limiting long term illness, lone parent households, and low social class. Constructed as a result of a survey carried out for London Weekend Television in 1990.

<sup>16</sup> The third level of geographical division used in the LS, after standard regions and counties. Equivalent to a non-metropolitan district, metropolitan district, unitary authority or London borough.

<sup>17</sup> This may require a request for linkage of x-files, which are restricted data (your CeLSIUS User Support Officer can advise you further on this).

## **Cross-sectional studies**

This section works through an example of a cross-sectional study using the socio-economic indicators in the LS.

### **Research question:**

*How has the relationship between Social Class and other socio-economic indicators (housing tenure and car access changed over the years 1971-2011 in different regions of England and Wales?*

To examine this, social class is cross-tabulated against a combined tenure/car access variable at each of the census years, 1971, 1981, 1991, 2001 and 2011, disaggregating the results by standard geographical region.

### Study population

The study population should consider people of working age (16-64), because no one outside this age range is expected to have an occupation that could be used to allocate them to a Social Class.

**[NB:** There will still be a small proportion within this population who do not record a current or most recent occupation on their census form, and so will not have a Social Class; these individuals are allocated to a separate 'undefined' group.]

Since the question also examines household characteristics, the study sample will exclude people living in communal establishments and visitors; the LS does not contain household information for these people on their usual place of residence.

### LS variables

At each census, we need the Social Class of our subjects.

In 1971 this is given by the variable SOCS7 (i.e. using the 'statistics' coding rather than the 'census' coding). For this variable, the codes 0 to 5 represent the six major Social Classes from I to V. We can combine the codes 7 (armed forces) and 9 (for miscellaneous other categories) into a single category called 'Other'. The code 8 represents people whose occupation is 'inadequately described'. This being the case, they might equally likely be placed in any of the Social Classes, so for this study we will combine such individuals with those who have no Social Class code at all, in a 'Missing' category. A similar process can be undertaken at the other censuses using the variables SOC8 (1981) and SCLAS9 (1991), both of which use very similar coding schemes.

In 2001 and 2011, the variables ROSERG0 and RGSC11 respectively can be used. However, while the categories are similar, the codings used for those categories are different. For example, in 2011, the codes 10, 20, 31, 32, 40 and 50 represent the six major social classes from I to V. Therefore, these would need to be recoded to 0 to 5, and 60 (Armed Forces) and -9 (Not applicable [Aged under 16, Student not at their term-time address, Never worked, Long-term unemployed, Full-time students])

could be combined into one group called 'Other' (i.e. similarly to what was done for Social Class in 1971). Finally, -6 (Missing) would remain as a separate category.

The research question entails comparing Social Class with the combination of housing tenure and car access. For this, a new variable needs to be derived for each census, which can be called TENCAR7, TENCAR8, TENCAR9, TENCAR0 and TENCAR11 for 1971, 1981, 1991, 2001 and 2011 respectively. This will be based on the values of TENURE (TENURE7, TENURE8, TENURE9, TENH0, TENH11) and CARS (CARS7, CARS8, CARS9, CAVH0, CAVH11) at each census. As was shown earlier, the coding scheme for housing tenure has changed over the years. For this study, people renting from housing associations or with their jobs can be combined with the 'private renter' category, and the New Town households can be combined with the local authority tenants to make a 'social rental' category. This allows comparison between the different census years. The car access variables simply give the number of cars available to a household, so they are similar at each census.

This will result in the following categories of TENCAR for each census year:

- Car access, owner-occupier
- Car access, private rental
- Car access, social rental
- No car access, owner-occupier
- No car access, private rental
- No car access, social rental

If either the housing tenure or car access variables is missing, then TENCAR is coded missing.

Finally, the research question requires the identification of the Standard Region in which each individual lives. In 1971 this is done using STRGX7, and the categories for standard region remained the same for 1981 (STRGXUA8) and 1991 (SRGUAEA9), although the order in which they were coded changed in 1991. Between 1996 and 1999, the Standard Regions were replaced by new 'Government Office Regions' (GORs) which had different boundaries and introduced a new region for London (previously part of the South-East region). Therefore, the eight Standard Regions that were used for the 1971, 1981 and 1991 Censuses became nine GORs for the 2001 and 2011 Censuses. A standardised variable GOR is available for all censuses (ask your CeLSIUS User Support Officer).

### Results for the Northern region

The data request for this study produced a very large number of tables, so rather than trying to examine them all, only two will be examined. The Northern region and the South-east region will be examined, to see how the distribution of the social indicators has changed over the years.

Tables 1-5 show tables of Social Class by the combined housing tenure/car access variable for the Northern region of England, for each of the census years 1971, 1981, 1991, 2001 and 2011 (see tables 1-5 in the associated Excel spreadsheet).

**[NB:** the boundaries of the North and North West regions were changed in 1999 and became the North East and North West regions. Therefore, although the results of the cross-sectional analysis for 2001 and 2011 are shown (tables 4 and 5), they are not discussed since they are not directly comparable.]

The column percentages (tables 1a, 2a and 3a) in the totals column show that the overall distribution of Social Class has flattened out, with people moving from the IIIM class up into classes II and IIIN. There has also been a big decrease in the proportion who are not coded to one of the main Social Classes I to V. Looking at the individual columns we can see the same pattern in most of the car/tenure categories.

Examining the row percentages (tables 1b, 2b and 3b), we can see that the major change from 1971 to 1991 has been a large increase in the proportion of people in the 'car access/owner-occupier' category group from 27% in 1971 to 63% in 1991. This has been matched by decreases in all the other categories, most notably the 'no car access/social rental' group.

A similar distribution is also seen in each of the Social Classes but is emphasised in Classes IIINM and below where the shift to the 'car access/owner-occupier' has been greater. Also, in Social Class V there were increases in both the 'no car access/owner-occupier' and 'car access/social rental' categories, whereas the proportions of the overall population in these categories decreased.

#### Results for the South-eastern region

Tables 6-10 show Social Class by the combined housing tenure/car access variable for the South Eastern of England, for each of the census years 1971, 1981, 1991, 2001 and 2011 (see tables 6-10 in the associated Excel spreadsheet).

**[NB:** the boundary of the South Eastern region changed in 1999 and became the South East and London. Therefore, although the results of the cross-sectional analysis for 2001 and 2011 (tables 9 and 10) are shown, they are not discussed here since they are not directly comparable.]

The column percentages (tables 6a, 7a and 8a) show more people in the higher Social Classes than we saw in the Northern region, but the change between 1971 and 1991 is similar. Again, there is an increase in the proportion of people in Social Classes II and IIIN, with a big corresponding decrease in those not coded to one of the Classes I to V.

Looking at the individual columns, we can see a greater than average shift out of the lower Social Classes for those in the 'no car access/private rental' and 'no car access/social rental' categories, but

without the corresponding growth of the middle Social Classes that we saw in the Northern region. There is also a greater than average shift into Social Class II for people in the 'car access/private rental' group, and a similar shift out of Class IIIN for those the 'car access/social rental' category.

The row percentages (tables 6b, 7b and 8b) show that the proportion of people in the 'car access/owner-occupier' group has risen, though the rise is much less marked than it was in the Northern region, going from 39% in 1971 to 68% in 1991. There were significant decreases in all the other categories, just as in the Northern region. The patterns in the individual rows are very different to those in the data for the Northern region. Here we see a decrease in the size of both the 'no car/private rental' and 'no car/social rental' categories for the middle and lower Social Classes.

## **Longitudinal studies**

### **Research question:**

*Do children living with higher social class parents in 1971 have higher educational qualifications in 1991? Is the pattern different for men and women and for the period 1991-2011?*

To examine this, the non-member's files for 1971 (for the period 1971 - 1991) and 1991 (for the period 1991 - 2011) need to be used to find out the Social Class of each study member's parents. This is an example of using parents' social class as a proxy for the social class of the child. Where both parents were working in 1971/1991, it is necessary to decide which parent to use as representative of the child, and also what to do with children whose parents did not have a Social Class assigned.

### Study population

The study subjects need to have been dependent children in 1971/1991 (i.e. aged 15 or under) and old enough to have completed a course of higher education by 1991/2011 (i.e. aged 22 or over). That means the subjects should be aged 2 –15 in 1971/1991. Furthermore, each subject should be living with one or more parents in 1971/1991, so that it is possible to use the parents' Social Class to represent the child's. Parents should not need to be excluded on the basis of age; as parents of children aged  $\geq 2$ , the parents will almost all be aged  $\geq 16$ . Any parents who are  $\geq 65$  in 1971/1991 can be included in the study, despite being retired, because in 1971/1991 they should have recorded their most recent occupation on the census form, and their Social Class is based upon this.

Since the question requires identification of parents it is only possible to consider private households, where the relationships between household members are known. Children in communal establishments are excluded, which introduces a possible bias here, because this will exclude children at boarding school, whose parents are likely to be in the higher Social Classes; or conversely, children in residential homes or hospitals whose parents are more likely to be in the lower Social Classes. However, the 1971 and 1991 Censuses were taken during the holidays for many such institutions, so the number involved are small: only 1,543 LS members aged 15 or under (1.17%) were in communal establishments in 1971.

Visitors also need to be excluded, since the relationship data in the LS will apply to the address they are visiting rather than their usual address, and therefore may not be appropriate; for example, a child might be visiting the parent with whom they do not usually live and on whom they are not dependent.

### LS variables

Where there is more than one parent, the parent with the highest social class needs to be identified, and their social class as the representative social class of the family as a whole. For example, if the father is in social class II (SOCS7 = 1) and the mother in social class IIIN (SOCS7 = 2), then the lower

value is used to represent the child<sup>18</sup>. This is only one method; we could alternatively choose to use the father's Social Class throughout, or the mother's, although this would exclude some children of lone parents.

Educational qualifications in 1991 and 2011 can be identified using QMLVHQT9 (1991) and HLQP11 (2011), which give the level of the highest qualification obtained by an individual. The codes for QMLVHQT9 are as follows:

- 0: No 18+ qualifications
- 1: Level 'a' – Higher degrees of UK standard
- 2: Level 'b' – First degree and all other qualifications of first degree standard
- 3: Level 'c' – Qualifications that are obtained at 18+ and above GCE
- 9: Not applicable/missing

The codes for HLQP11 are as follows:

- 10: no academic or professional qualifications (England and Wales and Northern Ireland)
- 11: Level 1: 1-4 O Levels/CSEs/GCSEs (any grades), Entry level, Foundation diploma, NVQ level 1, Foundation GNVQ, Basic/Essential Skills (England and Wales and Northern Ireland).
- 12: Level 2: 5+ O Levels (Passes)/CSEs (Grade 1)/GCSEs (Grades A\* - C), School Certificate, 1 A Level/2-3 A/S Levels/VCEs, Intermediate/Higher Diploma, Welsh Baccalaureate, Intermediate Diploma, NVQ level 2, Intermediate GNVQ, City and Guilds Craft, BTEC First/G.
- 13: Apprenticeship (England and Wales and Northern Ireland).
- 14: Level 3: 2+A Levels/VCEs, 4 A/S Levels, Higher School Certificate, Progression/Advanced Diploma, Welsh Baccalaureate Advance Diploma, NVQ Level 3; Advanced GNVQ, City and Guilds Advanced Craft, ONC, OND, BTEC National, RSA Advanced Diploma (England and Wales and Northern Ireland).
- 15: Level 4: Degree (BA, BSc), Higher Degree (MA, PhD, PGCE), NVQ Level 4-5; HNC, HND, RSA Higher Diploma, BTEC Higher level, Foundation degree (NI), Professional Qualifications (Teaching, Nursing, Accountancy) (England and Wales and Northern Ireland).
- 16: Other: Vocational/Work-related Qualifications, Foreign qualifications/Qualifications gained outside the UK (Not stated/level unknown) (England and Wales and Northern Ireland).
- 6: Missing.
- 9: No code required (Aged <16, Students not at their term-time address).

This study needs to divide the LS members into those with degree-level or higher qualifications (1 and 2 in 1991, and 15 in 2011) and those without (0 and 3 in 1991, 10-14 and 16 in 2011). Therefore, for each study period (1971-1991, and 1991-2011) two new variables need to be derived:

- one giving the highest social class of the subject's parents; and
- one showing whether or not the subject has a degree-level qualification.

---

<sup>18</sup> The lower value represents the higher social class.

These variables can then be cross-tabulated and broken down by sex for each of the study periods.

### Results

1971-1991 (see table 11 in the associated Excel spreadsheet).

The results of this study were as follows. For all study subjects, the results show a clear gradient, with a greater proportion of people obtaining higher education qualifications by 1991 when their parent's Social Class in 1971 is higher. Looking at the results disaggregated by sex, the tables show that overall, a higher proportion of men in the study population had obtained higher education qualifications by 1991, but the pattern with respect to parent's Social Class in 1971 is much the same for both men and women.

1991-2011 (see table 12 in the associated Excel spreadsheet).

The results of this study for the period 1991-2011 were as follows. For all study subjects, the results again showed a clear gradient, with a greater proportion of people obtaining higher education qualifications by 2011 when their parent's Social Class in 1991 is higher. Looking at the results disaggregated by sex, in contrast to 1971-1991, the tables show that overall, a higher proportion of women in the study population had obtained higher education qualifications by 2011 compared to men, which was the opposite to that which was seen in 1971-1991. However, the social gradient in the achievement of a degree qualification or higher was similar to that seen in 1971-1991.

## **Summary**

The LS is a very good source of occupational data for individuals based on their census records. We have shown how these data are converted into socio-economic measures, which are themselves available as LS variables.

Even where an individual has no current occupation, the LS may still provide occupational data because the census records the individual's most recent occupation, and where this is missing we can use occupation from previous censuses at which the LS member was present.

The LS also contains socio-economic measures other than those based on occupation:

- individual-level data on education;
- household-level data on tenure, amenities and occupation of other household members; and
- area-level measures of deprivation.