1961-1981 Census Microdata

Guide to the 1961-1981 Samples

Describes data for 1961-81

Version produced for:

1961 and 1981 Census Microdata Teaching Datasets for Great Britain (1%)

1961 and 1981 Census Microdata Safeguarded Household files for Great Britain (0.95%)

1961 and 1981 Census Microdata Safeguarded Individual files for Great Britain (5%)

1961 and 1981 Census Microdata Secure files for Great Britain (9%)

Produced by Jo Wathan on behalf of the EEHCM Project

EXTERNAL

29 June 2017

Version 2.0

UK DATA ARCHIVE

UNIVERSITY OF ESSEX

WIVENHOE PARK

COLCHESTER

ESSEX, CO4 3SQ

T +44 (0)1206 872001

Funded by ESRC Grant Ref.: ES/K005731/1

Contents

User Guide to the 1961-1981 Census Microdata Samples

1. Introduction

This document constitutes an introduction user guide to the 1961 – 1981 Census Microdata Samples, which were produced under the auspices of the Extending and Enhancing Census Microdata project funded by the Economic and Social Research Council. Users of the files are advised to consult both this document and documentation specific to any files used; this additional documentation includes a codebook and glossary for the file.

1961 was the first census which used computers to process outputs. The 1961-1981 Census Microdata Samples are files which have been recovered from available *existing digital tapes* which were generated by census offices as part of the process.

2. What files are available?

Four different files are intended for each Census.

- A large 9% secure file which contains sufficient information to constitute personal data.
- A large 5% individual safeguarded file. This file contains information on larger local authorities.
- A 0.95% household safeguarded file. This links household members together to allow individuals to be understood within their household context.
- A 1% individual file which is available under Open Government Licence which can be used as a taster file.

This document is produced for the first files produced (1981). This document will be updated with each data release.

3. The Extending and Enhancing Historical Microdata Census Microdata Project

The files were created by a project known as Extending and Enhancing Historic Census Microdata Samples (EEHCM), which was funded by the Economic and Social Research Council.

3.1. The project team

The project was led by Prof Matthew Woollard at the University of Essex and was a collaboration between The University of Essex, CMIST at the University of Manchester and the Census Offices. The team included:

- *Project Management and Data Recovery*: Tanvi Desai, Graham Jolliffe and Arne Wolters (all at UK Data Archive)
- *Specification, Consultation and Enhancement*: Jo Wathan (Co-Investigator), Lee Bentley and John McLoughlin (all at CMIST, University of Manchester).

The project also involved collaboration and oversight from the Office for National Statistics and National Records of Scotland.

3.2. Project scope

The project was structured in two phases. In the first phase the team sought to recover existing digital records, to perform some quality assurance and to return these digital records to the Census Offices for archiving and further use.

Secondly, the project sought to produce researcher microdata files akin to the Samples of Anonymised Records (1991-2001) or Census Microdata (2011). This work was undertaken in response to a small scale consultation, feedback from the project's advisory group and in response to feedback from the Census Offices.

3.3. Project management and oversight

The project was conducted with a high degree of scrutiny. During the duration of the funded project was under the oversight of the Office for National Statistics. Oversight took the form of a project board and routine reporting. The Office for National Statistics housed computing facilities which enabled much of the data recovery work to be undertaken.

3.4. User needs

Additionally the project reported to an advisory group which met twice during 2013, this group included census offices, census researchers, population historians, an IPUMS representative and others. Many of these individuals also attended a meeting in late 2014 which provided an update marking the end of the funded project.

The project undertook a small scale non-probability consultation survey to provide insight into user needs which had 28 responses. Most respondents were interested in the data for academic research and levels of interest in 1961, 1971 and 1981 data were very similar. Additionally a small number of expert users were contacted individually to seek advice on the treatment of individual variables such as ethnicity, occupation, migration and social grade.

Users of historic files have different needs, most notably some users require variables which are as authentic to the data collection period as possible. On the other hand most of the respondents to the consultation survey preferred data which enable maximum comparison over time. While, effort has been made to try to meet both sets of needs, this has not always been possible either due to data availability or confidentiality issues.

4. Source of the data

The project was only possible as existing digital records existed. Each of the 1961-81 Census Microdata Sample files were produced from these existing digital records. Where the digital record contained multiple files, the files recovered were those identified as definitive in that they contained a full set of records (subject to data quality issues as described below) for the full set of topics. Importantly, it should be noted that the digital records are distinct from the hard copy census records which are kept securely and data from which are expected to be released a hundred years after they were collected.

Not all data captured in the forms were available for the team to reuse; identifying information had been removed, which may have limited some recovery efforts. In 1961, for example, the file available for reuse included only 5 year bands rather than age in years, rather than date of birth.

Additionally, and importantly, these digital records produced precede current data archival standards and were created before microdata sets for secondary use were anticipated. The data had been transferred from one format to another over time. A process of data recovery and quality checking

was necessary to maximise their utility. Users should be aware that the data quality is imperfect and varies from file to file this problem is most marked for 1971 where over 10% of cases are missing.

Following sections describe major known issues.

4.1. Ten percent files

The datasets have been drawn samples from the 10% fully coded files as the 100% files would have omitted critical hard to count socio-economic information, such as occupation.

As with later censuses, not all information was collected and/or processed for all respondents. This principle applies as some information is harder to capture than other. In particular, not all questions are 'closed' tick box questions; questions that relate to addresses, occupation and industry for example tend to be 'open' write-in questions. It has been the case in most censuses these hard to code questions are not coded for the full population for resource reasons.

Accordingly, in all of the years available additional information is available for a 10% sample, than for the full population.

- In 1961 two separate forms were used, a long form delivered to ten percent of the population and a short form delivered to the other ninety percent. The selection of the ten percent was designed to be one which was systematic. Enumerators were required to deliver a long form to every tenth household.
- In 1971 and 1981 a single version of the form was used for all one hundred percent. However, hard to code questions were processed only for a ten percent sample.

4.2. Data quality issues

Each of the files has issues affecting the quality of the data. Some relate to the quality of the original census data (measurement error and issues relating to the selection of the 10% sample) and others relate to the recovery process (file loss and interpretability). None of the files have been adjusted for any bias.

4.2.1. Issues with the 10% sampling

In 1961, a systematic sample was planned. Each enumerator was given a set of forms to deliver in the order they should be delivered. The set included long forms to be completed by 10% of households and shorter forms for the remaining 90%. The long forms were interleaved in every tenth place, with the start point determined at random. There is evidence that this procedure was not followed, with the result that the sample is believed be biased such that households of size one and large households are under-represented.

Users should be aware that 1961 Census reports contain "bias corrections", which address biases as compared with the 100% data with respect to the number of persons and sharing stated, these have not been incorporated into the data. Ideally data would have been weighted to deal with these known biases, however this was not within the scope of the data recovery project. The size of the biases is illustrated well by the following table:

Table 1 Percentage excess in 10% sample when compared with one tenth of the full count figures

Persons in the household	Number of rooms occupied							All households
	1	2	3-4	5	6	7-8	9 and over	
1	-12	-11	-9	-7	-4	-1	+11	-8
2-6	-7	-1	0	+1	+5	+10	+22	+2
7 and over	-28	-27	-15	-10	-9	-7	+1	-10
All households	-10	-5	-1	0	+3	+8	+19	0

Source: Census 1961 Great Britain Summary Tables pxxxiv

More information is available in the Census 1961 Great Britain Summary Tables (General Register Office London and General Register Office Edinburgh 1966). Users should be aware that published tables were adjusted by bias factors (listed in Appendix C of the Summary Tables).

4.2.2. Coverage of the recovered source files

A benchmarking process was undertaken in the first stage of data recovery. Further information is given a separate file for each year. The following summarises features of the regional benchmarking. This benchmarking was undertaken by comparing the 10% file from which the census microdata files were drawn with published tables.

In some instances the available, recovered records did not provide a complete and unbiased representative sample of the population. In particular, source tapes had been produced in geographical order. The original tapes had been transferred to a number of different files. Accordingly any failure to retrieve all of the complete source files affects specific geographic areas. The most striking under-coverage issue is often expressed as the total omission of some areas from the recovered files.

1981: Overall coverage is believed to be very good. Only one region (Isle of Wight – not normally counted as a region) had a population more than 1% different than expected from the published figures.

1971: While coverage is believed to be good in most areas, some large blocks of data were not available for recovery, which affects three counties in particular: Huntingdon and Peterborough, Kent and Lancashire. The former two are entirely missing while around approximately 80% of Lancashire is missing. Lancashire was a very large county in 1971, second only in population to Greater London. The overall loss of coverage for England and Wales is 13.2% accordingly this file cannot be considered to be representative of the whole country.

1961: Coverage by region is believed to be good. Three regions were undercounted by more than 1%: South Western (1.1% most notable in Somerset), Northern (1.2%) and London and South Eastern (1.9% most notable in Middlesex).

These shortcomings are specific to the digital records used as source files for the datasets (rather than original paper files which were not available to the project team).

4.2.3. Contemporaneous Quality Checks on the census data

From 1961 onwards data quality studies were undertaken after census enumeration had taken place. Some of the non-sampling error issues have been identified in published reports. The following section describes the most notable aspects found. In each quality check a survey was undertaken with a sample of individuals. The results obtained from the census were then compared with those from the survey to establish the rate at which results differed using the two methods. Not all questions were included in the relevant Quality Checks let alone reported. Accordingly, omission from the lists below does not indicate the absence of measurement error.

1961

A multistage post enumeration survey was undertaken in England and Wales only. Findings are therefore generalisable only to England and Wales, but may be of some use when considering GB

more widely. The results of this survey were considered alongside administrative sources such as registration data to assess the quality of the census returns.

While the level of enumeration for the 100% sample was considered accurate overall, capturing recent immigrants poses particular problems. Figures from the Home Office suggested that as many as two hundred and six thousand West Indians may have been in the country in 1961, while the Census picked up one hundred and seventy two thousand – an undercount of around 20% (of the census figure).

The allocation of those who wrote 'Ireland' or 'Eire' were allocated to Ireland 'part not stated', however the post enumeration survey suggests that these people were generally born in the Republic of Ireland.

The number of households without a cold water tap is believed to be understated, however it was difficult to assess the extent of the problem as the group is small. Similarly, the number of households without a water closet it is believed to be considerably underestimated (to the extent that 1.5 million should have been enumerated as compared with the 1.0 million returned).

Length of residence is believed to be somewhat overstated by approximately a year, however this has not affected the number of migrants.

The number households renting from a private person is believed to be understated.

In terms of occupation it is believed that some men who were assigned as coal miners, electrical engineers or engineers should have been assigned to other occupational units within the same order. It is believed that the number of women in the private domestic service industry is overestimated. It is believed that there was some confusion of skilled, semi-skilled and unskilled manual workers in the socio-economic classification. The number of women classified as economically inactive is overstated.

Finally, it is believed that the number of people who are recorded as having completed education aged 17-19 is understated, while those recorded as having completed education aged 14 is overstated.

Sources: Census 1961 Great Britain Summary Tables (HMSO, 1966), Census 1961 Great Britain General Report . Further information is available in the Census General Report (HMSO, 1968).

1971

Information about the Quality Check of the 1971 census can be found in volume 3 of the Census 1971 General Report (Office of Population Surveys, 1983). The following table summarises the misclassification rate of each available topic available in the report.

Table 2: Errors estimated as part of the 1971 Quality Check

Topics	Error rate	Nature of issue
Date of birth (difference of more than one month)	1.0%	Around one third of errors are believed to be simple counting errors. Around 40% of the estimated error relates to error of less than a year. Overall there is very small effect on the age distribution.
Respondent's country of birth	0.9%	Main sources of known error arose where birthplace had not been stated or mis-classification of individuals other than the form filler as having been born in England when this was not the case or lack of precision for this same group of people.
Year of Entry into the United Kingdom	13.7%	The largest source of error was omission, which is likely to particularly affect the "before 1940" group. Additionally it was believed that the question was misinterpreted to mean year

		arrived *to settle*.
Country of birth of father/ mother	4.5% /5.1%	Country of birth was not always answered. The quality check also identified confusions between home nations. It is not possible to know whether this problem also affected foreign countries as the survey sample too small to pinpoint problems with less common countries.
Economic position	7.2%	Some errors were believed to have arisen as students and housewives could also validly have additional roles and it was not possible for form fillers to prioritise answers. Additionally there is some evidence that there was confusion over the meaning of 'sick' and 'wholly retired'.
Employment Status	9.7%	The greatest source of error was where the question was unanswered. However, the report indicates that most of the errors were corrected in editing.
Dates of first marriage	8.7%	5.6% of cases had the start date wrong (often by 1 year or 1 month out, suggesting a simple error). 3.1% of cases omitted to answer start date in error. Additional error was present with date of end of marriage, where date of end was not recorded in error.
Date of birth of children born alive in marriage	8.9%	Where individuals had no children the form might have been erroneously left blank. Amongst those who had children, the most common error was that the incorrect data was entered.
Number of children born alive in marriage	8.8%	It is believed that omissions arose in relation to children of a former marriage, children being absent, not part of the household, deceased, or already included on the form. Overcounts arose where children were counted even though they were adopted or born out of marriage (and therefore should not have been included).

1981

This information about the quality check is drawn from the Census 1981 General Report for England and Wales (Office of Population Census and Surveys, 1990). It is based on section 19.23 which lists the measures which had the highest gross error rates according to the Quality Check.

Table 3: Errors estimated from the 1981 Quality Check

Gross error rate	**Topic**	**Nature of issue**
28.6%	Number of rooms	Most errors related to a difference of one room. The inclusion of bathrooms or exclusion of full-sized kitchens were particular problems.
20-25%	Occupation	Errors arose due to lack of detail or coding errors. It is recognised that there is some judgement involved in coding occupation.
16%	Socio-economic group	This variable is derived from occupation, industry and employment status and inherits errors from these topics.
13%	Social class Female error rate: 10.0 Male error rate: 14.4	This variable is derived from occupation, industry and employment status and inherits errors from these topics. No individual class has more than a 0.5 percentage point difference.
9.8%	Employment status	The 10% sample may have underestimated those who are self-employed by around 5%.
8.6%	Means of travel to work	Half of this error is due to a confusion of three categories of car occupant.
7.8%	Economic position	The errors were higher for women (11.2% cf 4.1%), whereby too few women were reported as economically active.
3.5%	Cars/vans	The percentage of households without a car is believed to be accurate.

3.2%	Tenure	Owner occupation was estimated to be overstated by 0.9% and rented from local authority understated by 1.0%.
2.5%	Type of accommodation	This relates to type of accommodation as captured by the enumerator
1.5%	Usual address one year ago	The number of people living at a different address one year previous is believed to have been under-counted by around 8%.
1.3%	Marital status	Remarried was undercounted and married was overcounted.
1.3%	Inside toilets	Shared use and no use are believed to be undercounted.

4.2.4. Variable label recovery

The extent to which labels were available for each variable in the collection varied from file to file.

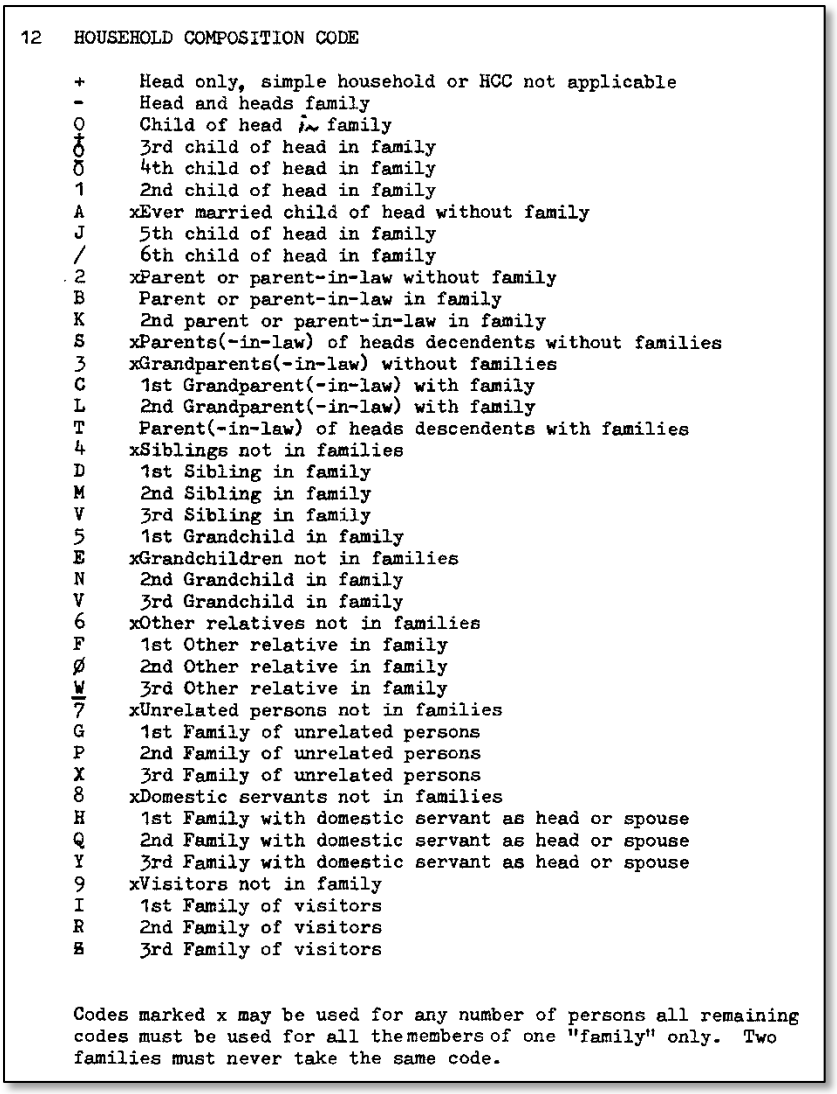
The data recovery team have worked with a number of sources in order to ensure that the data are presented and labelled correctly these include

- labels provided by ONS,
- contemporary documentation some of which were provided by the IPUMS team at the University of Minnesota,
- records from The National Archives,
- published census outputs
- and analyses of the data themselves.

Additionally, we consulted with staff in the Longitudinal Study in considering 1971 and 1981 data. Overall, we believe the process has been successful.

However in a small number of instances we include in the sample files a variable which continues to pose some problems of interpretation.

Fig 1: 1961 Household Composition Code



12 HOUSEHOLD COMPOSITION CODE

+	Head only, simple household or HCC not applicable
-	Head and heads family
O	Child of head in family
Ø	3rd child of head in family
Ö	4th child of head in family
1	2nd child of head in family
A	xEver married child of head without family
J	5th child of head in family
/	6th child of head in family
2	xParent or parent-in-law without family
B	Parent or parent-in-law in family
K	2nd parent or parent-in-law in family
S	xParents(-in-law) of heads decendents without families
3	xGrandparents(-in-law) without families
C	1st Grandparent(-in-law) with family
L	2nd Grandparent(-in-law) with family
T	Parent(-in-law) of heads descendents with families
4	xSiblings not in families
D	1st Sibling in family
M	2nd Sibling in family
V	3rd Sibling in family
5	1st Grandchild in family
E	xGrandchildren not in families
N	2nd Grandchild in family
V	3rd Grandchild in family
6	xOther relatives not in families
F	1st Other relative in family
Ø	2nd Other relative in family
W	3rd Other relative in family
7	xUnrelated persons not in families
G	1st Family of unrelated persons
P	2nd Family of unrelated persons
X	3rd Family of unrelated persons
8	xDomestic servants not in families
H	1st Family with domestic servant as head or spouse
Q	2nd Family with domestic servant as head or spouse
Y	3rd Family with domestic servant as head or spouse
9	xVisitors not in family
I	1st Family of visitors
R	2nd Family of visitors
Z	3rd Family of visitors

Codes marked x may be used for any number of persons all remaining codes must be used for all the members of one "family" only. Two families must never take the same code.

Most notable was the difficulty posed by household type in 1961. The original coding included some unconventional symbols (see fig 2) which did not transfer well as the data were ported to different formats over time. Despite extensive efforts to unpick the variable using all available documentation sources and reverse engineering efforts it was not possible to recover the variable labels. This omission has made it impossible to undertake a raft of variable derivation which might otherwise have been possible.

A question mark in final labels within the datasets indicates that the team were unable to confirm what they believed was a likely label.

A small number of unlabelled values remain. Key variables of this type may have been included in the data, despite their obvious shortcomings, as not to preclude further work to interpret them in the future.

An example of this sort of issue relates to usual residence and place of work in 1971. Notes received from the census office suggested that standard geographical coding applied, however no labels were available for codes which do not represent a given geographical area. Suggested labels have been attached on the basis of codes which were expected to be present and which fit patterns in the data. Value 3000 has been assigned as 'outside Britain' on the basis that 41% of employed persons who usually reside overseas take this value, compared with only 0.03% of employed persons who normally reside in the Britain. Value 3500 has been assigned on the basis of relationship to occupational detail; occupations which are disproportionately likely to take the value include, bricklayers (29%),

commercial travellers (36%), builders (36%) and plasterers (38%), all of whom are disproportionately likely to be working in different locations. The suggested labels fit expected distributions and instructions to respondents, which inter alia instruct respondents to give the location of a depot (if they worked out of one) rather than 'no fixed place'.

4.2.5. Known issues

4.2.5.1. General quality assurance issues for 1961

Although there is a reasonable level of documentation, some interpretation has been necessary when assessing how the data are stored either to label the data, to align coding schemes that vary across countries or to tease apart data in a stored single field which are best stored in more than one field. Where work of this type is necessary we have routinely sought to assure the quality of resulting data against other sources, particularly census tables.

However, there are some particular issues for 1961 which limits this activity.

- First, there are fewer published sources with which to compare data. We are typically limited to the 1961 Census tables. These are fewer in number than for 1971 or 1981 and, at the time of writing, are not available digitally, which means that any comparator data needs to be digitised.
- 1961 published tables were produced with only very limited access to computing facilities. Accordingly it is known that the published tables were subject to corrections post tabulation. We would not expect these corrections to be reflected in the microdata sources.

4.2.5.2. Country of Birth in 1961

Country of birth in 1961 was subjected to a degree of data manipulation in order to ensure that coding was applied consistently across Great Britain. However, in checking that the resulting distribution was correct it was found that although the distribution was of the right shape, there are some notable discrepancies between the published figures and those in the microdata. It is not clear at present which of several reasons explain these differences

- contemporary sampling and data quality issues which are not reflected in the published tables
- sampling and data quality issues in the files from which the EEHCM sample was drawn. In particular, the data have not been adjusted for known sampling bias.
- coding or sampling issues with the EEHCM sample

It should be noted that some of the numbers on which the percentage is based are relatively small and we would expect a reasonable amount of variation to arise from sampling. The 37% overcount of Swedes accounts to an overcount of about 191 cases.

Table 4: Country of birth comparison of 9% secure sample and Table 9 Census 1961 Summary Tables

Country of Birth (outside British Isles)	secure file	pop estimate (based on sample)	published figure from Table 9 summary tables	percentage difference (undercount in secure file)
Ghana	407	4,522	4,854	6.8
Nigeria	1,141	12,678	13,676	7.3
Rhodesia and Nyasaland	751	8,344	6,713	-24.3
South Africa	3,869	42,989	40,263	-6.8
Canada	5,499	61,100	56,611	-7.9

Ceylon	881	9,789	9,617	-1.8
Cyprus	3,018	33,533	42,283	20.7
India	14,437	160,411	165,869	3.3
Malaya	1,031	11,456	10,429	-9.8
Pakistan	2,132	23,689	31,891	25.7
Singapore	1,015	11,278	10,477	-7.6
Australia	3,966	44,067	39,263	-12.2
New Zealand	1,498	16,644	14,455	-15.1
Gibraltar	753	8,367	89,36	6.4
Malta	2,319	25,767	25,742	-0.1
Kenya	715	7,944	6,741	-17.9
Sierra Leone	147	1,633	2,216	26.3
Tanganyika	218	2,422	2,401	-0.9
Uganda	238	2,644	2,217	-19.3
Other Commonwealth territories in Africa	561	6,233	6,725	7.3
Aden	269	2,989	4,026	25.8
Hong Kong	952	10,578	10,878	2.8
Other Commonwealth territories in Asia	90	1,000	912	-9.6
British Guiana	717	7,967	10,889	26.8
Jamaica	6,545	72,722	100,410	27.6
Trinidad and Tobago	684	7,600	9,273	18.0
Other Commonwealth territories in the Caribbean	3,473	38,589	52,305	26.2
Other Commonwealth territories in America	70	778	782	0.5
Commonwealth territories in Oceania	77	856	959	10.8
Austria	2,793	31,033	31,298	0.8
Belgium	1,431	15,900	15,499	-2.6
Czechoslovakia	851	9,456	10,318	8.4
Denmark	809	8,989	7,480	-20.2
Finland	223	2,478	2,156	-14.9
France	3,011	33,456	31,495	-6.2
Germany	11,806	131,178	127,912	-2.6
Greece	669	7,433	7,268	-2.3
Hungary	1,448	16,089	18,272	11.9
Italy	6,564	72,933	87,243	16.4
Netherlands	1,695	18,833	16,479	-14.3
Norway	644	7,156	5,257	-36.1
Poland	10,186	113,178	127,246	11.1
Portugal	279	3,100	2,992	-3.6
Romania	407	4,522	4,608	1.9
Spain	1,795	19,944	21,363	6.6
Sweden	467	5,189	3,783	-37.2
Switzerland	1,437	15,967	13,174	-21.2
Yugoslavia	976	10,844	11,847	8.5
Other countries in Europe	197	2,189	1,708	-28.2

Sudan	93	1,033	1,032	-0.1
United Arab Republic	1,894	21,044	20,249	-3.9
Other countries in Africa	849	9,433	9,389	-0.5
Argentina	457	5,078	4,630	-9.7
Brazil	239	2,656	2,505	-6.0
Chile	193	2,144	2,013	-6.5
Colombia	45	500	473	-5.7
Peru	81	900	838	-7.4
United States	9,665	107,389	102,295	-5.0
Venezuela	70	778	851	8.6
Other countries in America	511	5,678	5,745	1.2
Burma	1,046	11,622	10,379	-12.0
China	901	10,011	9,852	-1.6
Indonesia	203	2,256	1,930	-16.9
Iran	369	4,100	4,035	-1.6
Iraq	313	3,478	3,570	2.6
Israel	430	4,778	4,559	-4.8
Japan	244	2,711	2,559	-5.9
Thailand	107	1,189	1,130	-5.2
Turkey	340	3,778	3,978	5.0
Other countries in Asia	533	5,922	5,515	-7.4
Countries in Oceania	19	211	208	-1.5
USSR	4,845	53,833	56,867	5.3
Born at sea	200	2,222	2,072	-7.3
Not stated	13,402	148,911		
Born in UK	1	11		
Total	4,614,477	51271967		
	Tot outside British Isles	**1583478**	1657158	4.4
	Tot outside British Isles	**1434567**	1495825	4.1

4.3. Population bases

As with other Census Microdata files, flexibility in the population base has been maintained where possible. This means that unlike a published table which will be produced for a single fixed population base (most commonly 'residents present' or 'usual residents') it may be possible to modify your choice of population base. Where this is so, the onus is on the user to determine a suitable population base and select only those cases which should be included. If this is not done, double counting of some individuals (especially residents who are away from their usual residence at the time of enumeration) may occur.

A variable called popbase is available in each dataset which should be used to define the population base.

Each of the files distinguishes between present residents, absent residents and visitors. This facilitates the distinction between the usual (aka 'de jure') and present (aka 'de facto') populations. These different groups were not treated equally in the census. Table 4 summarises which groups were fully enumerated versus partially enumerated.

Table 5: Extent of enumeration of population base components in households

Year	Population group(s) to whom most or all questions were asked	Population groups for whom limited data were collected
1961	Present residents, present visitors	Absent residents
1971	Present residents, present visitors,	Absent residents
1981	Present residents, present visitors (including those who arrived on Monday who had not been enumerated elsewhere) and absent residents	

The extent to which the population base can be recast depends on the availability of necessary data. Partially enumerated groups may have been asked a much smaller number of questions. So for example in 1961, in the 10% household sample forms absent persons were only asked approximately one third of the number of questions as compared with present persons. These questions were largely limited to issues of economic activity; migration, fertility and qualification questions were not asked.

Accordingly it is not possible to be entirely flexible throughout the whole period despite the presence of the population base information. Specifically, it is not possible to do full analysis on the usual residents (or 'de jure') population base for all time points, which has been used widely in from 1991 as full information was not collected from absent residents in each year.

Absent resident information was only collected for private households in each of the three censuses.

4.4. Households and Communal establishments

Individuals enumerated as part of a communal establishment are distinguished from those enumerated as part of a private household using the non-private variable in each year.

Members of large households and members of communal establishments are treated normally in the secure household file and individual files, with households being identifiable in the former. However, large households constitute a disclosure risk so additional protections have been imposed for the safeguarded household files. Households larger than 9 residents and members of communal establishments have been treated in a manner consistent with some IPUMS files rather than existing census microdata files for 1991 onwards. In these files, those enumerated in large households or communal establishments are included in all files including the household files. However, they are sampled as individuals and no household ID is included.

From 1971 the term dwelling was not used in instructions to enumerators rather the accommodation that *households* occupied were classified according to the amount of privacy enjoyed. (OPCS 1979: 13).

4.5. ID variables

A person ID is present in all files.

Each household file contains a household ID which has been generated in the process of 'flattening' the file from one which was structured by household, to one which gives a case for each person.

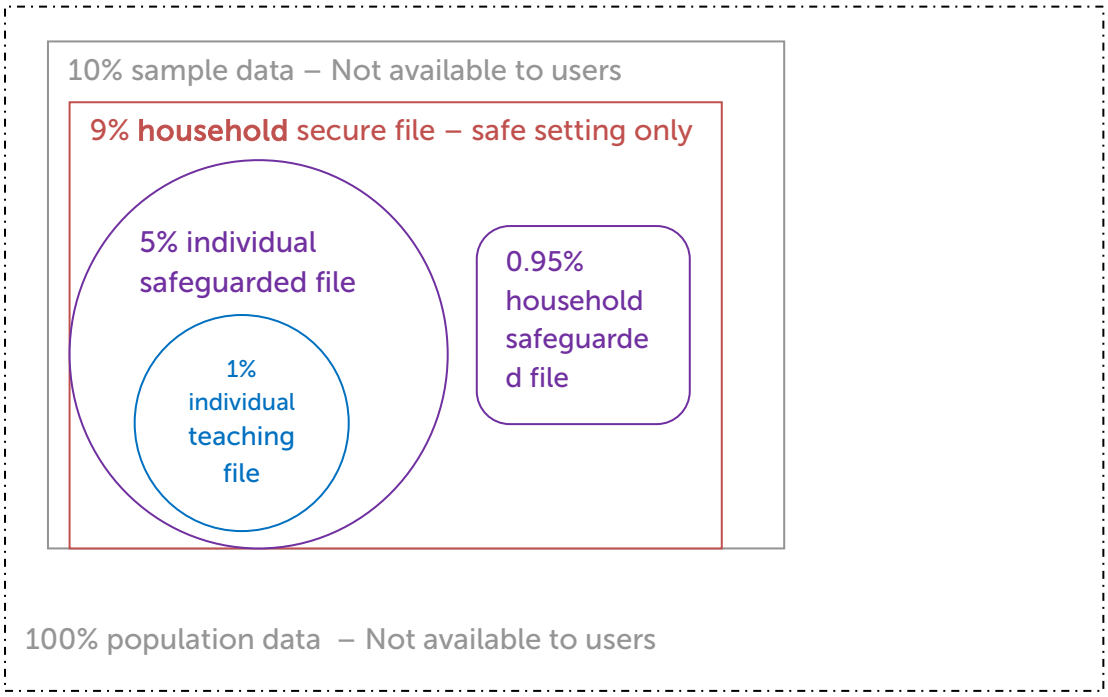
The ID variables were randomised, so are not ordered by geography below which the file is available.

Data were not stored in a manner as to enable the recreation of family units in all years. Accordingly it was not possible in all years to generate a family ID variable. This is included where it was available.

4.6. The sample process

The following diagram demonstrates the proposed relationship between the files. It is not to scale. Sample members cannot overlap across files from distinct samples. Samples, however, can be nested as any extra detail is only available to the users of the more detailed file (to whom by definition the extra detail is already available).

Fig 2: Sample design



Files represented by circles are individual files only, files represented by a rectangle contain linked household groups

The household secure file is available in a safe setting only, safeguarded files are intended for use by registered users and the individual teaching is designed for access under the Open Government Licence.

Sample sizes were agreed in discussion with ONS. The 9% secure file is designed to provide some uncertainty as to whether individuals from whom 10% data were collected[1] will be present in the output file. The size of the safeguarded household file constitutes a 'small' file. The other two files are equivalent in size to their 2011 equivalents.

4.6.1. The sampling method

Stratification

Following advice, we sought to generate samples which were stratified by household type and area. This should provide a slightly more efficient sample than would be possible using the method used in cross-sectional microdata files from 1991 which were stratified by geography alone.

[1] Members of the 10% file could have known if they were members in 1961.

Clustering

Against this, the files are also subject to clustering at the household level due to the sub-sampling of individuals from a limited pool of households. This feature was an unavoidable consequence of taking high sampling fractions from the complete 10% samples. As household ID numbers were not allowable in individual level files for reasons of confidentiality there is no effective PSU variable to allow users to readily adjust for this feature of the data in the individual files. Clustering can markedly reduce the precision of estimates, particularly in analyses using variables with less intra-household variation. In principle, we would expect the effect to be less than that of using a household survey without adjustment for clustering as the sampling method provided protection against the selection of entire households. However, the effect for some estimates can still be quite marked. Fuller information is given in Appendix B.

Sampling was undertaken using the following steps:

The Secure file was sampled first:

1. The 10% file was cleaned, header information (relating to area and household) was attached and variables were derived.
2. Household types were created (see the strat variable where available) that were large enough to facilitate stratification. Communal establishments were also identified as a separate 'household type'.
3. Communal establishments and households were separated and sampled separately.
4. To generate the 9% household sample, one case per household was identified to represent each case when defining the main sample.
5. These household cases were ordered by geography and household type
6. One in ten households was then systematically excluded. This had the effect of implicitly stratifying an otherwise randomly selected sample by area and household type.
7. The household members were merged back on to the selected 9% of households.
8. *Individuals* living within communal establishments were ordered by geography and then at random within geography. One in ten individuals was systematically excluded. No attempt was made to sample at the level of the communal establishment and communal establishments are not identified to enable these linkages to be made.
9. The two samples were combined to produce the Secure File.

The Safeguarded Household file was sampled from the Secure File:

10. Data were sorted by strata and geography and one person per household was selected to represent the household.
11. Individuals living in communal households or households with more than 9 residents were excluded from the following four steps.
12. A systematic sample 0.95% of households was selected in two steps.
13. One ninth of the households in the secure file were flagged by selecting the n_1th case in every nine, the number n_1 having been selected randomly.
14. Nineteen out of twenty of households were flagged by selecting all but the n_2th case in every twenty households (the number n_2 having been selected at random).
15. Households which were flagged by both of the previous 2 steps 12 and 13 constitute the sample of households of size 9 or under.
16. Household members were attached to selected households to produce a 0.95% sample of members of households with 9 or fewer residents.

17. *Individuals* living in households with more than nine residents, or in communal establishments were selected systematically using a similar two step procedure which selected 0.95% of individuals. In other words, there has been no attempt made to select households of size greater than 9 in their entirety.
18. The two samples of individuals within households of 9 or smaller and of those in larger households or communal establishments were combined to create the full sample.

The Safeguarded Individual file was selected third.

19. The Safeguarded Individual file was selected from the 9% secure file in such a way that any cases in the safeguarded household file were excluded from the safeguarded individual file.
20. Any case flagged for possible inclusion in the household file at step 13 was removed, with the result that 8% of cases remained.
21. 5 cases (the n_3th, n_4th, n_5th, n_6th and n_7th) out of 8 were systematically flagged for inclusion from the sample.
22. The flagged individuals constituted a 5% sample.

The Open Teaching contains one fifth of the cases in the Individual file.

23. This was done by selecting the cases flagged as being in the n_4th position of each systematic selection sequence at step 21.

5. Key discontinuities in concepts 1961-81

This section highlights some key differences in the way that key concepts were measured over the period 1961-1981.

Census data are used by public authorities in order to assess trends over time. Because of this there is always a desire to maintain sufficient consistency over time to enable comparisons between periods to be made. However, this factor is necessarily balanced against the need at any time to accurately capture the characteristics of the contemporary population using the best methods possible. The period 1961-2011 was one in which there were considerable social changes and the census methods have had to reflect these.

A very valuable resource for all users comparing data across years in the period 1971 to 1991 is "Longitudinal Study 1971-1991" by Lin Hattersley and Rosemary Creeser (1995, HMSO: London) which describes issues of census methodology and definition for this period in greater detail than is available here.

5.1. Geography

The amount of geographical detail on each of the output files varies according to the access mode of the file. The basic principle applied is;

Table 4: Geography in the microdata files

File type	Lowest level of geography
Files available under open government licence	Regional geography (based on Government Office Region)
Safeguarded files	Large borough/district

Controlled files	Ward or true local authority district.

It should be noted that geographical areas do not stay stable over time and throughout the entire period there have been numerous minor changes of boundaries.

5.1.2. Major local government reorganisations 1974 and 1975

The most major change of geography comes in England and Wales 1974 following the Local Government Act 1972 and a year later in Scotland following the Local Government (Scotland) Act. A uniform two-tier approach was introduced with counties at the higher level and either boroughs or districts at the lower level. The number of districts was slashed. One implication of the change is that authorities went from being distinguished between urban ones and rural ones to unified ones which contained both rural and urban areas. While the distinction was not a clear cut one, towns had routinely been identified as separate authorities from their rural surroundings. The 1971 census microdata are stored with geography appropriate to 1971 (i.e. unlike some outputs the data have not been recast for the post 1974 reorganised districts).

5.1.3. Greater London

Greater London was created in 1965. This affected a large number of authorities in the London area. Greater London was composed of parts of Essex, Hertfordshire, Kent and Surrey together with Middlesex and London.

5.1.4. dGOR/dgorno

Following a request from potential users for a derived variable which allowed comparison over time was attached where possible which gives approximated Government Office Region (or GOR, as used in 2001).

GOR is approximated in the sense that the classification is based on the best fit of each contemporary local authority to 2001 GOR using a look up kindly provided by Humphrey Southall of the Vision of Britain team (www.visionofbritain.org.uk) which provided best fit 2001 GOR for 1961 areacodes.

Linkage was done at this local authority level so that users who are interested in original contemporary geography will be able to access this either at local authority level in secure files, or grouped local authority level (see below) as well as being able to use a geography at grouped or regional level which offers considerable consistency over time.

This coarser grouping, together with variations in data quality mean that the population sizes of approximated GOR differs a little from Vision of Britain GOR. The extent of this difference is described below.

Best fit GOR for district (2001)	EEHCM 1961 9pct	Vision of Britain 61	1961 100% estimate	% Difference
East	339,039	3754784	3767100	-0.328
East Midlands	303,276	3325518	3369733	-1.330
London	705,324	7975558	7836933	1.738
North East	237,699	2649896	2641100	0.332
North West	622,168	6898871	6912978	-0.204
South East	518,011	5728416	5755678	-0.476
South West	329,998	3693029	3666644	0.714

Wales	238,666	2644023	2651844	-0.296
West Midlands	429,098	4756450	4767756	-0.238
Yorks & Humberside	422,698	4679582	4696644	-0.365
Scotland	468,490	5179344	5205444	-0.504
Total	4,614,467	51285471	51271856	0.027

Vision of Britain 61 values are taken from GOR total population 1801-2011 population statistics pages from A Vision of Britain website: www.visionofbritain.org.uk

% difference is calculated as the difference between the sample count divided by 0.09 and the population estimate given on Vision of Britain divided by the Vision of Britain estimate.

Grouped local authorities nest within GOR region and these groupings were created with consistent codes for **1971** and the groupings then used to generate estimated GOR for 1971.

5.1.5. Grouped local authority

In order to protect respondent confidentiality it is not possible to identify areas with populations below one hundred thousand unless data are in a secure environment. Accordingly, it has been necessary to group some local authorities together for the safeguarded files, this has been done to provide as much consistency with 1991 SAR areas as possible as these geographical groupings have been used in both 1991 and 2011.

Lists of the local authorities in each grouped area are listed in the appropriate spreadsheet file.

For **1981** this information is contained in the EEHCM CMGeography.xls file which forms part of the documentation. The equivalent file for **1961** is called EEHCM-CM61Geography.xlsx.

The ability to create consistent groupings has been limited by the need to retain as much authentic information as possible to enable comparison with published figures and other contemporary sources. It is not possible to present competing geographies together in safeguarded or open government files, as, f used together, they provide more information than is allowable for the data type as boundary mismatches produce 'slivers' which are areas too small to be identified. Users who interested in using the contemporary local authority local district or county boundaries are advised to explore the possibility of using the secure files which do contain this detail.

Local authorities in **1981** and 1991 were very similar. Local authorities prior to 1981 precede the 1970s local government reorganisations, and the process was more complex. Each authority was first associated with a best fit 1991 geography. Information used to generate these best fits was as follows;

- A look up table provided by Office for National Statistics listed all authorities in England and Wales in 1961 and the authorities that they went into after reorganisation. If a 1961 authority moved wholly into a 1991 authority this was the best fit.
- Equivalent lookups for Scotland were generated from the local government acts and Scotland's places website.
- Where local authorities were split information about the split was obtained from the relevant statue or F. A Youngs. (1979) A Guide to the Administrative Units of England, (Royal Historical Society) London.
- Where a local authority was split with named wards going into one authority and named wards going into another, the population size of those wards at latest census before reorganisation

was considered and the best fit was taken as the largest population drawing on population figures from CasWeb or Vision of Britain.

- Where a local authority was split on the basis of parishes, the most recent pre-organisation population figures available from Vision of Britain.
- In two or three instances no standard boundary information was given but rather the split was described in terms of areas bounded by a road, in this case a decision was made by eyeballing a contemporary map with sufficient detail to see buildings.

These areas were grouped by their best fit 1991 authorities into groupings used in 1991 SAR Areas. In some areas, particularly areas where population had increased greatly over the period 1961 – 1991 or where data had not been recovered some additional grouping was necessary.

Accordingly grouping has been done on a population best fit basis, limited by region and contemporary county. Users should note that the resulting geographies are combinations of contemporary authorities. These may not fully map on to 1991 or 2011 grouped authorities however they are best approximations.

Users who wish to link aggregate data to microdata are advised to use the local authority or grouped local authority files.

County is not available in safeguarded or open data as many counties did not meet the population threshold for release.

5.2. Occupational Classifications

Occupational classifications tend to be revised on a decennial basis, accordingly each census dataset is based on a different occupational classification. This necessarily reflects considerable changes to the occupational structure since 1961.

Table 5: Occupational classification 1961-81

Year	Classification/Reference	Notes
1961	General Register Office (1960) **Classification of Occupations 1960**. London (HMSO)	No digital census outputs are available for 1951, consistency with previous versions is therefore not considered.
1971	Office of Population Censuses and Surveys (1970) **Classification of Occupations 1970** London (HMSO)	This classification is closely comparable with those in the previous 2 censuses, however certain groups have been expanded to form subgroups, while boiler scalers, chimney sweeps, company directors and coopers, hoopmakers and benders groups were abolished[2].
1981	Office of Population Censuses and Surveys (1980) **Classification of occupations 1980.** London (HMSO)	While references have been found to a good degree of continuity being maintained between 1970 and 1980 classifications documentation to support the claim has not been located at the time of writing.

[2] Census 1971 England and Wales, General Report, Part 1 Defninitions (1979) HMSO, London, p29

5.3. Social class

Social class was introduced into official statistics in 1913 and was first used in the census results in a form we would recognise as the occupation based Registrar General's Social Class in 1921 (Rose 1995). It has been the subject of much criticism since it was first introduced in 1921 as it lacks a clear rationale. However, it remains one of the more consistent variables across the time period despite regular reclassifications of individual occupations.

The most notable discontinuity in the classification came in 1981 when the classification was recast as an indicator of occupational skill rather than occupation per se (Rose 1995). While this could have constituted a major conceptual shift only 7% of cases were reclassified.

In **1961** it is possible to distinguish between manual, non-manual and agricultural occupations as the Classification volume (General Register Office 1960) lists a manual, non-manual or agricultural status for each occupational class. This information has been retained for those individuals in Class III to distinguish between III-N and III-M to facilitate comparisons with the later version of the classification. It should be noted that Class III agricultural jobs have been combined with Class III manual jobs for this process.

5.4. Family units and household relationships

Each secure and household safeguarded file contains a household ID variable. With the exception of 1961 household files also contain a family unit identifier. Household files contain samples of entire households so that household members can be linked to each other in order to facilitate additional analysis of household characteristics and intra-household patterns.

Household composition analysis was undertaken in broadly the same manner in 1961 and 1971. The Household Classification Analysis algorithm resulted in a household relationship variable that describes the relationship *of the family unit* to the head of household. This approach is different from the equivalent variables in microdata produced in 1991 or later which indicates the relationship *of an individual* respondent to the head of household or household reference person. This approach does not have quite the flexibility that an individual approach does as the relationship of individual respondents to the head is not always clear. This has limited the scope of the EEHCM team to generate additional household classifications.

5.4.1. Chief Economic Supporter

This concept is used in some years and might be considered analogous to the current concept of a Household Representative person. In **1971** the Chief Economic Supporter (CES for short) was selected from members of the household who were 15 years of age or over, who either the head of the household or related to the head of household, using the rules described below, drawn from the Census Definitions volume for England and Wales (OPCS 1979: 23);

a) Employment status is considered first. Those in full-time employment (that is who worked for more than 30 hours in the week before the census) or out of employment were selected before those in part-time employment, who in turn were selected before any others.
b) Among those selected by rule (a) above, position in family was considered next, married men or widowed in divorced persons families being considered before other members of families or persons not in families.
c) Among those selected by rules (a) and (b), sex was considered next, males being selected before females.
d) Among those selected by rules (a), (b) and (c), age was considered next, older persons being selected before younger.

If these rules finally select two or more persons, the person whose name appears first on the form was selected as CES. In those rare cases where the head of household and all related persons were aged under 15 the head was selected as CES.

5.4.2. Head of household

Because in earlier censuses household relationship is recorded on a family, rather than individual, basis it has been difficult or impossible to identify the head of household in some datasets.

For example, for **1971**, the Head of household was taken to be the person reported as head of household in the form unless that person was

a) under 15 years;
b) not usually resident at the address or enumeration, or;
c) a resident domestic servant of the household, or a member of such a domestic servant's family;

in which case the first stated resident member over 15 years of age who was not a domestic servant was taken to be head. If this was not possible an under 15 year old could be taken as head as a last resort, however a domestic servant was not taken as head.

From ONS (1979: 22)

While this definition is given, head of household is not recorded in the recovered data in 1971 nor are variables relating to head of household given. This further limits the utility of the family relationship to the head of household variable.

Published household composition tables in 1971 do not use the Head of household concept, using CES instead.

In 1961, Although the concept of the Chief Economic Supporter was used in a tabulated output (e.g. Table 39 of the GB Summary tables) this was not refected in the dataset.

Similarly, Head of household is defined in the Summary Tables as follows;

"This is the person so described on the census schedule by the householder making the return. Where the head was present on census night, he/she should have normally made the return by placing his/her name at the head of column A. When the head was absent, the acting head should have indicated the identity of the head in the list of absent persons in Part III of the census schedule."

However, head of household was not in the source data used to draw the Census microdata samples. It seems likely, given the definition above that persno (in the secure file) and dvepersnoh=1 (in the safeguarded household file) might constitute a proxy for identifying a head of household. Initial exploration of this suggests that people who are person number 1 are disproportionately married men.

5.5: Data enhancements - limits

The data have been exposed to extensive recoding and enhancement. Virtually all variables which have been manipulated by the team start with the letters dv or simply d. In many cases this manipulation has been simply to reduce detail to a level which is acceptable for the data access route with detail being much more extensive in secure files than safeguarded files, and detail in open files being highly limited.

While it has been possible to add some derived variables the team's ability to enhance the data has been necessarily limited due to data availability, resourcing and appropriateness. Questions that we take for granted in 2011 were not necessarily asked 50 years ago, where they were, they may have been coded and stored in a manner which do not lend themselves to repurposing to current requirements.

For example;

Carstairs before 1981 While it was possible to add a measure which worked like Carstairs to the 1981 data, the necessary data were not available for 1971 and earlier due to differences in the output design, the availability of digital data and in the definition of population base. It should be noted however that Carstairs as presented in the data was calculated on a UK basis, with percentiles relating to relative positions in the UK. This will produce different results than had the figures been produced for separate countries.

It should be noted, that Carstairs was a measure created and evaluated in the 1980s (see Carstairs and Morris 1991), it is not at all clear that the measure would have been appropriate to earlier periods. In particular it is not clear that we could consider car ownership to be a reliable measure of deprivation when car access was less common.

Distances between place of work/previous residence Euclidean distance calculations require detailed locations for both addresses. Where these detailed locations were not present in the source files it was not possible to calculate distances.

Household derived variables in 1961 data have only very limited household type information as derived variables proved impossible to create with any degree of certainty. This problem is also present to a more limited degree in 1971; however the 1971 data arrived with a large number of household variables already calculated.

These issues been considered alongside the need to ensure that data enhancements do not pose a disclosure risk as described below.

6. Statistical disclosure Control considerations

Confidentiality is of paramount importance when considering any output from the Census, including datasets. While the EEHCM team designed samples, samples have only been released for use where the data owners at the relevant census offices have been satisfied that the resulting files do not constitute personal data. In the case of the secure files, data will not be released, but are rather only available to Approved Researchers in a controlled setting because the data are deemed to be too detailed for release.

The principles involved in producing microdata from Census records had been well rehearsed for contemporary samples (Marsh et. al (1988), Dale and Elliot (2001)). The general principles might be summarised as follows;

- No direct identifiers such as name, address or date of birth are included in the files– indeed these direct identifiers were not available on the source digital records.
- ID numbers are specific to the file. If an individual appears in the secure file and the safeguarded file it is not possible to easily link cases using a common ID. In practice, a user of the secure file would be prevented from doing this due to restrictions on use.

- Easy to access files have much smaller sample fractions. This is a longstanding disclosure technique (e.g. Dale et. al 2000 p27). A low sampling fraction means the chances of any particular person being in the sample is small.
- Details on files were limited to those appropriate to the data access type. In particular geographical detail is limited in easier to access files. For example, local authority is grouped in the Individual safeguarded file and omitted from the open file.
- The amount of detail available on individual variables have also been traded against each other, so for example it was not considered possible to include full detail on both industry and occupation. This approach reflects advice relating to the release of non-census microdata (see Disclosure Control Guidance for Microdata Produced from Social Surveys November 2014 http://www.ons.gov.uk/file?uri=/methodology/methodologytopicsandstatisticalconcepts/disclosurecontrol/policyforsocialsurveymicrodata/gssgsrdisclosurecontrolguidanceformicrodataproducedfromsocialsurveysfinalnovtcm77391500.doc <last accessed 25/09/16>)
- Files which allow household members to be linked together either contain less information on other variables. So for example the safeguarded household file is available only at region level.

6.1. Data ageing and data quality as natural additional protection

In some places it has been possible to provide a little more detail than would be possible from a contemporary census due to the age and quality of the data. The age of the data provides protections which are not typically available to up to date data.

6.1.3. Mortality of form filler

As data age, so do the data subjects and with data which is between 35 and 55 years old many will have died.

In 2014, the team looked at the decennial life tables for 60-62 to 2000-02 to calculate the probability of survival of all heads of household for each of the samples. Heads of households are selected as we would normally expect the head of household to complete the form (indeed in 1961, 66 and 71 the form is addressed to the head of household or acting head).

Probability of survival for each year of age was calculated from decennial life tables for 1960-62 to 2000-2002(ELT 12, ELT 13, ELT 14, ELT 15, ELT 16) and Lx for the period 2001-14 was based on the latest available of these tables. This probability was then multiplied by the number of heads of households at each age and summed to give the number of expected survivors. The exception is 1961 as only quinary age is available on the data, probability of survival was calculated for the average age of each age group.

Table 6: Probability of survival to 2014 for England

Heads of household in	Male (%)	Female (%)
1961	8.6	4.6
1971	24.9	12.3
1981	42.7	28.4

Source: tables ELT12, ELT13, ELT14, ELT15, ELT16

This exercise demonstrates that a high proportion of form fillers will not be alive at the time of data release.

This has three implications. First, the pool of potential intruders drawing on their own past experience is limited. Second, in the great majority of cases surviving data subjects will have been described by someone else, with the result that they are less likely to remember details about filling in the form, and errors arising from proxying may apply (see for example Dawe and Knight 1997). Finally, surviving data subjects of the older censuses are likely to have been children at the time of enumeration who are less likely than the form fillers would have been to remember whether in 1961 they received and E10 (long form) or E90 (short form).

6.1.4. Availability of matching data

A major driver for the re-assessment of statistical disclosure approaches around the time of the 2001 census was a growing concern about the availability in the data environment of potential matching databases. The statement on the confidentiality of 2001 Census microdata noted;

"Protecting the confidentiality of details about individual people becomes more difficult with each Census, as the amount of accessible and publicly available information about individuals increases. More information can now be matched statistically with the Census"

If this is true and newer data is riskier as more matching data is available, by extension, we would expect less contemporaneous data to be available to match on to old data. This conjecture is supported by law. The ability of organisations to keep personal data for long periods is restricted by the data protection acts which state that personal information (which might contain matching keys as well as identifying information) can only be kept for as long as is necessary. Necessity itself is often limited by other law, most commonly the limitations acts. So, for example, many actions under contract or tort are limited to six years in England and Wales (Limitations Act 1980).

6.1.5. Memory loss – forgetting one's own past

"For most events the passage of time results in memory decay"
Beckett et.al. 2001.

Much of the literature on long term recall relates to retrospective studies where authors conclude that useful data can be obtained using retrospective method. However, behind the desire to improve retrospective methods using techniques such as lifegrids, is the simple issue that we forget, a lot.

A useful study of long term memory of living conditions was undertaken by Berney and Blane (1997) who compared retrospective recall against historical records created as part of the a study fifty years previously, which provides a good model for considering census microdata up to fifty three years old.

Table 7: Recall rates based on a small sample of Boyd Orr and Mass Observation study subjects approximately fifty years later

Topic	Measure	Proportion exact/n (%)
Father's occupation	Description	20/28 (71)
Address	Exact address	23/47 (49)
House type	Description	9/16 (56)
No. of rooms in house	Number	11/58 (58)
No. of children living in the house	Number	19/22 (86)
Domestic water facilities inside/outside	Hot/cold and	15/16 (94)
Toilet facilities	Location	17/17 (100)

Source: summarised from Berney and Blane (1997)

6.1.6. Memory loss – forgetting others

When one is considering the potential for spontaneous recognition potential intruders (intentional or not) may draw up on their knowledge of their own or acquaintances' circumstances, so ability to remember details about the people around us is also relevant.

A useful US study by Bahrick et. al (1975) demonstrates the impact of memory on recall of names of individuals from one's past. The study addressed the number of names of classmates groups could remember based on length of time since graduation. They found that those who had graduated in the previous year (n=50) were able to freely recall the names of 47 classmates. By contrast the number of classmates freely remembered by individuals who had graduated a mean 48 years earlier was only 20. This was despite the potential for year books to act as aide memoires.

6.1.7. Churn

Added to the ability to recall information is the known effect of data churn, that is the tendency for data to become outdated as data subjects circumstances change. Information for the early 1990s about the degree of churn in data arising from the dynamic nature of people's lived experience was based on an analysis of the British Household Panel Study (Elliot and Dale 2001: 435ff). They cite Buck et. al. (1994) who discovered that in the first year of the study (1991-1992) 14% of households experienced a change in a composition and that 62% of those in work changed their detailed occupation or 42% changed their occupational group during that year.

With older data, churn not only makes trying to match current identifiable data to the old anonymised data impossible. It also means that it much harder to accurately remember one's circumstances on the night of the census so long ago. In other words, trying to remember changing circumstances is like trying to shoot at a moving target.

It also suggests that there is some protection in precision; the more precise the classifications, the less likely it is that a potential intruder will be able to get a precise match as more detailed data is less stable.

7. Documentation

A codebook has been generated from the data to describe the range of values stored for each variable and the codes associated with those values. This document is specific to *each individual file* and is the authoritative guide of the content of the file.

A glossary is presented in spreadsheet format for *each year*. This spreadsheet contains notes about the source of the variable. The following information is provided:

- whether it has been drawn directly from the existing digital data, or whether the variable was derived by the recovery team
- The variable name, label and a description
- Relevant notes, including those relate to problems in interpreting the variable or data quality issues
- The applicability of the variable (ie. The universe to which it applies).

Each year contains a benchmarking file providing some prima facie evidence about how the 10% sample datasets used as source data for the files compare with the 100% published tables and a spreadsheet which contains information about the grouping of local authorities.

Additionally the 1961 data contain an abstract from the Census 1961 Great Britain Summary which gives explanatory notes for many of the terms used.

8. References

Bahrick H.P. , Bahrick, P.O and Wittlinger, R.P. (1975) Fifty Years of Memory for Names and Faces: A Cross-Sectional Approach **Journal of Experimental Psychology: General Vol 104 No. 1 54-75**

Berney, L. R. and D.B. Blane (1997) Collecting Retrospective Data: Accuracy of Recall after 50 years judged against historical records **Soc Sci Med 45 (10)** 1519-1525

Britton M and Birch F. 1981 **Census Post-Enumeration Survey, An enquiry into the coverage and quality of the 1981Census in England and Wales**. HMSO (London 1985).

Buck, N. Gershuny, J., Rose, D. and Scott, J. (1994**) Changing Households, the British Household Panel Study**, 1990- 1992. Colchester: Centre for Micro-social Change.

Carstairs, V. and Morris, R (1991**) Deprivation and Health in Scotland** Aberdeen (Aberdeen University Press)

Dale, A. and Elliot, M (2001) Proposals for 2001 samples of anonymized records: An assessment of disclosure risk **Journal of the Royal Statistical Society Series A (Statistics in Society)** 164(3):427-447 · February 2001

Dale, A., Fieldhouse, E. and Holdsworth, C (2000**) Analyzing Census Microdata** London (Arnold)

Dawe. F. and I Knight (1997) "A study of proxy response in the Labour Force Survey" **Survey Methodology Bulletin (No.40), January 1997.**

Hattersley, and R. Creeser (1995) **Longitudinal Study 1971 - 1991: History, organisation and quality of data** HMSO (London)

Catherine Marsh, Chris Skinner, Sara Arber, Bruce Penhale, Stan Openshaw, John Hobcraft, Denise Lievesley and Nigel Walford (1991) 'The case for samples of anonymised records from the 1991 census' **Journal of the Royal Statistical Society. Series A (Statistics in Society)** , Vol. 154, No. 2 (1991), pp. 305-340

Office of Population Censuses and Surveys (OPCS) (1979) **Census 1971 England and Wales, General Report, Part 1 Definitions** HMSO (London)

Office of Population Censuses and Surveys (1983) **Census 1971: General Report Par 3 Statistical Assessment** HMSO (London)

Rose, D (1995) 'Official Social Classifications in the UK' **Social Research Update Issue 9** July 1995, Sociology at Surrey, available from sru.soc.surrey.ac.uk/SRU9.html <last accessed 05/06/2017>

Youngs, F (1979) **A Guide to the Administrative Areas of England** (vols 1 and 2) Royal Historical Society (London)

Appendix A: Abstract from the Classification of Occupations 1980

Summary of the OPCS 1980 occupation groups with condensed KOS headings

Order 1 Professional and related supporting management; senior national and local government managers

001 Judges, barristers, advocates, solicitors
- 0 Judges, barristers, advocates, solicitors

002 Accountants, valuers, finance specialists
- 1 Chartered and certified accountants
- 2 Cost and works accountants
- 3 Estimators
- 4 Valuers, claims assessors
- 5 Financial managers
- 6 Underwriters, brokers, investment analysts
- 7 Taxation experts

003 Personnel and industrial relations managers; O and M, work study and operational research officers
- 1 Personnel and industrial relations officers
- 2 O and M, work study and OR officers

004 Economists, statisticians, systems analysts, computer programmers
- 1 Economists, statisticians, actuaries
- 2 Systems analysts, computer programmers

005 Marketing, sales, advertising, public relations and purchasing managers
- 1 Marketing and sales managers and executives
- 2 Advertising and PR executives
- 3 Buyers (retail trade)
- 4 Buyers and purchasing officers (not retail)

006 Statutory and other inspectors
- 1 Environmental health officers
- 2 Building inspectors
- 3 Inspectors (statutory and similar)

007 General administrators—national government
- 1 General administrators—national government (Assistant Secretary level and above)
- 2 General administrators—national government (HEO to Senior Principal level)

008 Local government officers (administrative and executive functions)
- 0 Local government officers (administrative and executive functions)

009 All other professional and related supporting management and administration
- 1 Company secretaries
- 2 Officials of trade associations, trade unions, professional bodies and charities
- 3 Property and estate managers
- 4 Librarians, information officers
- 5 Legal service and related occupations
- 6 Management consultants
- 7 Managers' personal assistants
- 8 Professional workers and related supporting management and administration n.e.c.

Note: n.e.c. = not elsewhere classified

Order 2 Professional and related in education, welfare and health

010 Teachers in higher education
- 1 University academic staff
- 2 Teachers in establishments for further and higher education

011 Teachers n.e.c.
- 0 Teachers n.e.c.

012 Vocational and industrial trainers, education officers, social and behavioural scientists
- 1 Vocational and industrial trainers
- 2 Education officers, school inspectors
- 3 Social and behavioural scientists

013 Welfare workers
- 1 Matrons, houseparents
- 2 Playgroup leaders
- 3 Welfare occupations n.e.c.

014 Clergy, ministers of religion
- 0 Clergy, ministers of religion

015 Medical and dental practitioners
- 1 Medical practitioners
- 2 Dental practitioners

016 Nurse administrators, nurses
- 0 Nurse administrators, nurses

017 Pharmacists, radiographers, therapists n.e.c.
- 1 Pharmacists
- 2 Medical radiographers
- 3 Ophthalmic and dispensing opticians
- 4 Physiotherapists
- 5 Chiropodists
- 6 Therapists n.e.c.

018 All other professional and related in education, welfare and health
- 1 Medical technicians, dental auxiliaries
- 2 Veterinarians
- 3 Driving instructors (not HGV)
- 4 Professional and related in education, welfare and health n.e.c.

Order 3 Literary, artistic and sports

019 Authors, writers, journalists
- 0 Authors, writers, journalists

020 Artists, designers, window dressers
- 1 Artists, commercial artists
- 2 Industrial designers (not clothing)
- 3 Clothing designers
- 4 Window dressers

021 Actors, musicians, entertainers, stage managers
- 1 Actors, entertainers, singers, stage managers
- 2 Musicians

022 Photographers, cameramen, sound and vision equipment operators
- 1 Photographers, cameramen
- 2 Sound and vision equipment operators

Note: n.e.c. = not elsewhere classified

023 All other literary, artistic and sports
1 Professional sportsmen, sports officials
2 Literary, artistic and sports workers n.e.c.

Order 4 Professional and related in science, engineering, technology and similar fields

024 Scientists, physicists, mathematicians
1 Biological scientists, biochemists
2 Chemical scientists
3 Physical and geological scientists, mathematicians

025 Civil, structural, municipal, mining and quarrying engineers
0 Civil, structural, municipal, mining and quarrying engineers

026 Mechanical and aeronautical engineers
1 Mechanical and aeronautical engineers
2 Design and development engineers (mechanical)

027 Electrical and electronic engineers
1 Electrical engineers
2 Electronic engineers

028 Engineers and technologists n.e.c.
1 Chemical engineers
2 Production engineers
3 Planning and quality control engineers
4 Engineers n.e.c.
5 Metallurgists
6 Technologists n.e.c.

029 Draughtsmen
0 Draughtsmen

030 Laboratory and engineering technicians, technician engineers
1 Laboratory technicians
2 Engineering technicians, technician engineers

031 Architects, town planners, quantity, building and land surveyors
1 Architects, town planners
2 Quantity surveyors
3 Building, land and mining surveyors

032 Officers (ships and aircraft), air traffic planners and controllers
1 Aircraft flight deck officers
2 Air traffic planners and controllers
3 Deck, engineering and radio officers and pilots, ship

033 Professional and related in science, engineering and other technologies and similar fields n.e.c.
1 Architectural and town planning technicians
2 Building and civil engineering technicians
3 Technical and related workers n.e.c.

Order 5 Managerial

034 Production, works and maintenance managers, works foremen
0 Production, works and maintenance managers, works foremen

035 Site and other managers, agents and clerks of works, general foremen (building and civil engineering)
1 Managers in building and contracting
2 Clerks of works

Note: n.e.c. = not elsewhere classified

036 Managers in transport, warehousing, public utilities and mining
- 1 Managers in mining and public utilities
- 2 Transport managers
- 3 Stores controllers
- 4 Managers in warehousing and materials handling n.e.c.

037 Office managers
- 1 Credit controllers
- 2 Office managers n.e.c.

038 Managers in wholesale and retail distribution
- 1 Garage proprietors
- 2 Butchers (managers and proprietors)
- 3 Fishmongers (managers and proprietors)
- 4 Other proprietors and managers (sales)

039 Managers of hotels, clubs, etc., and in entertainment and sport
- 1 Hotel and residential club managers
- 2 Publicans
- 3 Restaurateurs
- 4 Club stewards
- 5 Entertainment and sports managers

040 Farmers, horticulturists, farm managers
- 0 Farmers, horticulturists, farm managers

041 Officers, UK armed forces
- 0 Officers, UK armed forces

042 Officers, foreign and Commonwealth armed forces
- 0 Officers, foreign and Commonwealth armed forces

043 Senior police, prison and fire service officers
- 1 Prison officers (chief officers and above)
- 2 Police officers (inspectors and above)
- 3 Fire service officers

044 All other managers
- 1 Proprietors and managers, service flats, holiday flats, caravan sites, etc.
- 2 Managers of laundry and dry cleaning receiving shops
- 3 Hairdressers' and barbers' managers and proprietors
- 4 Managers n.e.c.

Order 6 Clerical and related

045 Supervisors of clerks, civil service executive officers
- 1 Civil service executive officers

Supervisors of—
- 2 Stores and despatch clerks
- 3 Tracers, drawing office assistants
- 4 Other clerks and cashiers (not retail)
- 5 Retail shop cashiers, check-out and cash and wrap operators

046 Clerks
- 1 Stores and despatch clerks
- 2 Tracers, drawing office assistants
- 3 Other clerks and cashiers (not retail)

047 Retail shop cashiers, check-out and cash and wrap operators
- 0 Retail shop cashiers, check-out and cash and wrap operators

Note: n.e.c. = not elsewhere classified

048 Supervisors of typists, office machine operators, telephonists, etc.
Supervisors of—
1 Typists, shorthand writers, secretaries
2 Office machine operators
3 Telephone operators
4 Radio and telegraph operators

049 Secretaries, shorthand typists, receptionists
1 Receptionists
2 Typists, shorthand writers, secretaries

050 Office machine operators
0 Office machine operators

051 Telephonists, radio and telegraph operators
1 Telephonist receptionists
2 Telephone operators
3 Radio and telegraph operators

052 Supervisors of postmen, mail sorters, messengers
Supervisors of—
1 Postmen, mail sorters
2 Messengers

053 Postmen, mail sorters, messengers
1 Postmen, mail sorters
2 Messengers

Order 7 Selling

054 Sales supervisors
Supervisors of—
1 Shop salesmen and assistants
2 Petrol pump, forecourt attendants
3 Roundsmen, van salesmen

055 Salesmen, sales assistants, shop assistants, shelf fillers, petrol pump, forecourt attendants
1 Shop salesmen and assistants
2 Shelf fillers
3 Petrol pump, forecourt attendants

056 Roundsmen, van salesmen
0 Roundsmen, van salesmen

057 Sales representatives and agents
1 Importers, exporters, commodity brokers
2 Market and street traders and assistants
3 Scrap dealers, general dealers, rag and bone merchants
4 Credit agents, collector salesmen
5 Sales representatives
6 Sales representatives (property and services), other agents

Order 8 Security and protective service

058 NCOs and other ranks, UK armed forces
0 NCOs and other ranks, UK armed forces

059 NCOs and other ranks, foreign and Commonwealth armed forces
0 NCOs and other ranks, foreign and Commonwealth armed forces

Note: n.e.c. = not elsewhere classified

060 Supervisors (police sergeants, fire fighting and related)
- 1 Police sergeants
- 2 Fire service supervisors
- 3 Prison service principal officers

Supervisors of—
- 4 Security guards and officers, patrolmen, watchmen
- 5 Traffic wardens
- 6 Security and protective service workers n.e.c.

061 Policemen, firemen, prison officers
- 1 Policemen (below sergeant)
- 2 Firemen
- 3 Prison officers (below principal officer)

062 Other security and protective service workers
- 1 Security guards and officers, patrolmen, watchmen
- 2 Traffic wardens
- 3 Security and protective service workers n.e.c.

Order 9 Catering, cleaning, hairdressing and other personal service

063 Catering supervisors

Supervisors of—
- 1 Chefs, cooks
- 2 Waiters, waitresses
- 3 Barmen, barmaids
- 4 Counter hands, assistants

064 Chefs, cooks
- 0 Chefs, cooks

065 Waiters and bar staff
- 1 Waiters, waitresses
- 2 Barmen, barmaids

066 Counter hands, assistants, kitchen porters, hands
- 1 Counter hands, assistants
- 2 Kitchen porters, hands

067 Supervisors—housekeeping and related
- 1 Housekeepers (non-domestic)

Supervisors of—
- 2 Other domestic and school helpers
- 3 Travel stewards and attendants
- 4 Hospital porters
- 5 Hotel porters
- 6 Ambulancemen
- 7 Hospital, ward orderlies

068 Domestic staff and school helpers
- 1 Domestic housekeepers
- 2 Nursery nurses
- 3 Other domestic and school helpers

069 Travel stewards and attendants, hospital and hotel porters
- 1 Travel stewards and attendants
- 2 Hospital porters
- 3 Hotel porters

Note: n.e.c. = not elsewhere classified

070 Ambulancemen, hospital orderlies
1 Ambulancemen
2 Hospital, ward orderlies

071 Supervisors, foremen—caretaking, cleaning and related
Supervisors of—
1 Caretakers
2 Cleaners, window cleaners, chimney sweeps, road sweepers
3 Railway stationmen
4 Lift and car park attendants

072 Caretakers, road sweepers and other cleaners
1 Caretakers
2 Cleaners, window cleaners, chimney sweeps, road sweepers

073 Hairdressing supervisors
0 Hairdressing supervisors

074 Hairdressers, barbers
0 Hairdressers, barbers

075 All other in catering, cleaning and other personal service
1 Railway stationmen
2 Lift and car park attendants
3 Launderers, dry cleaners, pressers
4 Undertakers
5 Bookmakers, betting shop managers
6 Service workers n.e.c.

Order 10 Farming, fishing and related

076 Foremen—farming, horticulture, forestry
1 Farm foremen
2 Horticultural foremen
3 Foremen gardeners and groundsmen
4 Agricultural machinery foremen
5 Forestry foremen
6 Other foremen in farming and related

077 Farm workers
0 Farm workers

078 Horticultural workers, gardeners, groundsmen
1 Horticultural workers
2 Gardeners, groundsmen

079 Agricultural machinery drivers, operators
0 Agricultural machinery drivers, operators

080 Forestry workers
0 Forestry workers

081 Supervisors, mates—fishing
0 Supervisors, mates—fishing

082 Fishermen
0 Fishermen

083 All other in farming and related
0 All other in farming and related

Note: n.e.c. = not elsewhere classified

Order 11 Materials processing; making and repairing (excluding metal and electrical)

084 Foremen—tannery and leather (including leather substitutes) working
Foremen—
1 Tannery production workers
2 Shoe repairers
3 Leather cutters and sewers, footwear lasters, makers, finishers
4 Other making and repairing, leather

085 Tannery and leather (including leather substitutes) workers
1 Tannery production workers
2 Shoe repairers
3 Leather cutters and sewers, footwear lasters, makers, finishers

086 Foremen—textile processing
Foremen—
1 Preparatory fibre processors
2 Spinners, doublers, twisters
3 Winders, reelers
4 Warp preparers
5 Weavers
6 Knitters
7 Bleachers, dyers, finishers
8 Menders, darners
9 Other material processing, textiles

087 Textile workers
1 Preparatory fibre processors
2 Spinners, doublers, twisters
3 Winders, reelers
4 Warp preparers
5 Weavers
6 Knitters
7 Bleachers, dyers, finishers
8 Menders, darners

088 Foremen—chemical processing
0 Foremen—chemical processing

089 Chemical, gas and petroleum process plant operators
0 Chemical, gas and petroleum process plant operators

090 Foremen—food and drink processing
Foremen—
1 Bakers, flour confectioners
2 Butchers, meat cutters
3 Fishmongers, poultry dressers
4 Brewery and vinery process workers
Other material processing—
5 Bakery and confectionery workers
6 Food and drink n.e.c.

091 Bakers, flour confectioners
0 Bakers, flour confectioners

092 Butchers
1 Butchers, meat cutters
2 Fishmongers, poultry dressers

Note: n.e.c. = not elsewhere classified

093 Foremen—paper and board making and paper products

Foremen—

1 Paper, paperboard and leatherboard workers
2 Bookbinders and finishers
3 Cutting and slitting machine operators (paper and paper products making)
4 Other material processing, wood and paper
5 Other making and repairing, paper goods and printing

094 Paper, board and paper product makers, bookbinders

1 Paper, paperboard and leatherboard workers
2 Bookbinders and finishers
3 Cutting and slitting machine operators (paper and paper products making)

095 Foremen—glass, ceramics, rubber, plastics, etc.

Foremen—

1 Glass and ceramics furnacemen, kilnsetters
2 Glass formers and shapers, finishers, decorators
3 Casters and other pottery makers
4 Rubber process workers, moulding machine operators, tyre builders
5 Calendar and extruding machine operators, moulders (plastics)
6 Man-made fibre makers
7 Washers, screeners and crushers in mines and quarries

Other making and repairing—

8 Glass and ceramics
9 Rubber
10 Plastics

096 Glass and ceramics furnacemen and workers

1 Glass and ceramics furnacemen, kilnsetters
2 Glass formers and shapers, finishers, decorators
3 Casters and other pottery makers

097 Rubber and plastics workers

1 Rubber process workers, moulding machine operators, tyre builders
2 Calendar and extruding machine operators, moulders (plastics)

098 All other in processing materials (other than metal)

1 Man-made fibre makers
2 Brewery and vinery process workers
3 Washers, screeners and crushers in mines and quarries

Other material processing—

4 Textiles
5 Bakery and confectionery workers
6 Tobacco
7 Food and drink n.e.c.
8 Wood and paper
9 All other (excluding metal) n.e.c.

099 Foremen—printing

Foremen—

1 Compositors
2 Electrotypers, stereotypers, printing plate and cylinder preparers
3 Printing machine minders and assistants
4 Screen and block printers
5 Printers (so described)

Note: n.e.c. = not elsewhere classified

100 Printing workers, screen and block printers
- 1 Compositors
- 2 Electrotypers, stereotypers, printing plate and cylinder preparers
- 3 Printing machine minders and assistants
- 4 Screen and block printers
- 5 Printers (so described)

101 Foremen—textile materials working
Foremen—
- 1 Tailors, tailoresses, dressmakers
- 2 Clothing cutters, milliners, furriers
- 3 Sewers, embroiderers
- 4 Coach trimmers, upholsterers, mattress makers
- 5 Carpet fitters
- 6 Other making and repairing, clothing and related products

102 Tailors, dressmakers and other clothing workers
- 1 Tailors, tailoresses, dressmakers
- 2 Clothing cutters, milliners, furriers
- 3 Sewers, embroiderers

103 Coach trimmers, upholsterers, mattress makers
- 0 Coach trimmers, upholsterers, mattress makers

104 Foremen—woodworking
Foremen—
- 1 Carpenters, joiners
- 2 Cabinet makers
- 3 Case and box makers
- 4 Pattern makers (moulds)
- 5 Sawyers, veneer cutters, woodworking machinists
- 6 Other making and repairing, wood

105 Woodworkers, pattern makers
- 1 Carpenters, joiners
- 2 Cabinet makers
- 3 Case and box makers
- 4 Pattern makers (moulds)

106 Sawyers, veneer cutters, woodworking machinists
- 0 Sawyers, veneer cutters, woodworking machinists

107 All other in making and repairing (excluding metal and electrical)
- 1 Labourers and mates to woodworking craftsmen
- 2 Dental technicians
- 3 Carpet fitters
- 4 Musical instrument makers, piano tuners

Other making and repairing—
- 5 Glass and ceramics
- 6 Wood
- 7 Leather
- 8 Clothing and related products
- 9 Paper goods and printing
- 10 Rubber
- 11 Plastics
- 12 All other (excluding metal and electrical) n.e.c.

Note: n.e.c. = not elsewhere classified

Order 12 Processing, making, repairing and related (metal and electrical)

108 Foremen—metal making and treating
Foremen—
1 Furnace operating occupations (metal)
2 Rollermen
3 Smiths, forgemen
4 Metal drawers
5 Moulders, coremakers, die casters
6 Electroplaters
7 Annealers, hardeners, temperers (metal)
8 Galvanizers, tin platers, dip platers
9 Metal making and treating workers n.e.c.

109 Furnacemen (metal), rollermen, smiths, forgemen
1 Furnace operating occupations (metal)
2 Rollermen
3 Smiths, forgemen

110 Metal drawers, moulders, die casters, electroplaters, annealers
1 Metal drawers
2 Moulders, coremakers, die casters
3 Electroplaters
4 Annealers, hardeners, temperers (metal)

111 Foremen—engineering machining
Foremen—
1 Press and machine tool setters
2 Centre lathe turners
3 Machine tool setter operators
4 Machine tool operators
5 Press, stamping and automatic machine operators
6 Metal polishers
7 Fettlers, dressers
8 Shot blasters

112 Press and machine tool setter operators and operators, turners
1 Press and machine tool setters
2 Centre lathe turners
3 Machine tool setter operators
4 Machine tool operators

113 Machine attendants, minders, press and stamping machine operators, metal polishers, fettlers, dressers
1 Press, stamping and automatic machine operators
2 Metal polishers
3 Fettlers, dressers

114 Foremen—production fitting (metal)
Foremen—
1 Tool makers, tool fitters, markers-out
2 Precision instrument makers and repairers
3 Watch and chronometer makers and repairers
4 Metal working production fitters and fitter/machinists
5 Motor mechanics, auto engineers
6 Maintenance fitters (aircraft engines)
7 Office machinery mechanics

115 Tool makers, tool fitters, markers-out
0 Tool makers, tool fitters, markers-out

Note: n.e.c. = not elsewhere classified

116 Instrument and watch and clock makers and repairers
1 Precision instrument makers and repairers
2 Watch and chronometer makers and repairers

117 Metal working production fitters and fitter/machinists
0 Metal working production fitters and fitter/machinists

118 Motor vehicle and aircraft mechanics
1 Motor mechanics, auto engineers
2 Maintenance fitters (aircraft engines)

119 Office machinery mechanics
0 Office machinery mechanics

120 Foremen—production fitting and wiring (electrical)
Foremen—
1 Production fitters (electrical, electronic)
2 Electricians, electrical maintenance fitters
3 Plant operators and attendants n.e.c.
4 Telephone fitters
5 Cable jointers, linesmen
6 Radio and TV mechanics
7 Other electronic maintenance engineers

121 Production fitters, electricians, electricity power plant operators, switchboard attendants
1 Production fitters (electrical, electronic)
2 Electricians, electrical maintenance fitters
3 Electrical engineers (so described)
4 Plant operators and attendants n.e.c.

122 Telephone fitters, cable jointers, linesmen
1 Telephone fitters
2 Cable jointers, linesmen

123 Radio, TV and other electronic maintenance fitters and mechanics
1 Radio and TV mechanics
2 Other electronic maintenance engineers

124 Foremen—metal working, pipes, sheets, structures
Foremen—
1 Plumbers, heating and ventilating fitters, gas fitters
2 Sheet metal workers
3 Metal plate workers, shipwrights, riveters
4 Steel erectors, benders, fixers
5 Scaffolders, stagers
6 Welders
7 Riggers

125 Plumbers, heating and ventilating fitters, gas fitters
0 Plumbers, heating and ventilating fitters, gas fitters

126 Sheet metal workers, platers, shipwrights, riveters, etc.
1 Sheet metal workers
2 Metal plate workers, shipwrights, riveters

127 Steel erectors, scaffolders, steel benders, fixers
1 Steel erectors, benders, fixers
2 Scaffolders, stagers

Note: n.e.c. = not elsewhere classified

128 Welders
0 Welders

129 Foremen—other processing, making and repairing (metal and electrical)
Foremen—
1 Goldsmiths, silversmiths, precious stone workers
2 Engravers, etchers (printing)
3 Coach and vehicle body builders
4 Oilers, greasers, lubricators
5 Electronics wiremen
6 Coil winders

130 Goldsmiths, silversmiths, etc., engravers, etchers
1 Goldsmiths, silversmiths, precious stone workers
2 Engravers, etchers (printing)

131 All other in processing, making and repairing (metal and electrical)
1 Coach and vehicle body builders
2 Galvanizers, tin platers, dip platers
3 Metal making and treating workers n.e.c.
4 Oilers, greasers, lubricators
5 Riggers
6 Electronics wiremen
7 Coil winders
8 Shot blasters
9 Other metal, jewellery, electrical production workers

—
1 Foremen (engineering and allied trades)
2 Trainee craftsmen (engineering and allied trades)

Order 13 Painting, repetitive assembling, product inspecting, packaging and related

132 Foremen—painting and similar coating
Foremen—
1 Pottery decorators
2 Coach painters (so described)
3 Other spray painters
4 Painters and decorators n.e.c., french polishers

133 Painters, decorators, french polishers
1 Pottery decorators
2 Coach painters (so described)
3 Other spray painters
4 Painters and decorators n.e.c., french polishers

134 Foremen—product assembling (repetitive)
Foremen assemblers—
1 Electrical, electronic
2 Instruments
3 Vehicles and other metal goods
4 Paper production, processing and printing
5 Plastics goods

135 Repetitive assemblers (metal and electrical goods)
1 Assemblers (electrical, electronic)
2 Instrument assemblers
3 Assemblers (vehicles and other metal goods)

Note: n.e.c. = not elsewhere classified

136 Foremen—product inspection and packaging
Foremen inspectors, viewers, examiners—
1 Metal, electrical goods
2 Textiles
3 Food
4 Rubber goods
5 Plastics goods
6 Woodwork
Foremen—
7 Packers, bottlers, canners, fillers
8 Laboratory assistants
9 Inspectors, sorters in paper production, processing and printing
10 Weighers
11 Graders, sorters, selectors n.e.c.

137 Inspectors, viewers, testers, packers, bottlers, etc.
1 Inspectors, viewers (metal, electrical goods)
2 Packers, bottlers, canners, fillers

138 All other in painting, repetitive assembling, product inspection, packaging and related
1 Laboratory assistants
Inspectors, viewers, examiners—
2 Textiles
3 Food
4 Rubber goods
5 Plastics goods
6 Woodwork
7 Inspectors, sorters in paper production, processing and printing
8 Assemblers in paper production, processing and printing
9 Assemblers (plastics goods)
10 Weighers
11 Graders, sorters, selectors n.e.c.
12 Painting, assembling and related occupations n.e.c.

Order 14 Construction, mining and related not identified elsewhere

139 Foremen—building and civil engineering n.e.c.
Foremen—
1 Bricklayers, tile setters
2 Masons, stone cutters
3 Plasterers
4 Roofers, glaziers
5 Handymen, general building workers
6 Railway lengthmen
7 Road surfacers, concreters
8 Roadmen
9 Paviors, kerb layers
10 Sewage plant attendants
11 Mains and service layers, pipe jointers
12 Construction workers n.e.c.

140 Building and construction workers
1 Bricklayers, tile setters
2 Masons, stone cutters
3 Plasterers
4 Roofers, glaziers
5 Handymen, general building workers
6 Builders (so described)

Note: n.e.c. = not elsewhere classified

141 Concreters, road surfacers, railway lengthmen
- 1 Railway lengthmen
- 2 Road surfacers, concreters
- 3 Roadmen
- 4 Paviors, kerb layers

142 Sewage plant attendants, sewermen (maintenance), mains and service layers, pipe jointers (gas, water, drainage, oil), inspectors (water supply), turncocks
- 1 Sewage plant attendants
- 2 Mains and service layers, pipe jointers

143 Civil engineering labourers, craftsmen's mates and other builders' labourers n.e.c.
- 1 Craftsmen's mates
- 2 Building and civil engineering labourers

144 Foremen/deputies—coalmining
- 0 Foremen/deputies—coalmining

145 Face-trained coalmining workers
- 0 Face-trained coalmining workers

146 All other in construction, mining, quarrying, well drilling and related n.e.c.
- 1 Miners (not coal), quarrymen, well drillers
- 2 Construction workers n.e.c.

Order 15 Transport operating, materials moving and storing and related

147 Foremen—ships, lighters and other vessels
- 0 Foremen—ships, lighters and other vessels

148 Deck, engine-room hands, bargemen, lightermen, boatmen
- 0 Deck, engine-room hands, bargemen, lightermen, boatmen

149 Foremen—rail transport operating
Foremen—
- 1 Railway guards
- 2 Signalmen and crossing keepers, railway
- 3 Shunters, pointsmen
- 4 Other foremen rail transport

150 Rail transport operating staff
- 1 Drivers, motormen, secondmen, railway engines
- 2 Railway guards
- 3 Signalmen and crossing keepers, railway
- 4 Shunters, pointsmen

151 Foremen—road transport operating, bus inspectors
- 1 Bus inspectors

Foremen—
- 2 Drivers of road goods vehicles
- 3 Other foremen road transport

152 Bus, coach, lorry drivers, etc.
- 1 Bus and coach drivers
- 2 Drivers of road goods vehicles
- 3 Other motor drivers

153 Bus conductors, drivers' mates
- 1 Bus conductors
- 2 Drivers' mates

Note: n.e.c. = not elsewhere classified

154 Foremen—civil engineering plant operating, materials handling equipment operating
Foremen—
1 Mechanical plant drivers, operators (earth moving and civil engineering)
2 Crane drivers, operators
3 Fork lift, mechanical truck drivers
4 Slingers

155 Mechanical plant, fork lift, mechanical truck drivers, crane drivers, operators
1 Mechanical plant drivers, operators (earth moving and civil engineering)
2 Crane drivers, operators
3 Fork lift, mechanical truck drivers

156 Foremen—materials moving and storing
Foremen—
1 Storekeepers, warehousemen
2 Stevedores, dockers
3 Goods porters
4 Refuse collectors, dustmen

157 Storekeepers, stevedores, warehouse, market and other goods porters
1 Storekeepers, warehousemen
2 Stevedores, dockers
3 Goods porters
4 Refuse collectors, dustmen

158 All other in transport operating, materials moving and storing and related n.e.c.
1 Slingers
2 Workers in transport operating, materials moving and storing and related n.e.c.

Order 16 Miscellaneous

159 Foremen—miscellaneous
Foremen—
Labourers and unskilled workers n.e.c.—
1 Textiles (not textile goods)
2 Chemicals and allied trades
3 Coke ovens and gas works
4 Glass and ceramics
5 Foundries in engineering and allied trades
6 Engineering and allied trades
7 Coal mines
8 Other
9 Boiler operators

160 General labourers
Labourers and unskilled workers n.e.c.—
1 Textiles (not textile goods)
2 Chemicals and allied trades
3 Coke ovens and gas works
4 Glass and ceramics
5 Foundries in engineering and allied trades
6 Engineering and allied trades
7 Coal mines
8 Other

161 All other in miscellaneous occupations n.e.c.
1 Boiler operators
2 All other in miscellaneous occupations n.e.c.

Order 17 Inadequately described and not stated

—
1 Inadequately described occupations
2 Occupations not stated

Note: n.e.c. = not elsewhere classified

Appendix B: Design Factors and Effects.

To establish the extent of the deviation of the sample design from simple random sampling design effects and factors were obtained in Stata 12 for a number of estimated proportions. Sample design was set using the svyset command with household ID set as the clustering variable and strat set as the stratifying variable. Geographical stratification was implicit and was not captured in this analysis. Analyses below are presented without adjustment for finite population, however sensitivity testing suggested that the fpc had little impact on the results. .

The design effect can be interpreted as the variance of an estimate under the sample design divided by the variance had the estimate been achieved under simple random sampling (srs), the design factor is the square root of the design factor which can be used to understand the impact of the sample design on a standard error.

Under simple random sampling both the design effect and design factors will be 1. A design factor above 1 means that the design is less efficient than simple random sampling; design factors well below 1 are less common and mean that the design is more efficient. Standard errors calculated under an assumption of simple random sampling can in theory be corrected by multiplying a standard error under the assumption of srs by an appropriate design effect. However, the design effects are specific to each estimate.

As we see in indicative design effects and factors below some of the standard errors generated using these data will diverge considerably from the assumptions of srs, with the most dramatic differences being those relating to country of birth. Standard errors for proportion with each country birth drawn on the secure files will be underestimated by factors of between 1.2 and 2.2 if standard errors are calculated under an assumption of srs. These are very similar to the sorts of design effects that have been found in earlier census microdata files.

At the other extreme we see that standard errors for proportion married will be overestimated by approximately two thirds if calculated under the assumption of srs. It would not be advisable to assume that the data have standard errors approximate to those produced under srs.

Little work, apart from the indicative figures, has been to assess design effects. Users who wish to explore these effects further are advised to use household files in which it is possible to explore the effect of households and strata (indicated in the strat variable) on the standard errors.

Design effects for the secure1981 file

Country of Birth (Grouped)	Proportion	Standard Error	Design Effect	Design Factor
UK	0.937	0.0002	2.162	1.471
Ireland	0.011	0.0001	1.535	1.239
New Comm'wlth:Caribbean	0.005	0.0000	1.971	1.404
India	0.007	0.0001	2.466	1.570
Pakistan	0.004	0.0001	3.575	1.891
Bangladesh	0.001	0.0000	4.693	2.166
New Comm'wlth: nec	0.011	0.0001	2.325	1.525
Old Commonwealth	0.003	0.0000	1.556	1.247
European Community (1981)	0.008	0.0000	1.530	1.237
Other	0.014	0.0001	2.121	1.456

Occupational Order	Proportion	Standard Error	Design Effect	Design Factor
Not applicable	0.396	0.0002	0.509	0.714
Missing	0.000	0.0000	1.198	1.094
Professional and related support	0.021	0.0001	1.048	1.023
Professional and related in educ	0.041	0.0001	1.209	1.100
Literary, artistic and sports	0.005	0.0000	1.128	1.062
Professional and related in scie	0.021	0.0001	1.027	1.014
Managerial	0.049	0.0001	1.135	1.065
Clerical and related	0.086	0.0001	1.028	1.014
Selling	0.031	0.0001	1.046	1.023
Security and protective service	0.011	0.0000	1.065	1.032
Catering, cleaning, hairdressing	0.057	0.0001	1.040	1.020
Farming, fishing and related	0.009	0.0000	1.148	1.071
Materials processing: making and	0.041	0.0001	1.093	1.045
Processing making repairing and	0.064	0.0001	0.983	0.991
Painting, repetitive assembling,	0.022	0.0001	1.061	1.030
Construction, mining and related	0.021	0.0001	1.040	1.020
Construction, mining and related	0.037	0.0001	0.991	0.996
Miscellaneous	0.014	0.0001	1.077	1.038
Inadequately described and not s	0.073	0.0001	1.044	1.022

Marital condition	Proportion	Standard Error	Design Effect	Design Factor
Not married	0.502	0.0001	0.410	0.640
Married	0.498	0.0001	0.410	0.640

Owner occupation	Proportion	Standard Error	Design Effect	Design Factor
Owner occupied	0.583	0.0004	3.413	1.848
Not owner occupied	0.417	0.0004	3.413	1.848

The1981 individual files

The individual file does not contain a household number, accordingly it is not possible to use a complex sampling procedure such as Stata's svy commands to adjust for the effect of household clustering. The following outputs were produced as part of the data processing so that some indicative information about the effect of sampling could be provided.

Although the sampling design used prevents the inclusion of whole households in the way that the household file does a large proportion of members of the 10% sample is used. This means that multiple members of households may be included and there some residual clustering effects are

inherited, however the effects are considerably less marked than they are for the household sample which reflects the lower level of clustering when less than two thirds of household members are selected.

Grouped country of Birth	Proportion	Standard Error	Design Effect	Design Factor
UK	0.937	0.0002	1.579	1.257
Ireland	0.011	0.0001	1.244	1.115
New Comm'wlth:Caribbean	0.005	0.0001	1.450	1.204
India	0.007	0.0001	1.748	1.322
Pakistan	0.004	0.0001	2.352	1.534
Bangladesh	0.001	0.0000	3.028	1.740
New Comm'wlth: nec	0.011	0.0001	1.674	1.294
Old Commonwealth	0.003	0.0000	1.273	1.128
European Community (1981)	0.008	0.0001	1.260	1.122
Other	0.014	0.0001	1.565	1.251

Occupational Order	Proportion	Standard Error	Design Effect	Design Factor
Not appliciable	0.396	0.0002	0.590	0.768
Missing	0.000	0.0000	1.123	1.060
Professional and related support	0.021	0.0001	1.017	1.009
Professional and related in educ	0.041	0.0001	1.081	1.040
Literary, artistic and sports	0.005	0.0000	1.058	1.029
Professional and related in scie	0.021	0.0001	1.008	1.004
Managerial	0.049	0.0001	1.049	1.024
Clerical and related	0.086	0.0002	0.993	0.997
Selling	0.031	0.0001	1.016	1.008
Security and protective service	0.011	0.0001	1.031	1.015
Catering, cleaning, hairdressing	0.057	0.0001	1.013	1.006
Farming, fishing and related	0.009	0.0001	1.074	1.036
Materials processing: making and	0.041	0.0001	1.039	1.019
Processing making repairing and	0.064	0.0001	0.986	0.993
Painting, repetetive assembling,	0.022	0.0001	1.028	1.014
Construction, mining and related	0.021	0.0001	1.022	1.011
Transort operating, materials mo	0.037	0.0001	0.992	0.996
Miscellaneous	0.014	0.0001	1.037	1.019
Inadequately described and not s	0.073	0.0002	0.961	0.980

Marital condition	Proportion	Standard Error	Design Effect	Design Factor
Not married	0.501	0.0002	0.501	0.502
Married	0.499	0.0002	0.498	0.499

Owner occupation	Proportion	Standard Error	Design Effect	Design Factor
Owner occupied	0.5829973	0.000451	2.24036	1.49678
Not owner occupied	0.4170027	0.000451	2.24036	1.49678