

Revised, February 25, 1999

New Survey of London Life and Labour
Codebook for the
Data Files (Release 2.0)

Roy E. Bailey

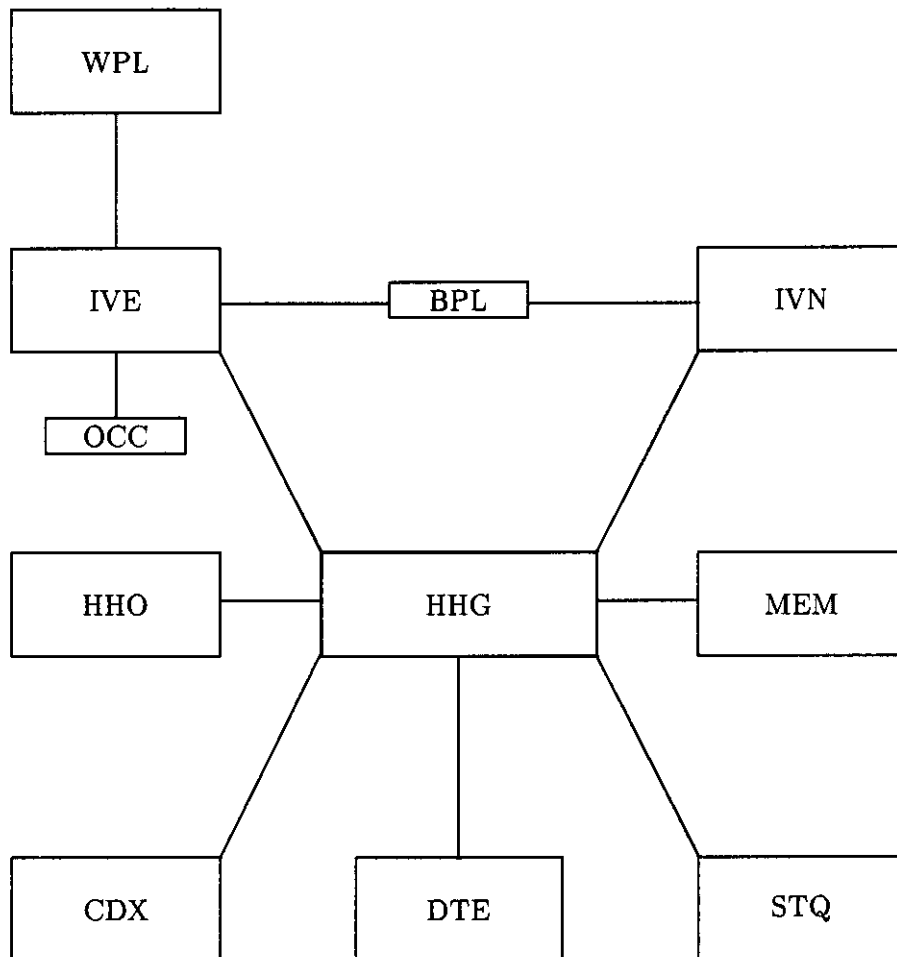
University of Essex

Anna Leith

London School of Economics

Acknowledgement: The computerisation, coding and analysis of the New Survey of London Life and Labour was funded under ESRC awards no. R000235697 and no. R000221981. Hamish James (Data Archive, University of Essex) kindly conducted a careful check of the data and enabled the correction of several errors in an earlier version of the codebook.

The New Survey Data Tables



New Survey Data Tables:

- HHG General household information, one record per card.
- HHO Other sources of income, one record per source.
- IVE Earners' information, one record per earner.
- IVN Non-earners' information, one record per non-earner.
- MEM Free-format text fields.
- STQ Street Quality codes, at least one record per card.
- CDX Card numbers for supplementary cards, zero or more records per card.
- WPL Work place location codes, one record per earner.
- DTE Date of interview (modified), one record per card.

Look-up Tables:

- OCC Occupation codes (related to IVE via a four character occupation code).
- BPL Birthplace codes (related to IVE and IVN via a seven character place code).

Each data table is related to the others via a unique, six-character code.

Contents

1: Overview	1
2: File Structures	2
2.1: Levels, Releases and Versions	2
2.2: Files Included at the Various Levels	2
2.3: A Note on Households and Cards	3
2.4: Free-Format Text Fields	4
2.5: Formats for Output Files	4
2.6: Money Fields	5
2.7: Non-Standard and Peculiar Entries	6
3: Summary of Files and Field Names	7
4: General Household Information (HHG)	11
5: Income From Other Sources (HHO)	21
6: Wage Earners' Information (IVE)	23
7: Non-Wage-Earners Information (IVN)	32
8: Supplementary Cards (CDX)	35
9: Free Format Text Fields (MEM)	35
10: Street Quality Codes (STQ)	36
11: Workplace Codes (WPL)	38
12: Interview Dates (DTE)	39
13: Occupation Codes (OCC)	40
14: Birthplace Codes (BPL)	43

List of Tables

1: General household information, one record per card: HHGxxxx.DBF	7
2: Other sources of income, one record per source: HHOxxxx.DBF	8
3: Earners' information, one record per earner: IVExxxx.DBF	8
4: Non-earners' information, one record per non-earner: IVNxxxx.DBF	9
5: Card numbers for supplementary cards: CDXxxxxA.DBF	9
6: Free format text fields: MEMxxxxA.DBF	9
7: Street quality codes: STQxxxxA.DBF	10
8: Workplace Codes: WPLxxxE.DBF	10
9: Interview Dates: DTExxxE.DBF	10
10: Occupation codes: OCCxxxxA.DBF (<i>Look-up table</i>)	10
11: Birthplace Codes: BPLxxxxA.DBF (<i>Look-up table</i>)	10
12: Non-standard entries in CARD_H	11
13: Non-standard entries in FILE_H	12
14: Borough codes and cases per borough	14
15: Names of Interviewers	16
16: Codes for tenure status	17
17: Number of persons	18
18: Number of bedrooms	19
19: Accommodation characteristics	19
20: Income from Other Sources	22
21: Sex of Wage-Earners	23
22: Relationship to Head of Household	25
23: Skill classification	28
24: Employment Status of Wage-Earners	29
25: Sex of Non-Wage-Earners	32
26: Non-standard card numbers	35
27: Street Quality Codes	37
28: Workplace Adjacency Codes	39
29: Amendments to Interview Dates	39
30: Occupation Order Codes	42

1: Overview

1. Cards for 26915 households have been coded. A total of 49445 individuals are entered as wage-earners, although by no means all of these were labour force participants. Of the total, 39797 are coded as employed or self-employed—this number excludes the unemployed, the incapacitated and those on strike as well as those not normally available for employment. A total of 44691 individuals are entered as non-wage-earners.
2. The data files are organised in sets at two 'levels', A and E (see section 2.1), according to the degree of coding. (Levels B–D are reserved for internal project use.) Level A files contain the data entered at the input stage, except that corrections are included and a very limited standardisation (in particular with respect to the spelling of interviewers' names) has been made. Level E files contain a limited number of variables for which numerical codes are assigned.

Excluded from the dataset are records for 1355 cards for middle class or Jewish households which were included *only* at the data input stage, not thereafter. In any case, these records are, for the most part, incomplete. All summary statistics reported in the codebook exclude the middle class and Jewish households.

3. The data files released into the public domain are the result of extensive checking and cleaning, as well as coding, which was carried out after completion of the data input from the cards. Even so, errors and inconsistencies remain. These will be investigated and corrected wherever possible.
4. The codebook is organised as follows: section 2 outlines the general structure of the data files; section 3 summarises the variables contained in each file. The remaining sections provide detailed descriptions and summary statistics of all the variables that appear in levels A and E arranged according to the file in which they appear.
5. A companion paper, *Computerising and Coding the New Survey of London Life and Labour*, (henceforth "*Companion paper*") has been prepared to provide detailed discussions of the most problematic aspects of the data and explanations of the coding principles adopted for Street Quality, Occupations and Birthplaces. Throughout the codebook and companion paper, the shorthand "*New Survey*" refers both to the published volumes of *New Survey of London Life and Labour* (9 vols. 1930–35, London: P. S. King & Son) and to the *New Survey* itself.
6. Caution should be exercised in interpreting some elements of the data. Notes accompany the variable definitions where specific caution is merited. The main weaknesses stem from:
 - (a) Failure of interviewers to adhere to the instructions (a summary of which appear in the *New Survey* vol. III, pp. 413–415). The biggest shortcoming, indeed the most severe problem overall, is that interviewers often failed to complete the card thus resulting in many missing values. A facsimile of a sample card, copied from *New Survey* vol. III, pp. 424–425, appears at the end of this codebook on page 45.
 - (b) Difficulties encountered by the interviewers in handling cases not envisaged, or allowed for, in the design of the card.
 - (c) Difficulties in reading the interviewers' handwriting.

7. Terminology. The term 'field' should be thought of as a column in a rectangular array of data and can be interpreted as corresponding to one variable. The term 'record' should be thought of as a row in a rectangular array of data and can be interpreted as corresponding to one observation (on a household, a source of income, a wage-earner, or a non-wage earner). The term 'table' is used interchangeably with the term 'file' for database files.

8. *Changes in Release 2.0.* Several amendments and extensions are incorporated in release 2.0 of the data:
- (a) File WPL (see page 38) has been added to provide information about the geographical location of each earner's workplace.
 - (b) File DTE (see page 39) has been added to clarify the date of interview (where this is needed and where it is possible).
 - (c) The employment status code (STAT_E, see p. 28) has been amended to signal those individuals for whom earnings are combined jointly with at least one other person in the same household. A total of 191 earners are affected.
 - (d) Amendments have been made to data for hours and earnings (HRLW_E, HRFT_E, EALW_E, EAFW_E). The corrections affect approximately 40 individuals in about 30 households.

2: File Structures

2.1: LEVELS, RELEASES AND VERSIONS

As the data are processed the relevant files can be thought to pass through various LEVELS beginning with the input level and passing through subsequent levels each designated with a letter code, A, B, etc. Levels A and E are to be released into the public domain. (Levels B-D were used for processing within the project. In addition, the input level files are not being released into the public domain.) Consequently, *this codebook documents levels A and E only.*

Updated file sets are made available in RELEASES at discrete times. *Release 2.0* refers to the first release of the data into to into the public domain.

The files in each version are identified with a three digit numeric value. For *release 2.0* most files are identified by 500. (The number has significance only for internal project use.)

Level A data files merge all the 192 input level file sets into one set.¹ Amendments to data have been made to correct errors found during the process of checking the data. A considerable amount of editing and a limited amount of standardisation, are incorporated in level A files.

Level E data files consist of all the numerically coded data. No alphabetic characters appear. It is expected that level E files will be used as the immediate point of departure for statistical work.

2.2: FILES INCLUDED AT THE VARIOUS LEVELS

Although there are many common features, differences do exist among the files between levels A and E. An important common element is a unique code (here called the IDNO) for each household. Where there can be more than one record per household (e.g. for earners), each record is identified by an extra 'individual' (person or income source) number. Hence, it is always possible to track every record uniquely across both levels.

File identification: Each file (at both levels) is identified by a three character prefix, e.g. HHG. Each *field* is associated with the file in which it appears by a *file identifier* which is the final character in its name, e.g. all field names ending in H (such as: PARL_H) appear in the HHG file. Note, in particular, that the household's IDNO appears with a different field name in each of the files, IDNO_H, IDNO_S, IDNO_E, etc.

¹An additional household was added during the checking and cleaning. For technical reasons this is attributed

The files are arranged as follows:

1. HHG This file contains information common to all members of the household. There is one record for each household.
File identifier for fields in HHG: H.
2. HHO This file contains information on the INCOME FROM OTHER SOURCES listed on each card, a separate record for each source in each household.
File identifier for fields in HHO: S.
3. IVE This file contains information on WAGE-EARNERS, a separate record for each wage-earner.
File identifier for fields in IVE: E.
4. IVN This file contains information on NON-WAGE-EARNERS, a separate record for each non-wage-earner.
File identifier for fields in IVN: N.
5. CDX This file contains information on supplementary card numbers written on many of the cards.
File identifier for fields in CDX: C.
6. MEM This file contains free-format textual information recorded on the cards.
File identifier for fields in MEM: M.
7. STQ This file lists street quality code(s) associated with the address entered on each card.
File identifier for fields in STQ: Q.
8. WPL This file contains codes to locate the workplace of each WAGE-EARNER, a separate record for each wage-earner.
(WPL was introduced in Release 2.0)
File identifier for fields in WPL: W.
9. DTE This file contains the date of interview for each household (including the year of interview, often omitted from the cards), one record per card.
(DTE was introduced in Release 2.0)
File identifier for fields in DTE: D.
10. OCC This file is a look-up table containing information (in particular, census occupation titles) associated with the coding of occupations.
File identifier for fields in OCC: O.
11. BPL This file is a look-up table that links the birthplace codes with formal place name descriptions.
File identifier for fields in BPL: B.

2.3: A NOTE ON HOUSEHOLDS AND CARDS

In almost all cases there is one card per household. In rare cases, a single household was so large that the interviewer needed to use a second card to make a complete record. In these cases both cards have been treated as a single household and assigned the same IDNO_H.

There are also cases where the interviewer may have erroneously treated the various occupants of a building as separate households and, accordingly, recorded them on separate cards. These are cases for which individuals described as lodgers appear to have been treated as sub-tenants. It is suggested they occur because some of the interviewers apparently experienced difficulties in making the distinction between lodgers and subtenants. The former were regarded

as part of the family with whom they shared both board and lodging; the latter as separate households, merely paying rent for the use of a room or rooms. In these cases the error may have been noted in comments added subsequently to one of the cards but, in any case, no attempt has been made to correct it and the original interviewer's work has been accepted at face value. Each of the cards has been treated as a separate household with its unique IDNO_H number.

2.4: FREE-FORMAT TEXT FIELDS

With three exceptions, all information has been entered into fixed length character fields. The exceptions are REMA_M, REMG_M, REMI_M. See section 9 (page 35). Initially, these fields appeared in HHG as REMA_H, REMG_H, REMI_H, but were subsequently relocated in MEM.

It should be noted that there is not complete consistency in the use of the three free-format text fields for recording particular types of information. As this information was copied directly from the cards, it depends on the practices of the individual interviewers. Whilst any additional comments about the accommodation are almost always to be found in REMA_M, and birthplaces in REMG_M, other information relating to household members, their occupations, *etc.* may occur in either REMG_M or REMI_M.

The free-format text fields have also been used to enter clarifying remarks by the data-inputter (for example, to indicate cases where there has been difficulty in reading the interviewer's handwriting). These remarks are typically preceded or followed by the signal *op note*, or *-Op*.

While the free-format text fields contain much valuable information, strict conventions were not enforced for the layout and format of the entries made in them. For this reason, rigid rule-based searches are prone to error. The information that the fields contain is more reliably extracted by simple visual searches.

2.5: FORMATS FOR OUTPUT FILES

The data are processed, at all levels, in dBase, DBF, format. Note that the level E files contain only numeric data and, hence, can be used in applications that are restricted to numeric input.

The data in the free-format (variable length) fields, REMA_M, REMG_M, and REMI_M were initially processed in a proprietary variant of the DBT file structure associated with dBase files. In the public release, the file structure has been translated to the standard DBT format. This format is suitable for importing into a variety of other application packages (e.g. Paradox or Microsoft Access).

In addition, the free-format text fields are available in an ASCII file named MEMxxxA.TXT. In this file each field entry begins with one line identifying the entry. This line contains a six character household number followed by a space and the name of the field (one of: REMA_M, REMG_M, REMI_M) followed by a space and the number of lines in the field (5 characters width, right justified). Following the identifying line, there are the requisite lines of text. Each line is a maximum 75 characters in length.

Example:

```
011994 REMI_M      3
Earner 01 was unemployed until a month ago. Now has employment
although it may be intermittent.
Earner 02 appears to have been out of work for over 6 months.
011997 REMG_M      4
A line of text.
Another line of text.
Some more text
```


Yet more text.

001204 REMA_M 1

House in very poor state of repair.

MEMxxxA.TXT contains only ASCII characters 32 to 126 with the following exceptions:

1. Carriage return, ASCII character 13.
2. Line feed, ASCII character 10.
3. The pound sign, £, ASCII character 156.

2.6: MONEY FIELDS

Common conventions affect all the money fields: RENT_H (rent paid), INOS_S (amount of income source), TRAN_E (transport expenses), EALW_E (earnings last week), EAFW_E (earnings, full week), and SIND_E (state insurance deductions).

The format of money fields differs between levels A and E.

For level A the rules are as follows:

- Each field is of 10 characters width, free format allowing numbers, letters and other keyboard characters (though inputters were discouraged from using the £ symbol).
- Numerical amounts are written in pounds, shillings and (old) pence, separated with a '/'. Thus, 4/12/6.25 is four pounds twelve shillings and sixpence farthing. If no '/' appears, the amount is interpreted as shillings alone. If one '/' appears, the amount is interpreted as shillings and pence if it is between two numbers, and as pence alone if there is no leading number. The presence of '/' indicates that the leading figure is an amount in pounds (e.g. 3// is three pounds). These rules were not always rigidly adhered to by data inputters but, in almost all cases, the intention is self evident on inspection.
- Unless otherwise indicated, the amounts entered are weekly. Where the amount clearly related to a period other than a week, a letter Y was appended for annual amounts; S for six monthly (half-yearly) amounts; Q for quarterly amounts; M for monthly amounts; D for daily amounts; and H for hourly amounts. The last two categories apply only to the earnings fields, EALW_E and EAFW_E. While this convention has been followed for the most part, there are some exceptions; in particular a few annual payments have not been identified as such and in some cases other forms of words (e.g. 'yr', 'pa' or 'P.A.') have been used to qualify a monetary amount.

For level E the rules are as follows:

- Each field is of 6 characters width, right justified and padded left with zeros.
- Values are coded as hundredths of (old) pence per week. (Example: 106225 equals 4 pounds, 8 shillings and 6 pence, one farthing.) This should not be interpreted as providing a high degree of accuracy; rather, it allows for the entry of farthings where these appear.
- Where possible, amounts for other than a week are suitably translated.
- Ranges (typically given with a dash, such as 11-15) are each coded as the average of the two extremes.
- Other words, or expressions, (e.g. 'approx') entered in the field are ignored except where a definite value can be inferred (e.g. 'nil' is interpreted as zero).

2.7: NON-STANDARD AND PECULIAR ENTRIES

Wherever possible, information has been entered as it appears on each card. Exceptions to this arise as a result of the problems outlined on page 1, paragraph 6. These exceptions are either described in this codebook or, where they relate purely to individual cases, are noted in the appropriate text field. Further details about particular problems of interpretation may be found in the *Companion paper*.

Some attempt has been made to standardise numerical entries, e.g. money values. This means that words, abbreviations or phrases appear in many cases for fields in which numerical entries alone would be expected. Where it is reasonably safe to do so, these words have been ignored; otherwise a missing value code (i.e. the field is filled with 9) is inserted in the coded files.

With the exception of accommodation characteristics, if a dash, '-', appears in a field it is interpreted as 'unknown' (and a letter X is entered). For the accommodation fields a dash is interpreted as 'none' (or zero). Also, for accommodation characteristics, ticks sometime appear against the facility in question. These ticks are normally interpreted as meaning 'yes' (or one), although in these cases a judgement was made in the context of what was written on the card and on the normal practice of the interviewer.

3: Summary of Files and Field Names

Tables 1 to 6 contain information on the variable names, and the field widths. Page numbers are provided for detailed descriptions of the fields which appear in the public release of the dataset.

Table 1: General household information, one record per card: HHGxxxx.DBF

Field	Input	A	E	Description	Page
IDNO_H	5	6	6	Household Identification Number	11
QURY_H	1	-	-	Query identification	-
CARD_H	5	5	-	Principal Card number	11
FILE_H	5	5	-	File number	12
NAME_H	40	-	-	Name of householder (omitted)	12
ADDN_H	20	30	-	Address: house identification.	12
ADDR_H	30	30	-	Address: Street	12
SQMC_H	-	1	-	Street quality match code.	13
BCRO_H	15	15	2	Borough	14
DATE_H	6	6	6	Date of interview	14
INTV_H	20	20	3	Interviewer's name	15
RENT_H	10	10	6	Rent paid	15
RTEN_H	1	1	2	House tenure code	15
NOPN_H	2	2	2	Number of persons	17
NOBR_H	2	2	2	Number of bedrooms	18
PARL_H	1	1	1	Parlour	18
KITC_H	1	1	1	Kitchen	18
SCUL_H	1	1	1	Scullery	18
PNTR_H	1	1	1	Pantry or Larder	18
BATH_H	1	1	1	Bathroom	18
YARD_H	1	1	1	Yard	18
GARD_H	1	1	1	Garden	18
ALLO_H	1	1	1	Allotment	19
REMG_H	10	-	-	General remarks (pointer)	-
REMA_H	10	-	-	Accommodation remarks (pointer)	-
REMI_H	10	-	-	Earner, etc. remarks (pointer)	-
REMO_H	10	-	-	Operator's Remarks (pointer)	-
OPDT_H	6	-	-	Date of record entry or amendment	-
OPIN_H	3	-	-	Operator's initials	-
CHEK_H	6	-	-	Internal check	-
CLAS_H	2	-	-	Class of household	-
ODDC_H	20	20	-	Odd characters on front	20
XTRA_H	10	-	-	Used to extend ADDN_H	-
XTRB_H	2	-	-	Reserved	-
PROJ_H	-	10	-	Internal Project Use	-

Table 2: Other sources of income, one record per source: HHOxxxx.DBF

Field	Input	A	E	Description	Page
IDNO_S	5	6	6	Household Identification Number	21
SRNO_S	2	2	2	Source number	21
SRCE_S	35	35	-	Description of source	21
INOS_S	10	10	6	Amount of income	21
SRCO_S	1	1	2	Code for Source of income	21
XTRA_S	10	-	-	Reserved	-
PROJ_S	-	10	-	Internal Project Use	-

Table 3: Earners' information, one record per earner: IVExxxx.DBF

Field	Input	A	E	Description	Page
IDNO_E	5	6	6	Household Identification Number	23
IVNO_E	2	2	2	Earners number	23
SEXX_E	1	1	1	Sex	23
AGEY_E	5	5	4	Age	23
RELA_E	15	15	2	Relationship to Head	24
RELN_E	-	-	2	Head Identification	26
BPLC_E	30	30	-	Birthplace	26
BPCC_E	-	7	7	Birthplace code	26
BPCN_E	-	-	1	Conurbation code	26
BPCM_E	-	1	-	Match code (assigned in coding)	27
OCCU_E	30	30	-	Occupation	27
OCCC_E	-	4	4	Occupation code	27
OCCS_E	-	-	1	Skill level	27
OCCO_E	-	-	2	Order (Census occupation grouping)	28
OCCM_E	-	1	-	Match code (assigned in coding)	28
STAT_E	3	3	1	Employment Status code	28
EMPL_E	30	30	-	Employer	29
PLWK_E	30	30	-	Place of Work	30
TRAN_E	10	10	6	Cost of transport	30
HRLW_E	8	8	5	Hours last week	30
HRFT_E	8	8	5	Hours, full time	30
EALW_E	10	10	6	Earnings last week	30
EAFW_E	10	10	6	Earnings in a full week	31
SIND_E	10	10	6	State insurance deductions	31
XTRA_E	10	-	-	Reserved	-
PROJ_E	-	10	-	Internal Project Use	-

Table 4: Non-earners' information, one record per non-earner: IVNxxxx.DBF

Field	Input	A	E	Description	Page
IDNO_N	5	6	6	Household Identification Number	32
IVNO_N	2	2	2	Individual Number	32
SEXX_N	1	1	1	Sex	32
AGEY_N	5	5	4	Age	32
RELA_N	15	15	2	Relationship to Head	33
RELN_N	-	-	2	Head Identification	33
BPLC_N	30	30	-	Birthplace	33
BPCC_N	-	7	7	Birthplace code	33
BPCN_N	-	-	1	Conurbation code	33
BPCM_N	-	1	-	Match code (assigned in coding)	34
PROJ_N	-	10	-	Internal Project Use	-

Table 5: Card numbers for supplementary cards: CDXxxxxA.DBF

Field	Input	A	E	Description	Page
IDNO_C	6	6	-	Household Identification Number	35
CARD_C	5	5	-	Supplementary Card number	35
PROJ_C	-	10	-	Internal Project Use	-

Table 6: Free format text fields: MEMxxxxA.DBF

Field	Input	A	E	Description	Page
IDNO_M	-	6	-	Household Identification Number	35
REMG_M	-	10	-	General remarks (pointer)	35
REMA_M	-	10	-	Accommodation remarks (pointer)	36
REMI_M	-	10	-	Earner, etc. remarks (pointer)	36

Table 7: Street quality codes: STQxxxA.DBF

Field	Input	A	E	Description	Page
IDNO_Q	-	6	-	Household Identification Number	36
QUAL_Q	-	2	-	Street Quality code	36
PRED_Q	-	1	-	Predominance code	37

Table 8: Workplace Codes: WPLxxxE.DBF

Field	Input	A	E	Description	Page
IDNO_W	-	-	6	Household Identification Number	38
IVNO_W	-	-	2	Individual Number	38
WPLC_W	-	-	7	Workplace code	38
ADCD_W	-	-	1	Adjacency code	38

Table 9: Interview Dates: DTExxxE.DBF

Field	Input	A	E	Description	Page
IDNO_D	-	-	6	Household Identification Number	39
DATE_D	-	-	6	Interview Date	39
AMND_D	-	-	1	Amendment quality code	39

Table 10: Occupation codes: OCCxxxA.DBF (*Look-up table*)

Field	Input	A	E	Description	Page
CENC_D	-	4	-	Modified Census Code	40
CODE_D	-	4	-	Numeric code	40
CLAS_D	-	1	-	Skill level	41
ORDE_D	-	2	-	Order	41
DESC_D	-	50	-	Description	43
COUN_D	-	5	-	Number of occurrences	43

Table 11: Birthplace Codes: BPLxxxA.DBF (*Look-up table*)

Field	Input	A	E	Description	Page
BPLC_B	-	7	-	Code identifying place name	43
CONU_B	-	1	-	Code identifying conurbation	44
DESC_B	-	50	-	Place Name	44
COUN_B	-	5	-	Number of occurrences	44

4: General Household Information (HHG)

- File: HHG Field: IDNO_H = *Household Identification Number*.

Levels: A and E.

Length: 6 characters, identical for levels A and E.

Description: A number assigned at data entry. It is unique to each household entered and establishes the relationships across data files. The first three digits are defined by the file set number determined for data entry. The second three digits were assigned at data entry in the sequence in which the cards were inputted. Some gaps in the sequence numbers appear as a result of subsequent editing. *The household identification number is unique to each household and should be used to form relationships among the data files.* It has no other significance.

- File: HHG Field: CARD_H = *Principal card number*.

Level: A only.

Length: 5 characters.

Description: A four-digit number which was stamped on each card at the time of the survey. In most cases (see below) the card number uniquely identifies the household recorded on the card within its borough. There is reason to think that the card numbers relate to the way in which the collection of the data was organised. They appear to have been designated in many boroughs largely according to the alphabetical sequence of the names of the streets from which addresses had been selected for inclusion in the survey. In some boroughs, subsequent alphabetical sequences of street names at the end of the card number sequence suggest that additional addresses were sometimes included at a later date, presumably when it was felt that the original sample had not yielded sufficient cases.

Some of the original card numbers were subsequently crossed out and replaced with an alternative. In each case the alternative has been treated as the principal card number and the original (i.e. deleted) values are recorded as textual information. In these cases the number may not be unique to the household on the card within its borough. In a very few cases an alphabetic character was appended to the number. These, together with other non-standard entries that appear in the database, are listed in Table 12.

See also, CARD_C, page 35.

Table 12: Non-standard entries in CARD_H

IDNO_H	CARD_H
039079	0392
043088	IT
046031	0872
046032	0593
055004	0416
063057	X
077122	F1366
135058	D1259
135061]1424
166031	X
171043	X
181104	B
246012	D1394
246021	D1485

- File: HHG Field: FILE_H = *File number.*

Level: A only.

Length: 5 characters.

Description: A running number assigned to each completed *working class* card at the time of the survey to determine its unique position in the filing sequence for each borough. All cards were sorted by borough and within each borough, by type of household, (e.g. single person, couple, with children, with working children, pensioners and widows). The numbering sequence starts at 1 in each borough. Non-standard entries are listed in Table 13.

Table 13: Non-standard entries in FILE_H

IDNO_H	FILE_H
115097	U
180094	=694
261095	#695

- File: HHG Field: NAME_H = *Name of householder.*

Level: A only.

Length: 40 characters.

Description: The name of the individual entered on the card for **Name**. Names have been entered in the form: Surname, First-name(s) (Title and/or Marital-status). The form of wording or abbreviation given on the card has been preserved at data entry.

Note: This field is omitted from the public release of the data.

- File: HHG Field: ADDN_H = *Address: House identification.*

Level: A only.

Length: 30 characters.

Description: The part of the address which identifies the dwelling place of the household. For most houses this is typically just a number. In some cases, e.g. workmen's flats in Peabody Buildings, the house identification includes the number (of the flat) together with identification of the building. For cases in which a house built originally for occupancy by a single household has become divided into multiple tenements, there is usually no individual identification of the tenements even though they appear on different cards.

House identification details have been entered as given on the card. Where a flat or tenement number includes a letter, this typically follows the number, and is entered in upper-case or lower-case as written on the card. Where the name of a building is given, this follows the identification of the individual flat or tenement. Abbreviations have been used where necessary: Bldgs or Bdgs for Buildings; Mans for Mansions.

- File: HHG Field: ADDR_H = *Address: Street.*

Level: A only.

Length: 30 characters.

Description: This contains the name of the street in which the dwelling is located.

The following abbreviations have been used in some cases: Rd for Road; St for Street; Sq for Square; Av for Avenue; Pl for Place; Terr for Terrace; Gdns for Gardens; Cresc for Crescent; Cir for Circle; Ln for Lane; Pk for Park; Bldg for Building; Mans for Mansions.

Note that there was some variation in the way the same street names were recorded on the cards by different interviewers. As part of the street quality coding process, considerable efforts were made to check and standardise the street names according to London County Council *List of streets and places within the administrative county, 1929-35*, and other sources. As a result, street names are standardised as far as possible. (See the *Companion paper* section 10, p. 28.) Where doubt remains a note has been added to REMI.M. In 41 cases an 'X' (followed by blanks) signifies an unknown address.

Also during the street coding process, it emerged that some of the addresses surveyed in particular boroughs were in streets that appear to have actually been part of another, neighbouring borough. For example, there are about 60 streets from which addresses were selected for the Camberwell sample, which appear from contemporary maps to have lain within St Paul's Deptford. Also, the division of Lambeth into North and South (for *New Survey* purposes only) resulted in the misallocation of some streets.

It is thought that these discrepancies may have arisen because of small boundary changes around that time; a comparison of the *New Survey* street survey maps (see section 10, page 36) with LCC street maps for the period does appear to show slight differences in the location of some borough boundaries. The number of streets affected is thought to be relatively small and no attempt has been made to reclassify them; the borough boundaries used in the *New Survey* have been accepted at face value.

- File: HHG Field: SQMC_H = *Street Quality Match Code*.

Level: A only.

Length: 1 character.

Description: a code that indicates when the street quality coding recorded for a particular household in QUAL_Q is one of three special cases. Normally SQMC_H contains D (to stand for Definite). The exceptions are:

1. SQMC_H=A: the street coding applies to an "adjacent" street to the one in which the household dwelling is located, because the actual street cannot be located definitely on the *New Survey* map. The name of the coded street is given in ADDR_H, after the actual street, and separated from it by a comma.
2. SQMC_H=C: the street coding applies to a "combination" of streets. These cases arise where ADDR_H contains more than one street name, separated by "and" or "or" because the entry in ADDN_H is a block (e.g. Peabody Buildings) that either fronts onto more than one street, or occurs in more than one location in the borough. In these cases, the address information given on the card did not allow a precise entry for ADDR_H to be determined.

Also, SQMC_H=C for cases in which there existed more than one street in a borough with the same name (e.g. 'High Street') and it was not possible to distinguish which applied to the dwelling in question. The street quality code(s) in these cases apply to both (or all) of the streets of that name in the borough.

3. SQMC_H=N: on the *New Survey* street map, the street appears with a name that is different from that in ADDR_H. The alternative name has been entered in REMI.M. In most cases the discrepancy has been verified as a name change in the 1929 LCC directory.

See the comments about QUAL_Q on page 36; also the *Companion paper* section 10, p. 28.

Table 14: Borough codes and cases per borough

Borough (Eastern Area)	Code	Cases	Date†	Borough (Western Area)	Code	Cases	Date†
Bethnal Green	01	1278	1929	Finsbury	18	558	1929
Shoreditch	02	701	1929	Holborn	19	471	1930
Stepney	03	1265	1929	Westminster	20	416	1930
Bermondsey	04	672	1930	Lambeth North	21	800	1932
Hackney	05	962	1929	Lambeth South	22	653	1932
Poplar	06	884	1929	Southwark	23	1045	1931
Stoke Newington	07	259	1929	Fulham	24	856	1930
Deptford	08	519	1931	Hammersmith	25	634	1930
Greenwich	09	680	1930	Islington	26	2040	1929
Lewisham	10	630	1931	St Pancras	27	1305	1932
Woolwich	11	636	1930	Chelsea	28	212	1930
Barking	12	392	1930	Hampstead	29	289	1931
East Ham	13	534	1931	Kensington	30	484	1930
Leyton	14	530	1930	Paddington	31	888	1931
Tottenham	15	0	1930	St Marylebone	32	328	1930
Walthamstow	16	0	1931	Battersea	33	847	1930
West Ham	17	1263	1929	Camberwell	34	1154	1931
				Wandsworth	35	910	1930
				Acton	36	686	1930
				Hornsey	37	422	1930
				Willesden	38	712	1931

†: Year in which interviewing was completed.

- File: HHG Field: BORO_H = *Borough*.
 Level: A and E.
 Length: 15 characters in A; 2 characters (code) in E.
 Description (level A): The name of the borough in which the dwelling falls.
 Description (level E): Two digit number identifying the borough.
 Coding: See Table 14. Note that no working class cards have survived for Tottenham and Walthamstow. (Some middle class and waste cards survive for Tottenham.)
- File: HHG Field: DATE_H = *Date of interview*.
 Level: A and E.
 Length: 6 characters in both levels.
 Description (levels A and E): The date at which the interview was conducted.
 Coding: Zero-filled where no date (or part of a date) is entered or illegible. Otherwise entered as Day, Month, Year (two characters each). For 5110 cards no date is available and for 5211 cards the year is unspecified. A rather imprecise guide to the date of interview appears in volume VI of the *New Survey* where, on page 36, a list of the year of completion for each borough is provided. The year of completion is given in Table 14 on page 14. Evidence from the dates written on the cards shows that practically all of the interviewing was carried out before 1931 or 1932 where this is the year of completion listed in Table 14.

Interview dates written on the cards suggests that, for several boroughs, a trivial amount of interviewing (at most three cards in any one borough) took place after the date listed in Table 14. The boroughs affected are Acton, Barking, Hammersmith, Islington,

Westminster and Willesden. For one borough, Greenwich, the evidence from the reimbursement of interviewers suggests that some interviewing may have taken place in early 1931.

Subsequent to release 1.0 of the data, efforts were made to estimate the year of interview where this was omitted. File set DTE contains a list of all interview dates with the insertion of estimates for years missing in DATE_H. where this was not entered on the card (See page 39, below, and the *Companion paper*, page 48).

- File: HHG Field: INTV_H = *Interviewer's name.*

Level: A and E.

Length: 20 characters in A; 3 characters (code) in E.

Description: Name of interviewer. In some cases the name of the interviewer is either not given or the entry is illegible. Where the interviewer can be identified, a standardised form of each distinct name appears in the level A file. A total of 171 distinct names has been identified although the handwriting on the cards is such that it is not possible to be completely certain about the form of all the names. It is probable that the total number of interviewers exceeded 171 but it is unlikely that further information could be extracted from the cards.

Coding: See Table 15 on page 16. Code 999 is used where no interviewer is identified or the entry is illegible.

- File: HHG Field: RENT_H = *Rent paid.*

Level: A and E.

Length: 10 characters at level A; 6 characters at level E.

Description: Amount entered for 'RENT' Many of the interviewers clearly had difficulty in dealing with non-standard cases and, moreover, were not able to adhere fully to the instructions they were given. Particular difficulties occurred with respect to: (a) owner-occupied dwellings (for which there was much confusion about the inclusion or non-inclusion of an amount for rates); and (b) cases in which an occupier sublet part of the accommodation (so that the rent figure appearing on the card may be the gross amount paid to the external landlord, as required, but is sometimes the net figure after allowing for receipts from subtenants). See the *Companion paper*, section 3, page 5, for details about the problems encountered and the ways in which they were addressed.

A variety of miscellaneous textual entries appear in the level A file. For example, Y denotes a payment per year, est to indicate an estimated amount, and abt for an approximate value.

Note that, apart from owner-occupied cases, the coding in RTEN_H (see below) clarifies the meaning of the amount entered in this field. *In owner-occupied cases, the value of the entry should be treated with particular caution.*

Coding: A money code. See section 2.6 on page 5.

- File: HHG Field: RTEN_H = *House tenure code.*

Level: A and E.

Length: 1 character at level A; 2 characters at level E.

Description: A code, determined at the point of data input, to denote the tenure status of the occupant. The letter codes at level A are as follows: Owner occupied (including leaseholders) = O; Rented = R; Rent free = F; Rented from employer = E; Other = Z;

Table 15: Names of Interviewers

Interviewer	Code	Cases	Interviewer	Code	Cases	Interviewer	Code	Cases
H T Allen	001	164	W A Griggs	059	70	Porter	117	108
B J Andrade	002	29	Gunn	060	73	Powell	118	19
P Arnold	003	35	R D Hale	061	218	E H Price	119	97
R Augless	004	16	P B Harding	062	18	Rabinovitch	120	66
Baily	005	49	T B Harris	063	18	Rason	121	1
S Baker	006	30	G H Hawkins	064	83	Evan Rees	122	50
Baran	007	198	J Hewitt	065	86	M A Rex	123	74
S F Barnes	008	34	R P Hill	066	563	D W Roberts	124	201
R A Barrett	009	329	W C Hill	067	81	Geo Royds	125	76
T S Barron	010	80	D J Hilliard	068	115	B W Rush	126	54
G E Bartlett	011	4971	E H Hinson	069	306	H Russell	127	29
Henry E Bartlett	012	95	Hodson	070	54	F Russen	128	4
A E Bates	013	85	W Holtor	071	83	S Saxby	129	385
A Bayliss	014	95	A Honey	072	60	G Scarborough	130	23
C E Beeson	015	76	J A Hooker	073	89	C Shales	131	81
F Bell	016	93	L L Hooker	074	103	F W Shaw	132	90
W H Bennett	017	105	J Hopker	075	877	H W C Sillitoe	133	6
Berridge	018	70	E Hunt	076	18	Simmonds	134	385
P L Bird	019	57	W F Hunt	077	75	J C Simpson	135	93
Blore	020	88	F B Jefferys	078	48	A J Smith	136	114
A G Boreham	021	55	F Jennings	079	109	H S H Smith	137	13
J Bowler	022	153	J W Jones	080	74	A T Spooner	138	5
F Boyfield	023	78	E Kentish	081	106	L Stallard	139	17
Broadbent	024	83	Kerman	082	264	F A Stammers	140	456
H Brown	025	149	E Kessel	083	35	H A Stanford	141	98
J H Brown	026	391	W B Kilgour	084	84	A Stapleton	142	44
A B Butt	027	87	G C King	085	69	Stein	143	87
W M Calder	028	64	N T King	086	258	J Stevenson	144	235
W Candle	029	106	A E Kirby	087	86	A C Stopps	145	60
Cant	030	30	Lake	088	36	W G Stratford	146	178
W H Castle	031	599	H Lake	089	35	F Strutt	147	45
A C Cates	032	58	J H Lansdowne	090	147	C Sullivan	148	120
J E Chandler	033	64	S Larkins	091	104	Theobald	149	36
J Charters	034	41	Leach	092	7	V B Thurgood	150	846
H E Chilton	035	71	J F Living	093	17	A E Timms	151	89
Christian	036	88	J Ludgate	094	905	C L F Tomkins	152	13
Clouter	037	84	R H MacKiffe	095	91	W Trowbridge	153	7
G F Cobb	038	37	Mackinnon	096	45	J R Tuffley	154	94
F P Coleman	039	84	L P Mahony	097	316	A Turmean	155	64
A W Cooks	040	159	W G Malcolm	098	50	R W Turner	156	74
Cornell	041	97	A M Marr	099	26	Tuteur	157	39
V H Craft	042	70	T I May	100	16	J H Vincent	158	145
H Cribb	043	55	W J McKay	101	97	C R Waite	159	98
E B Crudge	044	169	McKenzie	102	3	W E Walsh	160	102
J W Davis	045	52	W Milliken	103	157	E C V Ward	161	81
W Dodgson	046	70	G Montgomery	104	131	B Waterhouse	162	136
A E Dowsett	047	123	W R Morgan	105	48	L A Wells	163	34
C Etherington	048	157	C W Morrison	106	131	Thos H Wenman	164	128
J Finch	049	102	C J O'Brien	107	54	F L Whiting	165	36
A I Ford	050	86	O'Connell	108	122	Wilkinson	166	3
G Foreman	051	56	J D O'Driscoll	109	124	W Wilkinson	167	76
H E Fox	052	110	H F Oldcorn	110	28	T E Williams	168	147
R E Friend	053	75	M Pamment	111	74	Miss Willsher	169	12
A J Ganiford	054	44	H Pearce	112	78	C Winch	170	59
Gaster	055	31	Pedley	113	34	A N Winter	171	874
K Gibberd	056	17	S Peett	114	79	X	999	2455
H E Green	057	194	H E Pile	115	16			
E Grierley	058	69	G C Plant	116	102			

Coding: See table 16 on page 17. Note that the second digit (at level E) identifies the tenure status while the first digit identifies other characteristics of the entry for 'rent'. The other characteristics are as follows:

1. The word 'net' or 'nett' appears alongside a rent figure. The implication is that the rent figure excludes amounts received from subtenants—such amounts should have been included (i.e. unless otherwise qualified, the rent should be interpreted as gross of receipts from subtenants.)
2. There is information (recorded in REMG_M or REMA_M) that the rent figure *excludes* rates. The rent figure should normally include rates.
3. There is information (recorded in REMG_M or REMA_M) that the rent figure records rates *only*.

Table 16: Codes for tenure status

	Code		Cases
	A	E	
Owner occupied (including leaseholders)	O	00	1128
Rented	R	01	25214
Rented ('net' or 'nett' appears)	R	11	130
Rented (rates excluded)	R	21	51
Rented (rates only)	R	31	2
Rent free	F	02	272
Rent free (rates excluded)	F	22	1
Rent free (rates only)	F	32	2
Rented from employer	E	03	31
Rented from employer (rates excluded)	E	23	1
Other	Z	98	10
Unknown	X	99	73

- File: HHG Field: NOPN_H = *Number of persons*.

Level: A and E.

Length: 2 characters at level A; 2 characters at level E.

Description: Number of persons who usually slept at the dwelling. In a small number of cases, this number differs from the total number of earners plus non-earners recorded on the front of the card. This usually occurs where members of the household are working away, but can also mean long-term absence in an institution or 'sleeping out'. The reason is indicated in a text field (typically REMI_M or, occasionally, REMA_M).

Where the interviewer has recorded the wrong number of persons on the card, this has usually been corrected by the inputter. Such an error typically occurred for one of the following reasons:

1. The interviewer did not follow the instruction to sum the number of earners and non-earners, including lodgers, but exclude absent earners. The instruction (page 415 of the *New Survey*, vol. III) was: "*The Number of Persons* should, of course, equal the number of occupied and unoccupied shown on the face of the card, including lodgers, but not absent wage-earners."
2. The interviewer had difficulty in interpreting the instructions with regard to absent non-earners or those who slept elsewhere (perhaps as a result of lack of space). There is no specific reference to these categories in the instructions, but other evidence

suggests they were to be excluded from the entry in NOPN_H. It appears that many interviewers made a judgement about the duration of absence: individuals who were away for extended periods were excluded but those who were absent for short periods were included.

Coding: See Table 17 on page 18. For household IDNO_H = 284079, NOPN_H = 6?, coded as 6.

Table 17: Number of persons

Number of persons	Cases
01	2964
02	6630
03	6055
04	4618
05	2863
06	1750
07	953
08	563
09	278
10	156
11	53
12	23
13	8
16	1

- File: HHG Field: NOBR_H = *Number of bedrooms.*

Level: A and E.

Length: 2 characters at level A; 2 characters at level E.

Description: Number entered for 'No. of Bedrooms'. Note that rooms were often used for purposes other than, or in addition to, their stated purpose. Hence, the entry in NOBR_H is not necessarily the total number of rooms that were used for sleeping in. Nor does it mean that that the bedroom(s) was used only for this purpose. Where there are apparently zero bedrooms (i.e. NOBR_H equals 00), it is almost invariably the case that there is a reference to one of the other rooms (typically, but not always, a parlour or kitchen) being used for sleeping.

Coding: See Table 18 on page 19. Several non-standard entries were coded according to a reasonable estimate. For example, - and N were coded as 0, and Y was coded as 1.

- Accommodation characteristics. The following fields have identical features for coding purposes:

- File: HHG Field: PARL_H = *Parlour.*
- File: HHG Field: KITC_H = *Kitchen.*
- File: HHG Field: SCUL_H = *Scullery.*
- File: HHG Field: PNTR_H = *Pantry or Larder.*
- File: HHG Field: BATH_H = *Bathroom.*
- File: HHG Field: YARD_H = *Yard.*
- File: HHG Field: GARD_H = *Garden.*

Table 18: Number of bedrooms

Number of Bedrooms	Cases
00	272
01	14823
02	8447
03	2883
04	315
05	35
06	2
07	0
08	0
09	2
12	1
Share	97
Other	98
Unknown	99

Table 19: Accommodation characteristics

Code		Parl.	Kitc.	Scul.	Pntr.	Bath	Yard	Gard.	Allo.
0	0	7051	4947	9230	15593	18366	12167	11236	20112
1	1	17004	19400	11405	3252	3276	6108	9780	254
2	2	395	50	10	4	0	17	327	0
3	3	10	1	0	1	0	1	0	0
4	4	1	0	0	0	0	0	0	0
S (Share)	7	32	308	2054	41	922	1781	1998	29
Z (Other)	8	0	0	2	0	0	0	0	0
X (Unknown)	9	2422	2209	4214	8024	4351	6841	3574	6520

- File: HHG Field: ALLO_H = *Allotment*.

Level: A and E.

Length: 1 character at level A; 1 character at level E.

Description (level A): The entry can indicate one of the following: the presence (Y or 1), or absence (N or 0) of the facility in question; a number (e.g. number of parlours if greater than one); or, that the facility was shared or the household had the use of it (i.e. S). The coded entries for accommodation characteristics should be treated with caution because the interviewers clearly used a variety of ways for handling non-standard cases they encountered. An attempt has been made to standardise the treatment of non-standard cases at the point of data entry. In particular:

1. Many approximations or qualifications appear for individual facilities. For example, 'roof drying ground' or 'downstairs' sometimes appear against yard; sometimes 'downstairs' appears against yard or 'small' appears against a facility. These comments have been recorded in the accommodation remarks text field, REMA.M. The general principle followed by the inputters in interpreting this information, was to consider whether or not it implied that the facility (or an equivalent facility) was available to the household. For example, the 'roof drying ground' mentioned above was interpreted as a shared used in most cases.

2. In some cases different facilities (e.g. kitchen and scullery) were bracketed together by the interviewers. In such cases the computer inputter was instructed to enter 1 for the first facility and 0 for the other(s). In each case a note is provided under accommodation remarks.

See the *Companion paper* section 4 p. 9 for more details about the problems of coding accommodation characteristics.

Coding (levels A and E): each entry contains a number if it appears; otherwise:

Level: Meaning:

A E

N 0 None, nil, -, or other indication of absence.

Y 1 Yes, tick or other affirmative entry.

S 7 Shared or 'use of'

Z 8 Other, not elsewhere specified

X 9 Unknown

Non-standard entries were combined with standard entries as appropriate in the coding for level E. For example, - and N were coded as 0, and Y was coded as 1.

In most cases the "unknown" code, 9 (letter X at level A), probably represents absence of the facility but this is uncertain because, in such cases, the space on the card has been left blank.

Summary statistics: See Table 19, page 19.

- File: HHG Field: ODDC_H = *Odd characters on face of card.*

Level: A only.

Length: 20 characters.

Description: Sequences of letters were written (probably during processing following the interview) on the front of most cards. They appear at the top, towards either the left-hand or the right-hand corners. The sequences were recorded as faithfully as possible during data entry but with little genuine understanding about their meaning. It is suggested that they are coding schemes; those on the right relate to the employment and relief status of the household; and those on the left to the status of the household in the dwelling it occupied, i.e. main occupant, subtenant, independent tenant. See the *Companion paper* section 8, p. 22 for additional information.

5: Income From Other Sources (HHO)

The 'other' in this context refers to income other than from earnings.

- File: HHO Field: IDNO_S = *Household Identification Number*.
 Level: A and E.
 Length: 6 characters.
 Description: See IDNO_H on page 11.
- File: HHO Field: SRNO_S = *Source number*.
 Level: A and E.
 Length: 2 characters.
 Description: A number assigned on inputting to identify the source number within the household. Always assigned even if only one source is present. The IDNO_S and SRNO_S are needed to identify each source uniquely.
- File: HHO Field: SRCE_S = *Description of source*.
 Level: A only.
 Length: 35 characters.
 Description: Text description of the source of other income.
- File: HHO Field: INOS_S = *Amount of income*.
 Level: A and E.
 Length: 10 characters at level A; 6 characters at level E.
 Description: Amount entered for the listed source.
 Coding: See 2.6 on page 5 for comments on non-standard entries for INOS_S in the level A file. A total of 15106 sources of income across all households are listed, 961 of which have an amount unknown, code 999999, and 4 of which are zero, code 000000.
- File: HHO Field: SRCO_S = *Source of income code*.
 Length: 1 character at level A; 2 characters at level E.
 Description: A letter code at level A assigned at the point of data input to identify source of income. At level E the code is expressed as a pair of digits. See Table 20 on page 22.
Double counting. In some cases an entry was made in sources of other income which, in fact, represented 'double counting' of income elsewhere recorded as coming into the household, typically included in the earnings of one of its members. Such cases have occurred in the context of income received from absent earners, lodgers, and relatives living with the family, for whom earnings information is also given. The categories of income affected are: 'Allowances from relatives' and 'Lodgers' An extensive check has been made of suspicious cases in the 'Allowances from relatives' category to eliminate the double counting, by omitting either the earnings of the absent person or the other income source (usually board payments) received from a 'resident' relative. The excluded amount is entered in the remarks field, REMI.M. Similar checks have not been made for the 'Lodger' category.

A significant number of cases appear in which the entry 'rent from lodger' is given as an income source, but there is no evidence of a lodger in the household. It is assumed that this arose from interviewers' difficulties in distinguishing between lodgers and subtenants. The income source has normally been coded as rent from subtenants in these cases.

See the Companion page 15 for further information.

Coding: See Table 20. The classification of income source should be transparent except to note that 'Pensions' were interpreted fairly narrowly to comprise: old-age, retirement, widows' & orphans' and war pensions.

Table 20: Income from Other Sources

Source	Code		Cases
	A	E	
Allowances from relatives	A	01	686
Trade Union benefits	B	02	20
Income from charity	C	03	34
Friendly Society benefits	F	04	38
Poor Relief	G	05	1078
Health Insurance benefits	H	06	385
Perquisites	K	07	495
Lodgers	L	08	221
Pensions	P	09	5102
Rent	R	10	5088
Income from Savings	S	11	114
Trading Profits	T	12	36
Unemployment Insurance benefits	U	13	1313
Property income	V	14	26
Workmen's compensation	W	15	35
Private Means	D	16	200
Military Reserve Pay	M	17	29
Odd Jobs	J	18	65
Scholarships, Grants, etc.	E	19	28
Other	Z	98	113

6: Wage Earners' Information (IVE)

- File: IVE Field: IDNO_E = *Household Identification Number*.
 Level: A and E.
 Length: 6 characters.
 Description: See IDNO_H on page 11.
- File: IVE Field: IVNO_E = *Earners number*.
 Level: A and E.
 Length: 2 characters.
 Description: A number assigned on inputting to identify the individual. Always assigned even if only one individual is present. IDNO_E and IVNO_E are needed to identify each earner uniquely.
- File: IVE Field: SEXX_E = *Sex*.
 Level: A and E.
 Length: 1 character.
 Description: At level A, the sex of the individual is entered as 'F' or 'M'. 'X' appears where the sex could not be determined.
 Coding: See Table 21.

Table 21: Sex of Wage-Earners

	Code		Cases
	A	E	
Female	F	0	18018
Male	M	1	31317
Unknown	X	9	110

- File: IVE Field: AGEY_E = *Age*.
 Level: A and E.
 Length: 5 characters at level A; 4 characters at level E.
 Description: Age of the individual in years. At level A, ages in years and months appear with the number of months following a decimal point (e.g. 2.11 means two years and eleven months). The inclusion of months is very rare for wage earners.
 Non-standard entries: see section 2.7, page 6. The commonest non-standard entry is an indication that the individual is an adult. This often appears as an entry such as 'A', or 'Adult', or 'F', or 'Full'. For household 158094, individual 02, AGEY_E erroneously contains the entry 'Nurse'. (At level E, the individual is coded as an adult of unknown age, code 9998.)
 Coding: At level E, the entry is in one-hundredths of years. (Example: 0825 is 8 years and 3 months.) Non-standard entries: 9998 means that there is evidence that the individual was an adult but that no age was entered.

A total of 188 individuals are coded as of unknown age, with code 9999, and 4767 as adult

- File: IVE Field: RELAE = *Relationship to Head*.

Level: A and E.

Length: 15 characters at level A; 2 characters at level E.

Description: See Table 22 on page 25. Relationship to person identified as head of household. In some cases no individual is expressly identified as the head of household but it is almost always possible to make inferences about relationships from other information on the card. In some cases the head is absent; more often the head is identified in other ways (e.g. the two adults in a family may be listed as 'husband' and 'wife' rather than 'head' and 'wife'; a woman living alone may be listed as 'spinster' rather than 'head'). In some cases, the interviewer has entered a name, not a relationship: this has been entered only in the absence of any information to infer the relationship.

As a general rule, the relationships of individuals have been entered as given on the cards; however, exceptions have been made in cases where it is clear that the interviewer has made a habit of misrecording or omitting information about a particular relationship. Sometimes no relationship at all is recorded for the leading male earner who is clearly the head or husband: an appropriate entry has been made. In other cases (often single person households) the relationship, lodger or subtenant, that is recorded against an individual clearly defines a relationship to another household rather than to someone on the card: "head" has been entered. Lastly, some interviewers regularly recorded persons who were the head or husband, as "father", and/or the wife as "mother". Where these are obviously straightforward families consisting of parents with unmarried (usually young) children, the parents have often been entered as husband (or head) and wife.

See the *Companion paper* section 6, p. 15 for additional information about relationships to the head of household.

Coding: See Table 22 on page 25. The values reported in the table are those in the level E files; some grouping of non-standard entries at level A was necessary.

Table 22: Relationship to Head of Household

Relationship	Code	Earners	Non-Earners
Head	01	14330	1416
Wife	02	5836	14409
Son	03	8832	13202
Daughter	04	8399	13194
Husband	05	8041	196
Mother	06	216	612
Father	07	140	101
Widow	08	1035	229
Widower	09	260	7
Sister	10	313	160
Brother	11	235	38
Niece	12	86	57
Nephew	13	71	41
Mother-in-law	14	14	92
Father-in-law	15	23	35
Grand-daughter	16	40	176
Grand-son	17	34	158
Grandmother	18	15	24
Grandfather	19	13	13
Sister-in-law	20	56	34
Brother-in-law	21	82	15
Step-daughter	22	61	54
Step-son	23	76	51
Daughter-in-law	24	6	32
Son-in-law	25	77	0
Aunt	26	7	13
Uncle	27	5	4
Spinster	28	224	4
Bachelor	29	105	2
Cousin	30	11	5
Single	31	174	10
Lodger	32	454	81
Named	33	31	4
Relative	34	20	100
Other	98	47	45
Unknown	99	76	77

- File: IVE Field: RELN_E = *Headship Status*.

Level: E only.

Length: 2 characters.

Description: A code, 01, to identify the head of household, all other entries being zero, 00. For the majority of households in which a head is explicitly identified, RELN_E and RELN_E are the same. Otherwise, RELN_E indicates the head of household. See the *Companion paper*, section 6, p. 15 for details.

- File: IVE Field: BPLC_E = *Birthplace*.

Level: A only.

Length: 30 characters.

Description: Entry for 'Birthplaces of Adults'. This item is particularly problematic because the design of the questionnaire cards did not encourage the interviewers to record information in a clear and consistent way. This not infrequently resulted in problems at the point of data input in interpreting was meant in particular cases.

The practice of individual interviewers varied widely; in many cases just one birthplace is given without information as to whom it applies; in others there is a list of places which are not, or are only partially, attributed to individuals. Some interviewers interpreted 'Birthplaces of Adults' as meaning husband and wife or the older generation only; others interpreted 'adults' as being all persons 21 years and over, or as 18, 16 or 14; and, some noted the birthplaces of every member of the household. The degree of detail given in respect of the places themselves also varies, from a general entry such as 'London' to a specific address.

In computerising the data the main aim has been to achieve consistency in the interpretation of birthplace data for all persons aged 18 years and over. This has been done through the application of certain rules, but also with reference to the practices of individual interviewers. See the *Companion paper* section 7, p. 18 for more information. Limited standardisation has been undertaken of the form in which a place name appears (i.e. correction of spelling mistakes and adopting consistent abbreviations); substantive amendments have not been attempted.

The birthplace information for persons aged under 18 years should be treated with extreme caution; it is both patchy and inconsistent. See the *Companion paper* section 7, p. 18.

- File: IVE Field: BPCC_E = *Birthplace code*.

Level: A and E.

Length: 7 characters.

Description: A numeric code (the same at levels A and E) to represent the birthplace of each adult (person aged 18 years or over) entered under BPLC_E. The birthplaces of persons under 18 years of age, where given, have not been individually coded; their code is: 9999997 =no code. Where BPLC_E provides no information at all about an individual's birthplace, the code 9999998 =no information, appears.

Coding: See section BPLC_B, p. 43 and the *Companion paper* section 12, p. 41.

- File: IVE Field: BPCN_E = *Conurbation code*.

Level: E only.

Length: 1 character.

Description: A numeric code which associates those birthplaces entered in BPLC_E that are geographically part of conurbations. The code in BPCN_E identifies the particular conurbation in which the birthplace is situated. Not all birthplaces are part of conurbations.

Coding: 1 = Greater London; 2 = West Midlands; 3 = Merseyside; 4 = Manchester; 5 = Sheffield; 6 = Leeds and Bradford; 7 = Tyneside; 8 = Clydeside. All other birthplaces are coded 0. See the *Companion paper* section 12.

- File: IVE Field: BPCM_E = *Birthplace match code*.

Level: A only.

Length: 1 character.

Description: A code to represent the quality of the match between the birthplace entered under BPLC_E and the code assigned in BPCC_E.

Coding: D=definite match; N=no match (covering cases in which BPCC_E is 9999997, 9999998 or 9999999); U=match uncertain (i.e. there is significant doubt about the accuracy of the entered code). See the *Companion paper* section 12.

- File: IVE Field: OCCU_E = *Occupation*.

Level: A only.

Length: 30 characters.

Description: Entry for 'Occupation'. There is considerable variation in the way that occupations were recorded on the cards. The same activity may be recorded in several different ways, e.g. sailor, seaman, able-seaman. There has been no attempt to standardise entries in OCCU_E at level A.

- File: IVE Field: OCCC_E = *Occupation code*.

Level: A and E.

Length: 4 characters.

Description: A code to represent the occupation entered under OCCU_E. Each code uniquely identifies a specific group of closely related occupations as defined by an occupational title. The codes and occupational groups are those listed in *Census, 1931: classification of occupations*, HMSO, 1934. Every attempt was made to provide a meaningful code for each person.

See the *Companion paper* section 11, p. 33.

Coding: The first three characters are as recorded in the census.

Fourth character at level A: E=employers; F=foremen; R=assisting relatives; U=unskilled; _= no extra character needed. See CENC_D, p. 40.

Fourth character at level E: 1=employers (E); 2=foremen (F); 3=assisting relatives (R); 4=unskilled (U); 0=space (_). See CODE_D, page 40.

Special codes (level A): _ _ _ X = not in labour force; _ _ _ Y = retired; * _ _ _ = occupation unknown.

Special codes (level E): 9997 =not in labour force (5412); 9998 =retired (981); 9999 =occupation unknown (547). (Numbers in parentheses are the number of recorded occurrences in each category.)

- File: IVE Field: OCCS_E = *Skill level*.

Level: E only.

Length: 1 character.

Description: A code that associates the occupations entered in OCCU_E with a socioeconomic class or 'skill level'.

Coding: 1 =higher professionals, managers and proprietors; 2 =other professionals, managers and employers; 3 =skilled and clerical workers; 4 =semi skilled workers; 5 =unskilled workers; 0 =unknown. See also CLAS_E, page 41.

Table 23: Skill classification

Socioeconomic class	Code	Cases
Higher professionals	1	49
Other professionals	2	308
Skilled and clerical	3	20562
Semi-skilled	4	10068
Unskilled	5	11515
Unknown	9	6943

- File: IVE Field: OCCO_E = *Occupation Order*.

Level: E only.

Length: 2 characters.

Description: A numeric code that associates the occupations entered in OCCU_E with one of 31 main orders, as given in the 1931 Census. The census groups occupational codes into orders largely on the basis of the type of industry or the material used in production.

Coding: see ORDE_D, page 41 and the *Companion paper* section 11, p. 33.

- File: IVE Field: OCCM_E = *Match code*.

Level: A only.

Length: 1 character.

Description: A code that indicates the quality of the match between the occupation entered under OCCU_E and the code assigned in OCCO_E.

Coding: D=definite match; N=no suitable match; A=approximate match (used where the match is not definite, but there is either little doubt that it is best possible one available, or the choice is between closely related alternatives); T=tentative match (used in cases for which significant doubt about the choice of code remains, particularly where the choice is between unrelated alternatives). See the *Companion paper* section 11.

- File: IVE Field: STAT_E = *Employment Status code*.

Level: A and E.

Length: 3 characters at level A; 1 character at level E.

Description: A code determined at the point of data input to identify whether the individual was participating in the labour force and, if so, whether the person was employed or off work for some reason. See Table 24 for permissible code letters.

Space was left for a maximum of three characters to be entered. In all but a small number of cases, a single character was entered. In the remainder, two characters were entered but in every case it was straightforward to resolve these into a single primary category for coding purposes. Some difficulty was encountered where it was obvious that the status of the individual had changed in the days immediately preceding the interview. In these cases, the code should reflect the status in the last complete working week, not the day of the interview.

The most problematic aspect of coding STAT_E arises when the individual became unemployed, or restricted to part-time employment, in the days immediately prior to the interview. If it is judged that the person was in employment during at least part of the last complete week prior to the interview, then STAT_E is coded E, i.e. to qualify as unemployed, they must have been at least one week without work. Dockers habitually in receipt of some unemployment benefit to supplement their casual earnings have been coded as employed, as have those receiving benefit because of enforced short time working.

An attempt has been made to exclude unemployment benefits from 'Earnings Last Week' (EALW_E, page 30) and to include the benefits as a separate other income source. It is possible, however, that unemployment benefits remain as EALW_E in a few cases.

Some interviewers followed the practice of including certain categories of adults as Earners who were not participating in the labour market (e.g. a wife 'at home' or pensioners who were heads of households). There was usually no attempt to reclassify these individuals at the point of data input (hence the appearance of a high proportion of non-participants, code N).

Only workers clearly described as running their own businesses or working for themselves have been coded as self-employed. Casual workers, jobbing tradesmen, *etc.* were coded as employed (numeric code 1) unless they were clearly described as being not in work at all whether due to unemployment, illness, *etc.*

In *Release 2.0* of the data files, STAT_E has been extended to identify individuals for whom the recorded earnings were clearly obtained jointly with one or more other members of the household. (See page 31.) At level A, the third character of STAT_E was changed from a blank to J for those cases in which earnings were combined. For some of these individuals, earnings entries (EALW_E and EAFW_E) appear while for others in the same household earnings are entered as missing. At level E, the code was changed to 6 for employed persons with joint earnings, and to 7 for self-employed workers with joint earnings.

Coding: See Table 24. (Note that the coding was amended for *Release 2.0*. See above.) The values reported in the table are those in the level E files; some grouping of non-standard entries at level A was necessary.

Table 24: Employment Status of Wage-Earners

Status:	Code		Cases
	A	E	
Not in labour force	N	0	6261
Employed	E	1	37543
Self-employed	S	2	2064
Unemployed	U	3	2429
Sick/Incapacitated	I	4	524
On strike	T	5	3
Employed, joint income	E _J	6	84
Self-employed, joint income	S _J	7	107
Unknown and Other	Z,X	9	430

- File: IVE Field: EMPL_E = *Employer*.

Level: A only.

Length: 30 characters.

Description: Entry for 'Employer' name. Note that there is substantial variation in the way employer details were recorded on the cards. The same employer may appear under several different forms, e.g. London General Omnibus Co, or LGOC. No attempt was made to standardise entries for employers' names across the dataset. However, the details of some employers were checked either during data entry or, more particularly, in connection with problematic cases encountered during the occupational coding. In some cases it was possible to correct the form of the employer's name and explanatory notes were made in the REMI_M text field. Non-standard entries: see section 2.7, page 6.

- File: IVE Field: PLWK_E = *Place of Work.*

Level: A only.

Length: 30 characters.

Description: Entry for 'Place of Work'. As with the employer's name, there is considerable variation in the way places were recorded. Non-standard entries: see section 2.7, page 6.

- File: IVE Field: TRAN_E = *Cost of transport.*

Level: A and E.

Length: 10 characters at level A; 6 characters at level E.

Description: Entry for 'Cost of Transport, Weekly'.

Coding: See 2.6 on page 5. A total of 8141 entries are zero, code 000000, 282 are a variant of 'cycles', code 999998, and 22648 are other or unknown, code 999999. A large number of non-standard entries appear, largely because the cost of transport to work was often not easily expressed in money terms.

- File: IVE Field: HRLW_E = *Hours last week.*

Level: A and E.

Length: 8 characters at level A; 5 at level E.

Description: Entry for 'HOURS, Last Week' Note that many variants appear including hours per day and times of work, usually with some abbreviated explanation of how the entry should be interpreted. In particular, a 'D' following the number of hours indicates that the entry is a daily figure. 'Last week' is interpreted as the last complete working week prior to the interview.

Coding: Hundredths of hours per week. (Example: 04850 equals 48 and one half hours.)
Non-standard entries: see section 2.7, page 6.

A total of 13670 cases are unknown, code 99999, and 411 are zero, code 00000.

- File: IVE Field: HRFT_E = *Hours, full time.*

Level: A and E.

Length: 8 characters at level A; 5 at level E.

Description: Entry for 'HOURS, Full Time' Note that many variants appear including hours per day and times of work, usually with some abbreviated explanation of how the entry should be interpreted. In particular, a 'D' following the number of hours indicates that the entry is a daily figure.

Coding: Hundredths of hours per week. (Example: 04850 equals 48 and one half hours.)
Non-standard entries: see section 2.7, page 6.

A total of 13229 cases are unknown, code 99999, and 54 are zero, code 00000.

- File: IVE Field: EALW_E = *Earnings last week.*

Level: A and E.

Length: 10 characters at level A; 6 characters at level E.

Description: Entry for 'EARNINGS Last Week'. See the comments under 'EARNINGS Full Week'. 'Last week' is interpreted as the last complete working week prior to the interview. In some cases it was found that unemployment benefit had been entered under this heading. Where this is believed to have occurred, the amount has been moved to a separate source of income and designated as unemployment benefit. It is possible, however, that a few errors of this sort remain.

For some households, two or more individuals were working collectively (e.g. a shop-keeper and his wife) and their earnings were not separated but entered for just one of them (typically, the head of household). Thus, for some individuals earnings are higher than should be attributed to that person and for others in the same household earnings are entered as missing. No attempt has been made to allocate the earnings among the relevant individuals but the STATE_E code has been modified to allow identification of these cases. (See page 29.)

Coding: See 2.6 on page 5. A total of 12249 cases are unknown (no numerical amount entered), code 999999, and 433 are zero, code 000000.

- File: IVE Field: EAFW_E = *Earnings in a full week.*

Level: A and E.

Length: 10 characters at level A; 6 characters at level E.

Description: Entry for 'EARNINGS Full Week'. See 2.6 on page 5.

In many cases where, for whatever reason, interviewers were unable to obtain information about earnings, an amount was entered subsequently for earnings (full week) and state insurance deductions based, it is presumed, on standard rates of pay. Such information usually appears in pencil. These entries were recorded in the appropriate fields and, in each case, a comment to the effect that the amounts were estimated or entered in pencil was entered as other information.

See above, under EALW_E, for those cases in which earnings were received jointly by two or more individuals.

In cases for which an employed wage earner's hours and earnings were the same in the last week as in a full week, some interviewers followed the practice of entering the amounts just once, in the respective full week field. Where there was clearly no doubt about the status of the person, these amounts were duplicated in the 'last week' fields at the point of entry. Some inconsistency in the data entry may have occurred in this respect although extensive checking has sought to minimise the problem.

Coding: See 2.6 on page 5 and 2.7, page 6.

A total of 8849 cases are unknown (no numerical amount entered), code 999999, and 45 are zero, code 000000.

- File: IVE Field: SIND_E = *State insurance deductions.*

Level: A and E.

Length: 10 characters at level A; 6 characters at level E.

Description: Entry for 'State Insurance Deductions'. Note that several variants appear, for example, references to superannuation payments.

Coding: See 2.6 on page 5 and 2.7, page 6.

A total of 13743 are unknown, code 999999, 2486 are zero, code 000000, 3512 are given as 'state' (or other indication of normal payment), code 999998, 40 are given as 'health insurance deductions only' code 999997.

7: Non-Wage-Earners Information (IVN)

The IVN file is intended to record information for all individuals identified as non-wage-earners on the cards. It should be noted, however, that certain interviewers sometimes recorded as non-wage-earners adults who would normally participate in the labour market but were temporarily not doing so because of unemployment, illness, temporary disability, *etc.* In some cases there is evidence of state benefit being received. An attempt has been made to identify these individuals and to reallocate the relevant information to the IVE file. Inevitably, some ambiguity remains about whether some individuals were unemployed (in the labour force) or non-participants (not in the labour force). Unless there was good reason to do otherwise, in doubtful cases individuals were left with the classification that appears on the record card.

- File: IVN Field: IDNO_N = *Household Identification Number.*
Level: A and E.
Length: 6 characters.
Description: See IDNO_H on page 11.
- File: IVN Field: IVNO_N = *Individual Number.*
Level: A and E.
Length: 2 characters.
Description: A number assigned on inputting to identify the individual. IDNO_N and IVNO_N are needed to identify each non-earner uniquely.
- File: IVN Field: SEXX_N = *Sex.*
Level: A and E.
Length: 1 character.
Description: Sex of the individual entered, at level A, as 'F' or 'M'. 'X' appears where the sex could not be determined.
Coding: See Table 25.

Table 25: Sex of Non-Wage-Earners

	Code		Cases
	A	E	
Female	F	0	30240
Male	M	1	14413
Unknown	X	9	38

- File: IVN Field: AGEY_N = *Age.*
Level: A and E.
Length: 5 characters at level A; 4 characters at level E.
Description: Age of the individual in years. The letter 'A' represents an adult where no age is entered. Ages reported in years and months are entered with the number of months following a decimal point (e.g. 2.11 means two years and eleven months).

Non-standard entries: The commonest non-standard entry is an indication that the individual is an adult. This often appears as an entry such as 'A', or 'Adult', or 'F', or 'Full'

Coding: One-hundredths of years. (Example: 0825 is 8 years and 3 months.) Non-standard entries: 9998 means that there is evidence that the individual was an adult but that no age was entered.

A total of 148 individuals are coded as of unknown age, with code 9999, and 3593 as adult with code 9998.

- File: IVN Field: RELA_N = *Relationship to Head.*

Level: A and E.

Length: 15 characters at level A; 2 characters at level E.

Description: Relationship to person identified as head of household.

The head of household is frequently entered as a wage earner (even when it is clear that the individual is not in the labour force).

In some cases, the interviewer has entered a name, not a relationship: this has been entered in the absence of any information to infer the relationship.

See the description of RELA_E, p. 24 and the *Companion paper* section 6, p. 15 for additional information about relationships to the head of household.

Coding: See Table 22 on page 25.

- File: IVE Field: RELN_N = *Headship Status.*

Level: E only.

Length: 2 characters.

Description: A code, 01, to identify the head of household, all other entries being zero, 00. For the majority of households in which a head is explicitly identified, RELA_N and RELN_N are the same. Otherwise, RELN_N indicates the head of household. See the *Companion paper*, section 6, p. 15 for details.

- File: IVN Field: BPLC_N = *Birthplace.*

Level: A only.

Length: 30 characters.

Description: Entry for 'Birthplaces of Adults'. See the description of BPLC_E on page 26 and the *Companion paper*, section 18 for more information about the interpretation of birthplace entries.

- File: IVN Field: BPCC_N = *Birthplace code.*

Level: A and E.

Length: 7 characters.

Description: A numeric code to represent the birthplace entered under BPLC_N. Each complete code uniquely identifies a country, a county or a specific place. See the comments for BPCC_E on page 26.

Coding: See section BPLC_B, p. 43 and the *Companion paper* section 12, p. 41.

- File: IVN Field: BPCN_N = *Conurbation code.*

Level: E only.

Length: 1 character.

Description: A numeric code which associates those birthplaces entered in BPLC_N that are geographically part of conurbations. The code in BPCN_N identifies the particular conurbation in which the birthplace is situated. Not all birthplaces are part of conurbations.

Coding: 1 = Greater London; 2 = West Midlands; 3 = Merseyside; 4 = Manchester; 5 = Sheffield; 6 = Leeds and Bradford; 7 = Tyneside; 8 = Clydeside. All other birthplaces are coded 0 for BPCN_N. See the *Companion paper* section 12.

- File: IVN Field: BPCM_N = *Birthplace match code*.

Level: A only.

Length: 1 character.

Description: A code to represent the quality of the match between the birthplace entered under BPLC_N and the code assigned in BPCN_N.

Coding: D=definite match; N=no match (covering cases in which BPCN_N is 9999997, 9999998 or 9999999); U=match uncertain (i.e. there is significant doubt about the accuracy of the entered code). See the *Companion paper* section 12.

8: Supplementary Cards (CDX)

This table is provided at level A only.

- File: CDX Field: IDNO_C = *Household Identification Number*.

Level: A only.

Length: 6 characters.

Description: See IDNO_H on page 11.

- File: CDX Field: CARD_C = *Supplementary Card number*.

Level: A only.

Length: 5 characters.

Description: In addition to the principal card number, CARD_H, referred to above, one or more extra card numbers appear on some cards. It is thought that the extra numbers were intended to be the principal card numbers of *other* households dwelling at the same address. Their presence on a card indicates a relationship with that household. The relationship may be that of main tenant to a subtenant (or *vice versa*), or joint tenants, paying rent to an outside landlord. See the *Companion paper* section 8.

Non-standard entries in CARD_C are listed in Table 26.

Table 26: Non-standard card numbers

IDNO_C	CARD_C
056052	(0100
056052	MC)
102104	M0655
107003	0602
168118)793
186018	B
235025	&

9: Free Format Text Fields (MEM)

This table appears at level A only. See sections 2.4 (page 4) and 2.5 (page 4). Note, in particular, that the text fields include only printable characters with two exceptions: (i) the pound sign, £, ASCII character 156, and (ii) soft carriage return, ASCII character 141.

- File: MEM Field: IDNO_M = *Household Identification Number*.

Level: A only.

Length: 6 characters.

Description: See IDNO_H on page 11.

- File: MEM Field: REMG_M = *General Remarks*.

Level: A only.

Description: Contains information written under 'GENERAL REMARKS, Etc.' on the back of each card. As far as it is legible, the information has been recorded as written on the cards.

- File: MEM Field: REMA_M = *Accommodation Remarks.*

Level: A only.

Length: variable.

Description: Contains information written under 'Remarks on Accommodation' on the back of each card. As far as it is legible, the information has been recorded as written on the cards. Non-standard entries, for example pantry = cupboard, are copied from the card.

- File: MEM Field: REMI_M = *Information Remarks.*

Level: A only.

Length: variable.

Description: Contains information written on each card that is not attributable to any other field. It is used in particular, to record additional details about household composition, earnings and income which appear on the front of the cards. As far as is legible, the information has been recorded as written on the cards. Also, the results of various (post-interview) calculations, which have been written on the lower right hand portion on the front of each card, may be referred to in REMI_M. These calculations have not been recorded in the database, but the information they contain has sometimes been used by the inputter in making inferences about income and rent.

10: Street Quality Codes (STQ)

This table appears at level A only. See the *Companion paper* section 10, p. 28 for a more detailed explanation of coding principles.

- File: STQ Field: IDNO_Q = *Household Identification Number.*

Level: A only.

Length: 6 characters.

Description: See IDNO_H on page 11.

- File: STQ Field: QUAL_Q = *Street Quality Code.*

Length: 2 characters.

Description: a numeric code representing the economic and social classification of the street (in ADDR_H) in which the household identified in IDNO_Q, was living. At least one code is applied to each address.

The coding is drawn from the results of the *Street Survey* (conducted as part of the *New Survey*) which classified the population on a street by street, rather than individual household, basis according to a series of broad socioeconomic grades. Its results were published as a series of coloured street maps (*New Survey* vols. IV & VII).

Where more than one code is given for a household, it indicates that different classifications applied in different parts of the street in which it lived, as shown by the colouring on the maps; but, as the maps do not specify house numbers, a specific code could not be determined in such cases. However, the predominance order code (see below, PRED_Q)

There are three types of special case, as indicated by a value entered in SQMC_H (see page 13): (a) "adjacent" streets in which the code (or codes) entered in QUAL_Q applies not to the first named street, but to the one in whose locality it lies; (b) "combination" streets in which the code (or codes) entered applies to more than one street; and (c) name changes in which the street name in ADDR_H was found to be completely different from that on the *New Survey* map. The special cases reflect problems encountered during the coding process. These are chiefly cases where the street in question existed but could not be identified on the *New Survey* map or where the address information given on the card is insufficient to allow a precise entry to be made in ADDR_H.

Where the street appears on the *New Survey* map but is uncoded, QUAL_Q= 98; where it cannot be identified on the map or in any other source, QUAL_Q= 99.

Coding: See Table 27 on page 37.

Table 27: Street Quality Codes

QUAL_Q	Colour on map†	Cases‡	Total‡
01	black (criminals)	90	148
02	blue (poor)	139	443
03	purple (unskilled)	7150	9064
04	pink (skilled)	13581	15922
05	red (middle class)	1391	2580
21	blue with black stripe	231	303
31	purple with black stripe	381	574
41	pink with black stripe	98	189
32	purple with blue stripe	1564	2223
45	pink with red stripe	2318	3508
61	purple with blue and black stripes	320	407
98	uncoded	123	126
99	unknown street	180	180

† See *Companion paper*, p. 29, for an interpretation of composite colours.

‡ See PRED_Q for an interpretation.

- File: STQ Field: PRED_Q = *Predominance Order*.

Level: A only.

Length: 1 character.

Description: An indicator of the predominance of the street quality code for the street in which the household was living. For some streets, more than one code was entered (see above, QUAL_Q). In these cases, PRED_Q reflects the position of the code in an ordering of importance.

Coding: A numeric value. 1 indicates the code which occurs in the greatest area of the street. When only a single QUAL_Q value appears, it is automatically coded 1. Ties are permitted and are coded with the same value. For example, a street might be associated with four QUAL_Q codes, 21, 03, 61 and 02. Suppose that the PRED_Q entries are 1, 2, 2 and 3. This indicates that 21 occurs in the greatest area of the street, followed by 03 and 02, and 01.

correspond to a smaller but approximately equal area; and 02 to the smallest area for the street in question.

Interpretation of the 'Cases' and 'Total' columns in Table 27. In the 'Cases' column, the numbers refer to the number of *households* in a street with the given QUAL_Q value *and* for which PRED_Q equals 1. The aggregate of 'Cases' across codes exceeds the total number of households as a result of ties. The values in the 'Total' column refer to the number of times that the code appears in the file (i.e. irrespective of PRED_Q).

11: Workplace Codes (WPL)

- File: WPL Field: IDNO_W = *Household Identification Number*.
Level: E.
Length: 6 characters.
Description: See IDNO_H on page 11.
- File: WPL Field: IVNO_W = *Individual Number*.
Level: E.
Length: 2 characters.
Description: A number assigned to identify the individual. Both IDNO_W and IVNO_W are needed for a unique identification of each earner.
- File: WPL Field: WPLC_W = *Workplace Code*.
Level: E.
Length: 7 characters.
Description: A code assigned to identify the geographical location of the individual earner's workplace. The coding scheme is the same as that for birthplaces. Consequently, the borough is the most detailed level of location provided. Of the 49445 earners, 38042 were assigned a location code, the remaining 11403 being assigned the code 9999999.
Coding: See section BPLC_B, p. 43 and the *Companion paper* section 12, p. 41.
- File: WPL Field: ADCD_W = *Adjacency Code*.
Level: E.
Length: 7 characters.
Description: The adjacency code identifies the geographical relationship between the borough of residence and borough of employment.
Coding: One of six codes was assigned to each earner as shown in Table 28 on page 39.

Table 28: Workplace Adjacency Codes

Code	Number of cases	% of all matches	Description
1	16434	43.2	Residence and workplace in the same borough
2	10235	26.9	Residence and workplace in adjacent boroughs
3	9917	26.1	Residence and workplace in non-adjacent boroughs within the London conurbation (as defined in the birthplace coding)
4	1053	2.8	Employment all over London
5	403	1.1	Employment outside London (mostly seamen)
9	11,403	n/a	No identified workplace (i.e. code 9999999)

12: Interview Dates (DTE)

- File: DTE Field: IDNO_D = *Household Identification Number*.
Level: E.
Length: 6 characters.
Description: See IDNO_H on page 11.
- File: DTE Field: DATE_D = *Interview Date*.
Level: E.
Length: 6 characters.
Description: The value of DATE_D is the same as DATE_H (in HHG) where the *year* appears in DATE_H. (See DATE_H, page 14.) Where no year was written on the card at character positions 5 and 6 in DATE_H appear as 00. For these records, DATE_D differs by including an estimate of the year in place of 00 for characters 5 and 6. (See the *Companion paper*, page 48.)
- File: DTE Field: AMND_D = *Amendment Quality Code*.
Level: E.
Length: 1 characters.
Description: A code assigned to indicate the confidence with which the year of interview was estimated. (See the *Companion paper*, page 48.)
Coding: One of four values was assigned as set out in Table 29.

Table 29: Amendments to Interview Dates

Code:	Cases:	Interpretation:
0	16590	No change to year
1	9296	Year inserted (reliable)
2	1024	Year inserted (unreliable)
9	5	Year doubtful

13: Occupation Codes (OCC)

This *look-up* table of codes appears at level A only. See the *Companion paper* section 11, beginning page 33 for a more detailed explanation of coding principles and scheme.

- File: OCC Field: CENC_0 = *Modified Census Code.*

Level: A only.

Length: 4 characters.

Description: A code made up of three digits and additionally, in some cases, a letter, that uniquely identifies a specific group of occupations as defined by an occupational title (see below DESC_0). The codes and occupational groups are those listed in *Census, 1931: classification of occupations*, HMSO, 1934, but with some modifications in the way the codes have been used. Each occupational group represents a number of closely related occupations which are individually listed in the Census. The occupation is also assigned to one of 31 main orders, largely on the basis of the type of industry or the material used in production, and also is assigned one of five socioeconomic classes or "skill levels" (see below, ORDE_0 and CLAS_0).

Where the fourth character of the code is a letter (i.e. not a space), it denotes a certain additional category of job, and may indicate an alternative occupational grouping or skill level (see CLAS_0) to the one which would be represented if the same first three characters are followed by a space. *Examples:* (i) 528_ = "other skilled workers (rubber)", 528U=unskilled workers in the rubber industry (e.g. a labourer in a tyre company); (ii) 613_ = "haulage and cartage contractors and managers", but denotes a small contractor who is likely to be working alone from home, and 613E="haulage and cartage contractors and managers" denotes one who has a substantial business.

One of the following letters may appear as the fourth character: E, F, R, U. They refer to: E=employer, proprietor or manager (defined as someone who employs others, runs premises that are separate from the home, or manages other people); F=foreman or supervisor of other people; R=assisting relatives (defined as someone who is employed by a member of the family); U=unskilled occupation (used only where the Census classification scheme does not provide a specific code for the occupation, other than 920, 930 or 940 (see below)). A numeric code may appear both with or without the fourth character meaning something different in each case.

To determine the appropriate code, each occupation appearing in OCCU_E was matched against the occupations listed in the census and the best possible representation was obtained. In cases of doubt a detailed set of rules for assigning a code was applied. Where it was impossible to find a meaningful code, one of the following codes denoting "no occupation", was used: Y (preceded by three spaces), _ _ _ Y=retired; X (preceded by three spaces), _ _ _ X=not in the labour force (e.g. wives at home); *_ _ _ =unknown.

- File: OCC Field: CODE_0 = *Numeric code identifying occupation.*

Level: A only.

Length: 4 characters.

Description: A wholly numeric code that uniquely identifies a group of occupations as defined by an occupational title (see below DESC_0). CODE_0 is used to represent the coded occupations in the IVE(E) file OCCU_E. The codes map one-to-one onto those in CENC_0. The first three characters (digits) are the same in each case. The fourth character is a numeric representation of the additional letters that appear in some CENC_0 codes.

Coding: Fourth characters: 0=no character (i.e. a space); 1=employers (E); 2=foremen (F); 3=assisting relatives (R); 4=unskilled (U)

Special codes: 9997=not in labour force; 9998=retired; 9999=occupation unknown.

- File: OCC Field: CLAS_0 = *Skill level*.

Level: A only.

Length: 1 character.

Description: a socioeconomic class or "skill level" assigned to each occupational grouping represented by a separate code in CENC_0 or CODE_0. The skill coding is based on the classifications given in *The Registrar-General's Decennial Report*, 1931, HMSO, Table 1, but with some modifications in the way they have been assigned in particular cases, notably occupations concerned with shop work, and where the fourth character denotes an employer, foreman, assisting relative, or unskilled occupation.

Coding: 1 =higher professionals, managers and proprietors; 2 =other professionals, managers and employers; 3 =skilled and clerical workers; 4 =semi skilled workers; 5 =unskilled workers; 9 =no code.

- File: OCC Field: ORDE_0 = *Order*.

Level: A only.

Length: 2 characters.

Description: Each occupational grouping represented by a separate code in CENC_0 or CODE_0 is assigned to one of 31 main orders, as given in the 1931 Census. ORDE_0 identifies the order. The Census groups individual occupational codes into orders largely on the basis of the type of industry or the material used in production.

Coding: See Table 30 on page 42.

Table 30: Occupation Order Codes

Code	Cases	Occupation Order
01	0	Fishermen
02	138	Agricultural occupations
03	0	Mining and quarrying occupations
04	57	Workers in the treatment of non-metalliferous mine and quarry products
05	167	Makers of bricks, pottery and glass
06	384	Workers in chemical processes; makers of paint, oils, <i>etc</i>
07	3216	Metal workers (not electro plate or precious metals)
08	89	Workers in electro plate or precious metals
09	966	Electrical apparatus makers and fitters (not elsewhere enumerated) and electricians
10	94	Makers of watches, clocks, and scientific instruments
11	497	Workers in skins and leather, makers of leather and leather substitute goods (not boots or shoes)
12	163	Textile workers
13	3523	Makers of textile goods and articles of dress
14	1222	Makers of foods, drinks and tobacco
15	1974	Workers in wood and furniture
16	658	Makers of and workers in paper and cardboard; bookbinders, <i>etc</i>
17	1001	Printers and photographers
18	1807	Builders, bricklayers, stone and slate workers; contractors
19	1206	Painters and decorators
20	272	Workers in other materials
21	451	Workers in mixed or undefined materials (not elsewhere enumerated)
22	7413	Persons employed in transport and communications
23	3902	Commercial, finance and insurance occupations (excluding clerks)
24	457	Persons employed in public administration and defence (excluding professional men, clerical staff and typists)
25	268	Professional occupations (excluding clerical staff)
26	163	Persons professionally engaged in entertainments and sport
27	4974	Persons engaged in personal service (including institutions, clubs, hotels, <i>etc</i>)
28	2172	Clerks and draughtsmen; typists
29	1823	Warehousemen, storekeepers and packers
30	554	Stationary engine drivers, dynamo and motor attendants
31	2891	Other and undefined workers
32	0	Retired or not gainfully employed — not used (Assigned code 99)
99	6943	Unknown occupation

- File: OCC Field: DESC_0 = *Description of occupation.*
Level: A only.
Length: 50 characters.
Description: The formal title of the occupational grouping attached to each code in CENC_0 and CODE_0 as it appears in the 1931 census.
- File: OCC Field: COUN_0 = *Number of occurrences.*
Level: A only.
Length: 5 characters.
Description: The number of occurrences in the IVE file of each unique occupational code.

14: Birthplace Codes (BPL)

This *look-up* table of codes appears at level A only. See the *Companion paper* section 12, p. 41 for a more detailed explanation of coding principles and scheme.

- File: BPL Field: BPLC_B = *Code identifying place name.*
Level: A only.
Length: 7 characters.
Description: A numeric code that uniquely identifies a country, county or specific place. The coding scheme is hierarchical. It is structured by country and, where applicable (see below), by county, and is based on the international boundaries and national administrative units that existed in 1931, or as close thereto as possible.

Each code is made up of three parts. The first two characters represent the country; characters three and four are the county; and the three righthand characters are the specific place name. All seven characters are needed to identify uniquely each place name or geographical area; the use of zeros for characters 5, 6 & 7 and for 3 & 4 allows an area, either county or country, as opposed to a specific place, to be represented. *Examples:* 0100000 =England; 0101000 =Bedfordshire; 0101002 =Bedford; 0600000 =Eire; 0612000 =Dublin County; 0612001 =Dublin City; 0000000 =Great Britain; 1800000 =France.

There are a few codes that have special meaning: the code 900 for characters 5, 6 & 7 indicates a rural area within the United Kingdom and Eire (see below). *Examples:* 0113900 =Essex rural areas, 0100900 =English rural areas ("the country"); the code 9000000 indicates "abroad" (no country specified); 9100000 indicates Jewish origin (no country specified); 9999998 indicates that there is no information on the card; and 9999997 that there is some information but no code is assigned.

Individual counties and place names are coded only for the countries of the United Kingdom: England, Wales, Scotland and Northern Ireland, and for Eire. These countries occur at the beginning of the scheme. A unique code is provided for each county borough, metropolitan or municipal borough and urban district. (In Scotland, the analogous areas are large and small burghs.) Rural districts (landward areas in Scotland) are not individually coded, but are represented by a single general code (900) within each county and country. Place names and area names outside the United Kingdom and Eire are covered by the appropriate country code.

- File: BPL Field: CONU_B = *Code identifying conurbation.*

Level: A only.

Length: 1 character.

Description: an additional numeric code assigned, where appropriate, to the place codes in BPLC_B, to associate them with conurbations. Only those codes in BPLC_B representing specific places that are geographically part of a conurbation, are assigned a conurbation code. CONU_B identifies the conurbation.

A conurbation is defined as a single very large urban area or city (population of 600,000 or over), usually identified by one particularly large county borough, in which a number of county boroughs, metropolitan boroughs and urban districts are contiguous, i.e. without rural districts in between. The conurbation code allows those contiguous boroughs and urban districts to be linked. (In Scotland, the appropriate areas are large and small burghs.)

Coding: Eight conurbations are identified: 1 = Greater London; 2 = West Midlands (Birmingham and Wolverhampton); 3 = Merseyside (Liverpool and Birkenhead); 4 = Manchester (including Salford and Stockport); 5 = Sheffield (including Rotherham); 6 = Leeds and Bradford; 7 = Tyneside (Newcastle, Gateshead, North and South Shields); 8 = Clydeside (Glasgow). All places outside of these areas are coded 0.

- File: BPL Field: DESC_B = *Title of place name.*

Level: A only.

Length: 50 characters

Description: the formal place name title attached to each place or area code. The name may be that of a country, a county or a specific place. In the case of United Kingdom and Eire counties and places, the form of name used is that which was in official usage in 1931, as given in the sources indicated above (see BPLC_B).

- File: BPL Field: COUN_B = *Number of occurrences.*

Level: A only.

Length: 5 characters.

Description: The number of occurrences in the IVE and IVN files of each unique birthplace code.

SPECIMEN OF HOUSEHOLD INVESTIGATION CARD

(a) Front of Card.

Name A— B— Address C Street
 Borough Barking

LONDON SCHOOL OF ECONOMICS, UNIVERSITY OF LONDON. Social Survey of London.
 A 0847 File No. 157

WAGE-EARNERS (including unemployed)						HOURS		EARNINGS				State Insurance Deductions	
Relationship	Age	Occupation	Employer	Place of Work	Cost of Transport Weekly	Last Week	Full Time	Last Week		Full Week		s.	d.
424 Husband	38	Labourer	E Company	F Street	2/6	32	50	s. 44	d. —	s. 55	d. —	s. 1	d. 4
NON-WAGE EARNERS	Sex ...	Age ...	Relationship	F	F	M	M						
				31	7	6	2						
				Wife	Daughter	Son	Son						
INCOME FROM OTHER SOURCES		From sub-letting 2 rooms		4/1½		M	44/-	55/-		8/3			
						wc ₂ in	36/1	47/1		4/1			
							32/1	32/1		5 . 3½		2 . 1	
Date.		7.1.30		Interviewer.		X—Y—							

(b) Back of Card.

GENERAL REMARKS, Etc.

Birthplaces of Adults *Husband—Barking*
Wife —Shoreditch

RENT, Weekly 8/3

(Including Rates. Receipts from Sub-tenants not deducted)

PERSONS AND ACCOMMODATION
 (This refers to family as on face of card)

No. of Persons 5
 „ „ Bedrooms 1
 Parlour No
 Kitchen 1
 Scullery Yes
 Pantry or Larder No
 Bath No
 Yard Yes
 Garden No
 Allotment No

Remarks on Accommodation
House in fair repair
Rooms rather small