The British Business Census of Entrepreneurs 1851-1911 (BBCE)

User Guide

Robert J. Bennett, Harry Smith, Carry van Lieshout, Piero Montebruno and Gill Newton

University of Cambridge

Department of Geography, and Cambridge Group for the History of Population and Social Structure, Downing Place, Cambridge, CB2 3EN, UK

January 2020

The British Business Census of Entrepreneurs 1851-1911 (BBCE): User Guide

Robert J. Bennett, Harry Smith, Carry van Lieshout, Piero Montebruno and Gill Newton

University of Cambridge

Contents

1. Introduction

The British Business Census of Entrepreneurs 1851-1911 (BBCE) is a major output from the ESRC-supported project ES/M010953: '*Drivers of Entrepreneurship and Small Businesses*', Principal Investigator Professor Bob Bennett, the authors of this *Guide*, and other researchers and assistants (listed in full in the Acknowledgements). Additional support from other funders came from the Leverhulme Trust, and Isaac Newton Trust.

The aim of the database is to provide to users an quality-controlled, consistent, and integrated database that is as complete as feasible for all business proprietors in the period 1851-1911 and aligned to modern definitions used up to the 2011 census. This includes as fully as possible all:

- o   Employers (those who employed others),
- o   Sole proprietor own account self-employed (those who employed no-one else),
- o   Directors of corporations (Limited Companies).

This is the de facto population of the self-employed. This *User Guide* gives the detailed specification of the information contained in the BBCE database.

The data and coding have been developed through a series of complex decisions. It is important that users familiarise themselves with the decisions involved so that the definitions and origin of the data are understood and thus mistakes of interpretation avoided. The background to how the decisions were made is contained in Working Papers and other publications, which are listed at the end of this *Guide* and on the BBCE website. Working Papers are referred to by their number; e.g. WP 1 (which outlines the project's database objectives and summarises how the various sources of information were assembled). The Working Papers are available as downloads, and several also have downloads of supplementary data, definitions used, and working material that should help other users. A summary of the main outputs from initial analysis of the database is given in the book-length study: *The Age of Entrepreneurship: Business proprietors, self-employment and corporations since 1851*.[1]

This Guide will be updated and extended as required in future editions: suggestions and any corrections are welcome; please send to rjb7@cam.ac.uk.

The BBCE website, which also gives links to all working papers, additional data downloads, and updates, is located at:

www.bbce.uk

2. Census and other data sources in BBCE

The main database derives mainly from the original population census manuscript records held by The National Archives (TNA) that have been transcribed by various genealogy providers. The manuscripts are the Census Enumerators Books (CEBs) for 1851-1901, and the individual household records for 1911. Chiefly the records derive from the transcriptions by FindMyPast (FMP), which were encoded into fields of variables by the *Integrated Census Microdata (I-CeM) Project*, which was also supported by the ESRC (grant RES-062-23-1629) and is available from UKDA.[2]

I-CeM version 1 is the main source for BBCE, which has then been supplemented and enriched. It has also been improved by numerous corrections undertaken by collaborators at the Cambridge Group for the History of Population and Social Structure Campop (Campop) based at the University of Cambridge in the Department of Geography and Faculty of History, by the authors, and especially by Alice Reid's team working on the *Atlas of Victorian Fertility Decline* (ESRC grant ES/L015463/1) with support from Kevin Schürer. This means that the BBCE uses the *Campop version* of I-CeM v.1. The effects of this for users of this *Guide* are summarised in Section 6 below. The BBCE database is aimed at integrating with I-CeM as far as possible in current and future versions.

---

[1] Robert Bennett, Harry Smith, Carry van Lieshout, Piero Montebruno and Gill Newton (2019) *The Age of Entrepreneurship: Business proprietors, self-employment and corporations since 1851*, Routledge, London.
[2] Higgs, Edward and Schürer, Kevin (University of Essex) (2014) *The Integrated Census Microdata (I-CeM)* UKDA, SN-7481. http://doi.org/10.5255/UKDA-SN-7481-1

In addition to the FMP/I-CeM database, the BBCE also uses the genealogy provider S&N:The Genealogist as a source to infill gaps in FMP/I-CeM records for 1851 (about 55,000 individuals or 25% of employers in England and Wales) and 1871 (all England and Wales individuals) (see WP 3 and WP 12). For 1861 transcription gaps in FMP occupation strings were infilled by the project team and assistants using searches of the original census manuscripts of the CEBs (about 35,000 individuals in England and Wales for 1861; see WP 1 and 3). For Scotland similar FMP/I-CeM truncations occur, with 8,000 individuals infilled from the original CEBs across all years (see WP 20). BBCE also amends I-CeM codes extensively for occupations, transcriptions with split lines, and other technical issues (see below). This means that BBCE differs from I-CeM in content and is more complete and accurate.

In addition, data linkage and enrichment were developed to code those individuals identifiable in the census database who were directors of corporations, together with information on their companies (about 18,000 individuals; see WP 14).

I-CeM data for BBCE were used under special licence from UKDA for names and addresses. Other users requiring access to names and addresses in BBCE or I-CeM should apply to UKDA for access to the I-CeM database. Access to S&N names and addresses of employers for 1871, if needed for research purposes, may be made available under a restrictive special licence by direct request to the PI.

3. Relationship between BBCE and I-CeM

To facilitate users, the records in BBCE can be linked with those in I-CeM through the individual unique person identifier (RecID) for each person available in BBCE that is also in I-CeM. This is the chief method that users should employ to work smoothly with the records in I-CeM and BBCE. However, some people identified in the BBCE data are additional to those in I-CeM (some in 1851; and for all BBCE for 1871 England and Wales). They have been given additional RecIDs unique to BBCE. For these additional BBCE people their attributes (gender, RELA, etc.) have been coded as far as possible equivalent to I-CeM codes, but users should be aware that these individuals do not have equivalent codes for all variables, especially for 1871. Further information and downloads are given for additional individuals in WP 12 and WP 21.

Where people in BBCE are identical to those in I-CeM only the additional data coding attached to people as entrepreneurs is provided in BBCE. Users will need to link this to I-CeM to be able to obtain the rest of the census information on each person (including names and addresses). The two databases are distinct for reasons of intellectual property and data licencing, but also because future updates in I-CeM or BBCE would otherwise have to be coordinated which would probably be impossible to achieve. The aim of both I-CeM and BBCE is to focus on maintaining a fixed link through the RecID so that, after any future change to the databases, users will be still be able to continue linking the same people between the two sources. We are very grateful to Eddy Higgs and Kevin Schürer for advice in these developments and for a licence to include some information in BBCE from I-CeM where this will help users to avoid confusion in some data fields. Where possible information obtained by the BBCE team that is useful to I-CeM has been licensed to Higgs and Schürer for inclusion in updated versions of I-CeM (for example

corrections of age, gender, and occupation codes of some individuals). However, to maintain replicability of our published research by future users against the same data version we have used, any supplementary data coding not in I-CeM Version 1 is also included in BBCE (and is identified separately) or is given in supplementary downloads. Additional downloads are noted under the variable descriptions and listed at the end of this *Guide*. While new versions of I-CeM may maintain exact matches of supplementary codes for each individual, this is not guaranteed, so that users of BBCE who are focused on the entrepreneur population will generally find BBCE a preferable source for these variables.

Users should be aware that the occupational coding in BBCE and I-CeM differs in some significant respects: BBCE gives codes for the main or first listed occupation, whereas I-CeM has tried to maintain a match to the coding conventions used in census *publications*. Coding for census publications usually gave precedence to farming over other occupations.[3] This affects how those with multiple occupations are coded; e.g. a '*cotton mill proprietor with 1,000 employees and farmer of 4 acres with 2 men*' would be coded to farmer by GRO clerks; and generally this would be followed by I-CeM. The I-CeM coding algorithm looked for farmers first, coded them, and for these people did not consider any other occupations further. This followed the GRO coding of dual occupations where farming always took precedence, regardless of it being mentioned first or last.[4]

In BBCE precedence has been given to the first entrepreneur occupation stated (which would be *cotton mill proprietor* in the above example). BBCE coding has generally followed the census instructions rather than census clerical coding - that the first occupation listed should be the main one. Additionally in BBCE for 1881 alone, in those few cases where two or more occupations each had other information such as employee numbers, this was used to order occupations in BBCE downward in order of importance or by workforce size (but this is usually the main of first occupation in any case). In addition, all portfolios identified in BBCE are given additional multiple codes, which are ordered by main, secondary, and so on. In I-CeM only one occupation code is given and this is not always the main one. Apart from business portfolios, many I-CeM occupation codes identify rank rather than economic status. BBCE gives precedence to business occupations, so that I-CeM coding has been replaced where this was not the main business activity for those of 'rank' (e.g. M.P., Lord, Alderman, etc. who had a main business activity), or non-entrepreneur occupations (e.g. workers such as postmasters who were employees of the Post Office, but had another main entrepreneurial activity)(see WP 1, WP 3, WP 4, WP 5).[5]

For workers, corrections to occupational coding in I-CeM were also made. The workers are not included in BBCE. However, workers are identified separately from entrepreneurs in downloads which allow the estimates of the whole economically active population to be used, usuing the

---

[3] See e.g. TNA, RG27/5, 'Instructions to clerks employed in classifying occupations and ages of the people', 1881, p. 3.
[4] Confirmed by advice from Kevin Schürer and Eddy Higgs.
[5] See Bennett, R.J., Smith, H., van Lieshout, C., Montebruno, P., Newton, G., 2019. *The Age of Entrepreneurship: Business proprietors, self-employment and corporations since 1851*, Chapter 11. Routledge International Studies in Business History, London and New York. http://doi:10.4324/9781315160375. Also see Radicic, D., Bennett, R.J. and Newton, G., 2017. Portfolio entrepreneurship in farming: Empirical evidence from the 1881 census for England and Wales. *Journal of Rural Studies*, 55, 289-302. DOI: 10.1016/j.jrurstud.2017.08.019

correct occupation codes. These downloads are given with WPs 9 and 9.2 for England and Wales, and WP 20 for Scotland. Also see below, under the variable: ENTOCCODE.

It is important that users read the I-CeM *User Guide* as well as this *BBCE Guide* and Working Papers to fully understand the detailed specification of the original census data, how it was coded in I-CeM and enriched, and how BBCE has then processed, coded, corrected, supplemented, and further enriched these data. BBCE takes I-CeM as the point of departure and I-CeM definitions are not repeated in this Guide unless essential for understanding. The specific detail of how the census data were collected for entrepreneurs is summarised in WP 2, while WP 3 and WP 4 for England and Wales, and WP 20 for Scotland, outline the detail of how BBCE identifies each field and infills the FMP/I-CeM gaps.

Summary of differences between BBCE and I-CeM

(see WP 1, WP, 3, WP 4, WP 12, for England and Wales, and WP 20 for Scotland). BBCE differs from I-CeM as follows:

- o Recodes and corrects occupations to economic active status, and where necessary orders multiple occupations to focus on economic activity and main occupation of entrepreneur activities
- o Adds occupational codes for each part of a portfolio of multiple entrepreneur activities
- o Corrects several million other occupation codes; those for entrepreneurs included in BBCE, those for workers are available as a separate downloads attached to WPs 9,9.2 and 20; other changes will be included in I-CeM Version 2
- o Adds entrepreneur data for 1871
- o Fills up truncations and gaps in I-CeM for
  - 1851: 55,000 records in England and Wales
  - 1861: 35,000 records in England and Wales
  - 1881: minor truncations
  - 8,000 records for 1851-81 in Scotland
- o Identifies several thousand transcriptions of occupational lines for entrepreneurs split between two or more individuals, and recodes to correct individual
- o Identifies all employers and masters 1851-81
- o Parses employer's employee numbers 1851-81 into fields by age and gender of workforce, etc.
- o Identifies all business partners and joint entrepreneurs identifiable in the census and codes by partner size, partner type, gender mix, family associations etc., and where possible links them together under the same firm code[6]
- o Identifies all employers running firms that employed household members
- o Corrects and recodes employment status in I-CeM for England and Wales in 1911; this will be included in I-CeM Version 2

---

[6] Bennett, R.J., 2016. Interpreting business partnerships in late Victorian Britain. *Economic History Review*, v. 69, p.1199-1227. doi:10.1111/ehr.12327

- o Adds own-account employment status missing in FMP transcripts in I-CeM for all individuals in Scotland 1891-1901; this will be included in I-CeM Version 2
- o Provides weights for misallocation and non-response bias in 1891-1911 England and Wales, and 1891-1901 in Scotland (available with WP 11)[7]
- o Provides estimates of the population of all employers and own-account individuals over 1851-81 that the census instructions were insufficient to identify (see below: variables EMPSTATUS_NUM and EMPSTATUS_IND). This uses a 'reconstruction' method to supplement the known entrepreneurs identified in extraction Groups (see below, variable: GROUP) (methods given in WP 9 and 9.2 for England and Wales; WP 20 for Scotland)
- o Identifies and codes into BBCE all company directors and enriches for census individuals that can be linked with *Directory of Directors* entries 1881-1911. [An additional data deposit is being made of the full list of all *Directory of Directors* entries (i.e. for those that can be linked to the census *and* those that cannot)]

4. Other issues important for BBCE Users

Additional data used by the ESRC '*Drivers of Entrepreneurship and Small Businesses*' project were created for company directors, transport infrastructure, land quality: see other Working Papers, which are being added to over time.

Note that the BBCE is focused on the economically active which have been pre-cleaned with inactive and those under 15 excluded. However, a few people that are identified in the census as inactive or have blanks for occupation fields are included in BBCE where enrichment has been used that shows they were actually economically active (e.g. for some company directors).

In general, BBCE includes only the entrepreneurs from the economically active, but users will find it useful to have a full list of all the economically active for comparison purposes, and for any study that seeks to focus solely on workers, or the inactive. The reconstruction estimates EMPSTATUS_IND and EMPSTATUS_NUM (see below) provide estimates of employers and own account, but at the same time have had to make estimates of who were workers. These worker status individuals are not included in BBCE as they are very numerous, but separate downloads of the RecIDs of these workers are available with WP 9 and 9.2 (England and Wales) and WP 20 (Scotland), as listed at the end of this *Guide*.

5. Other materials

This *Guide* defines all variables in BBCE. Background to how the decisions were made to prepare these variables is given in Working Papers listed at the end of the *Guide* and on the BBCE website. The BBCE data deposit has a summary list of all variables in the UKDA documentation. The Working Papers are available as downloads, and several also have downloads of supplementary data, definitions used, and working material to help other users. One of the main outputs from analysis of the database gives further information on

---

[7] see Bennett. R.J., Smith, H. and Montebruno, P. (2018) The Population of Non-Corporate Business Proprietors in England and Wales 1891-1911, *Business History*, 2018. doi.org/10.1080/00076791.2018.1534959

interpretation, related literature, and potential avenues for future research: Bennett et al. (2019) *The Age of Entrepreneurship: Business proprietors, self-employment and corporations since 1851.*

Important aspects of the BBCE for researchers to bear in mind are:

- o Almost all data in BBCE can be linked to I-CeM via the RecID, which is common between the two data deposits.

- o However, all BBCE data for England and Wales in 1871 (see WP 12), and some individuals in England and Wales in 1851 (see WP 21), are additional to those in I-CeM as they were omitted in the FMP source transcripts. In addition to the BBCE codes, the England and Wales additional people in 1851 have been coded to give their I-CeM variables: SERVTS (Total number of servants), HHS (Size of household), and RELA (Relationship to household head). These codes and people are not in BBCE, but are given as downloads with WP 21.

- o Some other data, other than the BBCE variables, have been adjusted in BBCE because of errors found in I-CeM coding for sex, age and marital status. These codes are not in BBCE, but are given as downloads with WP 21. Using these should allow near-exact replication of the research reported by the ESRC *Drivers of Entrepreneurship and Small Business* project; these will be included in I-CeM Version 2.

- o Adjustment Weights for employment status should be applied by users to the 1891-1911 EMPSTATUS variable to compensate for non-response and misallocation biases 1891-1911. Weights to adjust the raw census statements on employment status are given by RecID (see download with WP 11). WP 11 also describes in detail how they were constructed; see also Bennett et al. (2019).[8]

- o There is an important amount of data lost from National Archives sources for the 1861 census. Adjustment weights need to be applied to this year to compensate. The ESRC *Drivers of Entrepreneurship and Small Business* project used the weights method developed by Jaadla et al. (2019) for women, but extended it to men and the whole population as detailed in WP 24, together with a download of the weights. This approach is recommended to other users.

- o Various spatial and other coding is available as downloads with the WPs listed at the end of this *Guide*. Researchers should find particularly useful spatial classification codes for:
  - ▪ Entrepreneur location type code for employers: see WP 8 and WP 10 for definitions and downloads by RSD.
  - ▪ Entrepreneur location type code for all entrepreneurs: see WP 8 and WP 10 for definitions and downloads by RSD.
  - ▪ Economically active location type code for all economically active individuals: see WP 8 for definitions and downloads by RSD

- o Workers by RecID. The BBCE database identified all entrepreneurs. The rest of the working population is also invaluable to know. These are identified by RecID for each year in downloads with WEP 9, 9.2 and 20. This also gives corrections to I-CeM Occodes for workers.

---

[8] Bennett. R.J., Smith, H. and Montebruno, P., 2019. The Population of Non-Corporate Business Proprietors in England and Wales 1891-1911, *Business History*, http://doi.org/10.1080/00076791.2018.1534959

o   Farmers' EMPSTATUS is identified in BBCE without data supplementation and reconstruction. A specific method to separate between farm employers, own account proprietors and farm workers is used based on their extraction GROUP.[9] The definition of the different categories for some farmers (GROUP 5) is derived from cut-offs for each RSD (parish in Scotland) based on a logit model. The cut-offs are provided as downloads with WP 9, and WP 20.

6. Replication of analysis: differences of BBCE, I-CeM v.1, I-CeM v.2, and Campop version of I-CeM

A useful guide to how BBCE works in practice is provided in the overview analysis given in Bennett et al. (2019a) *The Age of Entrepreneurship*, as well as other publications. BBCE users will be able to compare any specific analysis they undertake against the aggregate results given in these publications and table sin the WPs. This should help users to trap any major errors in using BBCE, and also allow other researchers to develop points of departure, test alternative methods and inferences, and develop the scholarship in this field further.

For most purposes near-exact replication will be achieved using BBCE data linked to I-CeM. However, users should be aware that exact replication is not possible because of the different versions of I-CeM that have been used. At the time of the BBCE creation only I-CeM v.1 was available at UKDA. As noted in Section 2, I-CeM v.1 was supplemented and enriched specifically for entrepreneurs in BBCE. But it was also improved by numerous corrections undertaken by the authors, collaborators at Campop, and support from Kevin Schürer. This means that the BBCE uses a Campop version of I-CeM.[10] Most of these improvements in the Campop version of I-CeM will be included in I-CeM v.2, but that version will also include additional improvements. This means that users will not be able to precisely replicate analyses in *The Age of Entrepreneurship* or other publications.

The key differences between I-CeM versions are summarised in WP 1. Some of the issues involved are summarised in the *I-CeM Guide v.2*.[11] I-CeM shortcomings derive from the FMP data or enrichment coding errors. There are four main categories:

(i) For some years there are duplicated records where FMP entered the same people two or more times in different versions of keying and these were not cleaned in I-CeM v.1. This especially affects 1861 where I-CeM v.1 contains about 250,000 duplicated records, with smaller numbers of duplicates in all years. These records have been deleted in the Campop I-CeM version, and the BBCE version for Scotland. In BBCE and related downloads the entrepreneurs and the workers that are identified by RecID are all corrected with no duplicates. However, users will have in the rest of their I-CeM v.1 data the original duplicates for some people (who will look different as having different

---

[9] See Montebruno, P., Bennett, R.J., van Lieshout, C., Smith, H. and Satchell, M., 2019. Shifts in agrarian entrepreneurship in mid-Victorian England and Wales. *Agricultural History Review*, 67, 71-108.
[10] Schürer, K., Higgs, E., Reid, A.M., Garrett, E.M. (2016) Integrated Census Microdata V.2 (I-CeM.2). Note: at the time of BBCE development this did not cover Scotland, which was cleaned separately by the BBCE project team.
[11] https://www.essex.ac.uk/history/research/icem/documents/icem-guide-version-2-2015.pdf

RecID, but are actually the same – entered twice). Users should therefore beware that they may be counting possible duplicates in spatial units or other analysis and eliminate these.

(ii) I-CeM v.1 has uncorrected occupational codes; users should use BBCE as the most accurate coding available until release of I-CeM v.2; but even then, users requiring economically-defined occupation should still use BBCE in preference (for entrepreneurs and workers). The is because, as noted earlier in this *Guide*, I-CeM coding gives farming precedence and not identifying main occupation, uses many ranks as occupations, does not code to portfolios, and is affected by split lines, etc.).

(iii) For 1891, 1901 and 1911, where the census has specific identifiers for EMPSTATUS as employer, own account and worker, there can be a significant number of individuals miscoded between categories for reasons deriving from both the original census enumeration process and the encoding method used in I-CeM for this information. Re-coding these requires various assumptions to be made and weighting to be applied. Most of the corrections made in BBCE will not be carried into new versions of I-CeM. Users should therefore continue to use BBCE as the most accurate source for employer and own account status.

(iv) Some of the additional enrichment coding added to the census by I-CeM had a significant degree or error and misplacement of records in I-CeM v.1. This affected:

- o Household classification and identification of heads
- o Conjugal family units
- o Place of birth[12]
- o Geographical location attributed to the household.

The corrections in the Campop I-CeM version which are used in *The Age of Entrepreneurship* and other publications will be included I-CeM v.2 and hence will correct most of these defects. However, users need to be aware that some errors persist and care needs to be exercised in all interpretations especially of the spatial units which can be assumed to be accurate only at the RSD level, and only to a percentage confidence level, leaving some significant mis-allocations for a few parishes within some RSDs.[13] However, even some of the RSDs contain mis-allocated individuals who should be attributed to other locations. This occurs because of detached and otherwise unattributable original census sheets in FMP. Some of these defects are irredeemable; however, future versions of I-CeM may improve some spatial coding beyond I-CeM v.2.

Section 2 summarises other detailed differences of coding in BBCE and I-CeM. Section 5 and WP 21 give details of downloads which allow additional people in BBCE for 1851 that are not in I-CeM in England and Wales, to be identified, as well as giving small corrections for errors in I-CeM that are adjusted in BBCE for sex, age and marital status. Using these downloads and exercising care with the four issues outlined above should allow near-exact replication of the research in *The Age of Entrepreneurship* and other publications.

---

[12] More detail in Smith, H., Bennett, R.J. and van Lieshout, C., 2019. Immigrant Business Proprietors in England and Wales (1851-1911), *Continuity and Change*, 34/2, 253-76.http://doi: 10.1017/S0268416019000171
[13] https://www.essex.ac.uk/history/research/icem/documents/icem-guide-version-2-2015.pdf

7. Variable descriptions

In the list below each variable in BBCE available through UKDA is defined: giving its basic data *Type*, *Maximum Length*, *Description* of its contents, and the *Coverage* available (by year, and country of Great Britain). Background details of how the data were derived is included where not clear directly from I-CeM or census originals; additional detail is referred to through Working Papers with additional data downloads available from the University of Cambridge depository (given with each WP). All these Working Papers and downloads are listed at the end of this *Guide* with DOI links. Some variables, such as name and address, cannot be downloaded but access to them can be obtained by application to the UKDA for a Special Licence: this process is for an I-CeM Special Licence for access to FMP records. They can then be linked with BBCE through the RecID.

To facilitate comparisons, the format of variable descriptions and terminology is kept aligned with I-CeM as far as possible.

Variable: YEAR

Variable label: Census enumeration year

Type: Numeric

Max length: 4

Description: Census year.

Coverage: E&W 1851-1911; Scotland 1881-1901, as in I-CeM

Variable: COUNTRY

Variable label: Code for country of enumeration

Type: Character

Max length: 3

Description: ENG England; WAL Wales; SCT Scotland; BBCE excludes 'Islands in the British Seas' (i.e. Isle of Man, Channel Islands)

Coverage: 1851-1911 England and Wales, 1851-1901 Scotland, as in I-CeM

Variable: RecID

Variable label: Unique person identifier

      Note: same as in I-CeM except for additional people identified in 1851, and all 1871

Type: Numeric

Max length: 8

Description: Key variable to link with I-CeM. Unique identifier for each person's record, within each country and year. As in I-CeM, with additional people added who are not in I-CeM given high numbers not used by I-CeM. The 1871 RecID is unique to BBCE and does not relate to any future I-CeM deposits for 1871.

Coverage: 1851-1911 England and Wales, 1851-1901 Scotland; 1871 given RecID for BBCE but no equivalent 1871 data available in I-CeM

Variable: 71S&NID

Variable label: Unique person identifier in S&N data for 1871

Type: Numeric

Max length: 8

Description: Additional unique identifier for each person's record, in 1871 England and Wales only. This is the S&N identifier 'uid', an internal S&N data ID not available on their public website, but an essential link to their database. S&N attempt to keep the 'uid' stable, but any user in future who wishes to identify the same people via S&N internal searches should check that that the 'uid' has not changed since the date when accessed for the BBCE (unchanged up to the end of 2018).

Coverage: 1871 England and Wales only

Variable: 71CENSUSREF

Variable label: Archival census reference for 1871

Type: Character

Max length: 50

Description: This is the reference to The National Archive census page of the record. It consists of the Series Code, which is "RG10" for 1871, followed by a 'Piece' number (usually covering a whole or part of a registration sub-district), and then a page reference indicating entries within the piece. This is derived from S&N transcripts. For other years users should consult the I-CeM Guide (p. 134) for CENSUSREF.

Coverage: 1871 England and Wales for S&N records only

Variable: SOURCE

Variable label: Data source

Type: Numeric

Max length: 1

Description: Source from which record of individual is obtained:

1	I-CeM version 1
2	Inspection of CEBs by project team
3	S&N: The Genealogist
4	Hybrid: S&N integrated with I-CeM

Coverage: 1851-1911 England and Wales, 1851-1901 Scotland

Variable: PARID71

Variable label: census parish number

Type: numeric

Max length: 6

Description: parish location in which person is enumerated as in S&N; parish dictionary as in I-CeM http://doi.org/10.5255/UKDA-SN-7481-1

Coverage: 1871 England and Wales only

Variable: PARISH71

Variable label: Parish name

Type: Character

Max length: 90

Description: Parish name of enumeration, as used in S&N (note some may differ from exact I-CeM names due to transcriptions and consolidation of names, but have been assigned to the correct parish in the I-CeM dictionary under PARID71; for method used see WP 12)

Coverage: 1871 England and Wales only, from S&N

Variable: URBAN_CLASS

Variable label: Urban-rural classification code

Note: see WP 6 for definitions, with associated download of full classification at parish level, and Smith et al. (2018).[14]

Type: Numeric

Max length: 1

---

[14] Smith, H., Bennett, R.J. and Radicic, D. (2018) Towns in Victorian England and Wales: a new classification. *Urban History*, 45, 4, pp. 568-594. doi.org/10.1017/S0963926818000020

Description: Type of Location in which person is located. Urban refers to towns with populations greater than 10,000 in that census year. Transition type 1 are non-urban parishes in a Registration Sub District containing type 1 parishes.[15] Transition type 2 are parishes in Registration Sub Districts with population densities greater than 0.3 people per acre. Rural are parishes in Registration Sub Districts with population densities less than 0.3 people per acre.

1	Urban
2	Transition type 1
3	Transition type 2
4	Rural

Coverage: 1851-1911 England and Wales

Variable: BURGH

Variable label: Urban-rural classification code

Note: see WP 6 for definitions; with associated download of full classification at parish level

Type: Numeric

Max length: 1

Description: Name of Parliamentary Burgh in which a person is located. Also indicates all individuals in a Burgh; those without a value lived outside the Burghs.

Coverage: 1851-1901 Scotland

Variable: TOWN

Variable label: Town name

Note: see WP 6 for definitions of towns; with associated download of full list and definitions at parish level

Type: character

Max length: 35

Description: Name of town in which person is located based on their parish of residence, as in URBAN; towns are those over 10,000 in population at the census date, as defined by Law and Robson, aligned as far as possible with I-CeM data (see WP 6).

Coverage: 1851-1911 England and Wales

---

[15] Law, Chris M. (1967), 'The growth of urban population in England and Wales, 1801-1911', *Transactions of the Institute of British Geographers*, 41, 125-43; Robson, Brian (1973), *Urban Growth: An Approach*, London: Methuen.

Variable: EMPSTATUS

Variable label: Status as entrepreneur code from census responses

Type: Numeric

Max length: 1

Description: Entrepreneur status among economically active, as reported in 1891-1911 censuses, and coded by I-CeM as EMPLOYCODE, with corrections for misallocation bias (WPs 4 and 11), infilled for miscoded EMPLOYCODE in 1911 I-CeM. For Scotland the missing own account data in 1901 not transcribed in FMP/I-CeM for Scotland is infilled (see WP 20).

1	Worker
2	Employer
3	Own account
4	Inactive or blank: these are individuals who are not identified as entrepreneurs; this also includes some partners and company directors who were not recorded as employers, own account or workers

Coverage: 1891-1911 England and Wales; 1891-1901 Scotland

Variable: EMPSTATUS_IND

Variable label: Status as entrepreneur code: method IND (Individual focus)

Type: Numeric

Max length: 1

Description: Entrepreneur status among economically active as supplemented for 1851-81. See WP 9.2 (WP 20 Scotland) for definition, with associated downloads of decision files, and workers in the rest of the economically active

For 1851-1881 censuses (both England and Wales, and Scotland) status uses original census responses GROUP supplemented with _estimates_ using method IND; see WP 9.2. Method IND assigns all individuals using a logit model with tailored cut-offs. It gives a smaller total than method NUM (see below) but does not use random allocation so each individual is given a status based on a score calculated using only available information in the census. It is preferable for identifying the most probable INDIVIDUAL entrepreneur status in the supplementation for 1851-81, but the total number identified probably underestimates the actual situation. All definitions of reconstructions with alternative estimates are provided in separate downloads listed with WP 9.2 (WP 20 Scotland), and at the end of this _Guide_. Company DIRECTORS and PARTNERS are coded as separate supplementary codes (later below). Both have an entrepreneur status code here based on their occupation as given in the census. All individuals identified as entrepreneurs within the GROUP variable (see below) are included with appropriate status as employers or own account.

1	Worker
2	Employer
3	Own account
4	Inactive or blank: these are individuals who are not identified as entrepreneurs in some EMPSTATUS estimates; this also includes some partners and company directors who were not recorded as employers, own account or workers

Coverage: 1851, 1861, 1881 England and Wales; 1851-1881 Scotland

Variable: EMPSTATUS_NUM

Variable label: Status as entrepreneur code: method NUM (Number focus)

Type: Numeric

Max length: 1

Description: Description: Entrepreneur status among economically active as supplemented for 1851-81. See WP 9 for definition, with associated downloads of decision files, and workers in the rest of the economically active.

For 1851-1881 censuses status: the original responses GROUP supplemented with *estimates*; see WP 9. Method NUM (only available for England and Wales; *method not applied to Scotland*) assigns all individuals using a mix of extrapolation, a logit model, and intelligence-led approaches based on secondary literature, which result in some individuals are identified by random assignment. It is therefore less satisfactory for identifying individuals than method IND (above), although the majority of individuals are the same under both methods, and both have all the extracted from GROUP 1-8. It is preferable for identifying the most probable NUMBER of entrepreneurs in the supplementation for 1851-81, but the individuals identified will sometimes be random assignments. It gives a larger number than method IND. All definitions of reconstructions with alternative estimates are provided in separate downloads listed with WP 9, and at the end of this *Guide*. Company DIRECTORS and PARTNERS are coded as separate supplementary codes (later below). Both have an entrepreneur status code here based on their occupation as given in the census. All individuals identified as entrepreneurs within the GROUP variable (see below) are included with appropriate status as employers or own account.

1	Worker
2	Employer
3	Own account
4	Inactive or blank: these are individuals who are not identified as entrepreneurs in some EMPSTATUS estimates; this also includes some partners and company directors who are not recorded as employers, own account or workers

Coverage: 1851, 1861, 1881 England and Wales

Variable: RELA10

Variable label: Relationship to household head, aggregated code

Type: Numeric

Max length: 2

Description: Relationship to head of family, derived from I-CeM variable RELA, aggregated to 10 categories. It is a simplification of the I-CeM RELA code for all years except 1871 (which was coded to RELA10 by BBCE directly from S&N transcriptions).

1	Head
2	CFU member (spouse, children)
3	Older generation (parent, uncle/aunt)
4	Siblings
5	Other family (grandchildren, cousin, niece/nephew, extended and other relatives)
6	Servants
7	Working title
8	Lodgers/boarders
9	Non-household (institution and military)
10	Unknown relationship/visitors

Coverage: 1851-1911 England and Wales; 1851-1901 Scotland

Variable: GROUP

Variable label: Extraction source of entrepreneur data code

Type: Numeric

Max length: 1

Description: How entrepreneurs were identified and extracted from occupation strings following census instructions [see WP 2 for census instructions, WP 13 (for comparisons with census publications), WP 3 (for method in 1851-61, 1881), WP 12 (for method in 1871), WP 20 for how groups were slightly adapted in Scotland (mainly for crofters), WP 9 and WP 9.2 (for comparisons with how used in reconstruction)], and journal publications listed at www.bbce.uk

1	All stating employer and any others with stated employees
2	All stating 'employer' but with no stated employees
3	Master etc. but with no employees
4	'Farmer' not stating employers or acres, treated as workers unless reporting portfolio occupations
5	Farmer giving acres but no stated employees
6	Owners or proprietors of business assets

7	Code NOT used in data deposit (covers non-entrepreneurs owners extracted in early stages of analysis)
8	Directors (all types of company directors); this does not include all those in the census available under DIRSOURCE (see below)

Coverage: 1851-81 all countries

Variable: ETOT

Variable label: Employees total

Type: Numeric

Max length: 6

Description: total number of employees employed of all types reported by employers; note that an earlier version of these data without RecID has been deposited as part of a firm size analysis.[16]

Coverage: 1851-81 all countries

Variable: EMEN

Variable label: Men employees

Type: Numeric

Max length: 6

Description: number of MEN employees employed reported by employers. Includes descriptors that were obviously men, such as tradesmen.

Coverage: 1851-81 all countries

Variable: EWOM

Variable label: Women employees

Type: Numeric

Max length: 6

---

[16] Piero Montebruno, Robert J. Bennett, Carry van Lieshout, and Harry Smith, (2019) A tale of two tails: Do Power Law and Lognormal models fit firm-size distributions in the mid-Victorian era?, *Physica A: Statistical Mechanics and its Applications*, https://doi.org/10.1016/j.physa.2019.02.054.
Data download supplementary data at: Mendeley Data, 2019. http://dx.doi.org/10.17632/86xkkncmw3.1

Description: number of WOMEN employees reported by employers. Includes descriptors easily identifiable as women such as laundress, shopwomen.

Coverage: 1851-81 all countries

Variable: EBOY

Variable label: Boy employees

Type: Numeric

Max length: 6

Description: number of BOY employees reported by employers.

Coverage: 1851-81 all countries

Variable: EGIRL

Variable label: Girl employees

Type: Numeric

Max length: 6

Description: number of GIRL employees reported by employers.

Coverage: 1851-81 all countries

Variable: ELAB

Variable label: Labourer employees

Type: Numeric

Max length: 6

Description: number of LABOURER employees reported by employers.

Coverage: 1851-1871 all countries; Scotland 1881

Variable: EAPP

Variable label: Apprentice employees

Type: Numeric

Max length: 6

Description: number of APPRRENTICE employees reported by employers.

Coverage: 1851-1871 all countries; Scotland 1881

Variable: EJOURNEY

Variable label: Journeyman employees

Type: Numeric

Max length: 6

Description: number of JOURNEYMAN employees reported by employers.

Coverage: 1851-1871 all countries; Scotland 1881

Variable: EMALE

Variable label: Male employees

Type: Numeric

Max length: 6

Description: number of MALE employees reported by employers. Includes descriptors where gender is stated but no indication of age, such as 'son', and all male contractions such as 'men and boys'.

Coverage: 1851-1871 all countries; Scotland 1881

Variable: EFEMALE

Variable label: Female employees

Type: Numeric

Max length: 6

Description: number of FEMALE employees reported by employers.[17] Includes descriptors where gender is stated but no indication of age, such as 'daughter', and all female contractions such as 'women and girls'

Coverage: 1851-1871 all countries; Scotland 1881

Variable: ECHILD

Variable label: Child employees

---

[17] More detail in van Lieshout, C., Smith, H., Montebruno, P. and Bennett, R.J., 2019. Female Entrepreneurship: Business, Marriage and Motherhood in England and Wales, 1851-1911, *Social History,* 44.4 https://doi.org/10.1080/03071022.2019.1656929

Type: Numeric

Max length: 6

Description: number of CHILD employees reported by employers. Includes young persons and contractions such as 'boys and girls'.

Coverage: 1851-71 all countries; Scotland 1881

Variable: EOTHER

Variable label: Other employees

Type: Numeric

Max length: 6

Description: number of all EOTHER types of employees reported by employers. Includes general non-gendered titles and occupations such as servants, clerks, assistants, bakers, hands; and contractions that include both males and females and/or unknowns, e.g. 'men, women, and boys' or 'males and apprentices'. EOTHER for 1881 in England and Wales includes the categories labourers, apprentices, journeymen, male, female and child employees. These were not extracted separately for this year and so are combined with any 'others' in this variable.

Coverage: 1851-81 all countries

Variable: ACRE

Variable label: Acreage of farm

Type: Numeric

Max length: 6

Description: number of ACRES reported by farmers.

Coverage: 1851-81 all countries

Variable: ENTOCCODE

Variable label: occupational code

Type: Numeric

Max length: 5

Description: Numeric code for occupational groupings in the I-CeM Occupational Matrix, with corrections to I-CeM individual's codes; re-coding to entrepreneur relevance; and, where multiple occupations, giving main occupation first. For codes see E. Higgs, C. Jones, K. Schürer and A. Wilkinson, *Integrated Census Microdata (I-CeM) Guide*, 2nd edition, Colchester, 2013), 163-83.

Coverage: 1851-1911 England and Wales; 1851-1901 Scotland, recoded from I-CeM

Note the BBCE data for ENTOCCODE is for entrepreneurs only; **workers** are identified in separate downloads, and their corrected OCCODES are given there. The workers identified depend on the employment status method used. Workers identified by method EMPSTATUS_IND are available for England and Wales in WP 9.2:

https://doi.org/10.17863/CAM.44406

Workers by method EMPSTATUS_NUM are available for England and Wales in WP 9:

https://doi.org/10.17863/CAM.44404

For Scotland for EMPSTATUS_IND in WP 20:

https://doi.org/10.17863/CAM.44409

Variable: EA17

Variable label: 17 sector codes

Note: see definitions and coding from Occode in WP 5

Type: numeric

Max length: 2

Description: Sector codes for 17 sector groupings; first 13 contain entrepreneurs and non-entrepreneurs; last 4 contain only non-entrepreneurs.

1	Farming and fishing (incl. farming labourers & family workers)
2	Mining & quarrying
3	Construction
4	Manufacturing (incl. gas, electricity & water manufacture)
5	Makers and dealers
6	Retail
7	Transport
8	Professional & business services
9	Professional & personal services
10	Agricultural produce, drink & tobacco manufacture
11	Food retailing
12	Lodging & refreshment
13	Finance & commerce
14	Public administration, military, clergy
15	Domestic service
16	Undefined general labourers
17	Persons of property with no stated occupation

Coverage: 1851-1911 England and Wales; 1851-1901 Scotland

Variable: EA51

Variable label: 51 sector codes

> Note: see definitions and coding from Occode in WP 5

Type: Numeric

Max length: 2

Description: Sector codes for 51 sector groupings; first 44 contain entrepreneurs and non-entrepreneurs; last 7 contain only non-entrepreneurs.

1	Farming, fishing, market gardeners, horse breeding & keeping
2	Coal mining
3	Other mining & quarrying, brickmaking, gravel, salt works
4	Construction operatives (masons, bricklayers, thatchers, plumbers etc.)
5	Machinery manufacture
6	Tool & weapons manufacture
7	Iron & steel manufacture, bolts & nails
8	Blacksmiths
9	Other metal manufacture (copper, tin, brass, whitesmiths, etc.)
10	Ship, road & rail vehicle manufacture
11	Earthenware & glass manufacture
12	Gas, coke, water & chemical manufacture
13	Leather, fur, hair & bone manufacture
14	Wood manufacture (sawyers, coopers, cane workers)
15	Furnishing manufacture (cabinet makers, french polishers, undertakers)
16	Printing & paper manufacture (paper, cardboard, printers, bookbinders)
17	Waterproof goods manufacture (floor & oil cloth, rubber etc.)
18	Woollen manufacture (woollen goods, carpets, blanket, flannel)
19	Cotton & silk manufacture (incl. ribbon, weaving, dyeing, bleaching etc.)
20	Other textile manufacture (flax, hemp, rope, jute, lace, tape, thread)
21	Clothing manufacture (tailors, milliners, hosiery, hats, gloves, umbrellas, buttons, leather)
22	Shoe, boot, clog manufacture
23	Agric. produce manufacture (millers, refiners, bakers, confectioners)
24	Drink & tobacco manufacture (maltsters, brewers, distillers, tobacco & pipes)
25	Watch & instrument manufacture
26	General manufacture (manufacturers, mechanics, artisans, machinists)
27	Ocean, inland and dock transport
28	Road & rail transport
29	Coal dealing
30	Timber, hay, corn & agricultural produce dealing
31	Clothing and dress dealing (drapers, hosiers, haberdashers)
32	Food sales (butchers, fishmongers, cheesemongers, millsellers, grocers)
33	Lodging & drink sales (wine & spirits, hotels, inns, coffee houses)
34	Communications (publishing, newsagents, stationers & telecoms)
35	Household & personal goods dealer (earthenware, glass, jewellers)
36	Ironmongers
37	Other retail (general shopkeeper, huckster, hawker)

38	Chemists, druggists
39	Merchants, banks, insurers & brokers
40	Other commerce (accountants, salesmen, travellers, officers of companies)
41	Construction management (builders & contractors)
42	Professions (barristers, solicitors, scientific pursuits)
43	Professions (doctors, dentists, artists, performers, education)
44	Personal services (washing & bathing, hairdressing, chimney sweeps)
45	Public administration, clergy
46	Military
47	Domestic service
48	Undefined general & factory labourers
49	Factory hand (textile, undefined)
50	Commercial clerks
51	Persons of property

Coverage: 1851-1911 England and Wales; 1851-1901 Scotland

Variable: OCCSTRING

Variable label: Occupation descriptor

Type: alphanumeric

Max length: 350   **NOTE need to allow this space in your software, which often exceeds defaults in Access or Excel**

Description: occupational descriptor strings from CEBs and S&N that are supplemental to I-CeM, as defined by SOURCE; see WP 1, WP 3, WP 12, WP 20 Scotland.

Coverage: England and Wales: 1851-1861 only for FMP/I-CeM gaps and truncations; 1871 from S&N for all data entries. Scotland: 1851-1901 only for FMP/I-CeM gaps and truncations

Variable: PORT

Variable label: Portfolio business

> Note: Individuals defined as portfolio if they run businesses in different sectors of Occode, for occupations defined in Woollard (2001)[18] for 1881, applied to all censuses

Type: Character

Max length: 1

Description: Y entrepreneur has more than one business activity; N only one sector of activity

---

[18] Matthew Woollard (2001) *The classification of occupations in the 1881 census of England and Wales*, Historical Censuses and Social Surveys Research Group Occasional Paper No. 1, Department of History University of Essex, Colchester. Bennett, R.J., Smith, H., van Lieshout, C., Montebruno, P., Newton, G., 2019. *The Age of Entrepreneurship: Business proprietors, self-employment and corporations since 1851, Chapter 11*. Routledge International Studies in Business History, London and New York. http://doi:10.4324/9781315160375

Coverage: 1851-1911 England and Wales; 1851-1901 Scotland

Variable: PORT1

Variable label: Portfolio business sector 1

Type: Numeric

Max length: 6

Description: ENTOCCODE of sector of **first** business activity, ranked in order of size (if available) or order of listing. Note that occodes for 1881 are at the level of 1881EW, for which see the I-CeM occupational matrix.

Coverage: 1851-1911 England and Wales; 1851-1901 Scotland

Variable: PORT2

Variable label: Portfolio business sector 2

Type: Numeric

Max length: 6

Description: ENTOCCODE of sector of **second** business activity, ranked in order of size (if available) or order of listing

Coverage: 1851-1911 England and Wales; 1851-1901 Scotland

Similarly for PORT3– PORT7

Variable: PORT8

Variable label: Portfolio business sector 8

Type: Numeric

Max length: 6

Description: ENTOCCODE of sector of **eighth** business activity, ranked in order of size or of listing

Coverage: 1851-1911 England and Wales; 1851-1901 Scotland

Variable: DIRECT

Variable label: Director of company

Type: Character

Max length: 1

Description: Y Individuals identified as directors, N if not

Coverage: 1851-1911 England and Wales; 1881-1901 Scotland. Small numbers before 1881

Variable: DIRSOURCE

Variable label: Director of company source code

Type: Character

Max length: 1

Description: Individuals identified from source of data (census or *Directory of Directors* (DoD)); see WP 14.

1	Called 'director' in census but with no confident DoD match found [all years 1851-1911]
2	Identified from DoD and matched in census, but not called 'director' in census [only for 1881 and later]
3	Called 'director' in census and matched with DoD [only for 1881 and later]

Coverage: 1851-1911 England and Wales; 1881-1901 Scotland; DoD-sourced 1881-1911 England and Wales; 1881-1901 Scotland

Variable: DIRNO

Variable label: Number of companies

Type: Numeric

Max length: 3

Description: Number of companies for which individual linked in census to DoD is identified as a director in DoD at census dates

Coverage: 1851-1911 England and Wales; 1881-1901 Scotland

Variable: DIR1

Variable label: director business sector 1

Type: Numeric

Max length: 2

Description: Main sector code of company's activity, or most frequent business activity if director has more than one, for individuals linked in census to DoD. Codes are identical to EA17 sectors, except for 18-22 which are disaggregations from EA17 as indicated, and 9 which combines 8 and 9 from the EA17; see WP 14.

1	Farming/estate work/plantations
2	Mines and quarries (non-coal)
3	Construction
4	Manufacturing
5	Maker dealer
6	Retail & wholesale
7	Transport
9	Personal & business services
10	Agricultural produce & dealing
12	Refreshment
13	Finance & commerce (including land investment)
18	Utilities & Municipal (from EA17 code 4)
19	Coal mining (from EA17 code 2)
20	Steel & chemical manufacture (from EA17 code 4)
21	Communications (from EA17 code 6)
22	Railway (from EA17 code 6)
0	Unknown

Coverage: 1851-1911 England and Wales; 1881-1901 Scotland

Variable: DIR2

Variable label: director business sector 2

Type: Numeric

Max length: 2

Description: Sector code of company's second most frequent business activity (if more than one), for individuals linked in census to DoD

Coverage: 1851-1911 England and Wales; 1881-1901 Scotland

Variable: DIRGEOG1

Variable label: director location 1

Type: Numeric

Max length: 2

Description: Main location of company's activity or most frequent (if more than one), for individuals linked in census to DoD; see WP 14

1	England and Wales
2	Scotland
3	Ireland
4	Not UK
5	Channel Islands and Isle of Man
6	Colonial (British)
7	European
8	USA
9	Foreign (other)
10	International (multiple territories; can include UK)
11	Unknown

Coverage: 1851-1911 England and Wales; 1881-1901 Scotland

Variable: DIRGEOG2

Variable label: director location 2

Type: Numeric

Max length: 5

Description: Location of company's activity that is second most frequent (if more than one), for individuals linked in census to DoD; see WP 14

Coverage: 1851-1911 England and Wales; 1881-1901 Scotland

Variable: PART

Variable label: Partner in a business

Type: character

Max length: 2

Description: Y Individuals identified as partner joint or other co-operator (as defined below), N if not; see WP 18.[19]

Coverage: 1851-1911 England and Wales; 1851-1901 Scotland

Variable: PARTID

Variable label: Partnership identifier for each business

---

[19] Bennett, R.J., 2016. Interpreting business partnerships in late Victorian Britain. *Economic History Review*, v. 69, p.1199-1227. doi:10.1111/ehr.12327

Type: Numeric

Max length: 5

Description: Code given to all partners in same business firm (where they can be linked) for given year (unique PARTID only to each year); see WP 18

Coverage: 1851-1911 England and Wales; 1851-1901 Scotland

Variable: PARTREF

Variable label: Code for partnership reference person

Type: Numeric

Max length: 1

Description: Code given to partners in same business by level of seniority. Reference person has all partnership information (when this is available) so that they can be used alone to analyse partnerships by excluding non-reference persons and hence removing duplication of business numbers. But where not all partners are known the only person coded is type 2; hence if a partner to represent each firm is required, then users should first filter on PARTID selecting all PARTREF 1, and then filter on the *rest* of PARTID to select PARTREF 2. Note: see WP 18 for definitions

1	Reference person: Senior (if stated as such), or that giving employee numbers or acres as required in census instructions, or that deduced as senior from household order or statements
2	Other partner (or not explicitly senior) or where rest of partners not known
blank	If unknown or has no other partners given so not deducible unless explicitly stated; usually these are coded as a non-senior partner (2). They can be used as a reference person only where no other partners are given.

Coverage: 1851-1911 England and Wales; 1851-1901 Scotland

Variable: PARTTYPE

Variable label: Type of partner

Type: Character

Max length: 5

Description: Type of partner deduced from their descriptor; uses all partners/individuals involved to inform PATYPE, not just one individual's entry (e.g. one individual saying 'with' and another saying 'partner' gives partner code, but two 'with' is de facto). Note: see WP18 for definitions

P	Explicit 'partner' or 'co-partner' statement; includes 'company with', 'conjunction with', 'in firm with', and 'employing X men with another'
J	'Joint' operator with others as explicit joint statement
D	de facto partner operating with others: 'with' statement, or duplicate or near-duplicate descriptor (often 'ditto'); includes 'master/mistress assisting with'
A	'Assistant/ing with' but no mention of partner/joint status or master/mistress ('master/mistress assisting with' is taken as partner); differs from 'operating with' above; but 'assistant' with no extra descriptor is excluded. This category is ambiguous and can be excluded if desired; it is small in number, and mainly occurs for 1911
O	Land owners operating together (only coded in 1881; *should usually be excluded from analysis*)
R	'Retired' where explicitly stated (a small number)
S	'Sleeping' where explicitly stated (a small number, none found in Scotland)
Sal	Only where 'Salaried' was explicitly stated (very rare statement, none found in Scotland)
Dec	Partner of recently deceased other person identified by explicit statement (very rare, coded to allow former partner status to be identified)

Coverage: England and Wales1851-1911; 1851-1901 Scotland

Variable: PARTN

Variable label: Number of partners in partnership

Type: Numeric

Max length: 1

Description: Number of partners that can be confidently identified as the whole partnership; where not complete or uncertain number is left bank; Note: see WP 18 for definitions

Coverage: 1851-1911 England and Wales; 1851-1901 Scotland

Variable: PARTGEN

Variable label: Gender of partners in partnership

Type: Character

Max length: 1

Description: Gender of partnership where all partners can be confidently identified as the whole partnership. Note: see WP 18 for definitions

M	All partners male
F	All partners female
X	Partners of both genders (at least one of each); even if all partners are not known, if one female and one male is identified then this is mixed, but males with unknown others, and females with unknown others are coded as unknown U. Partners with an unknown other, but stating 'with father', 'with daughter', etc., are coded to gender
U	Gender mix unknown because not all, or insufficient, partners are identified, or descriptor is unclear: e.g. male 'with proprietor' as an unknown other does not have clear gender mix. This also occurs where I-CeM has a 'U' code for gender

Coverage: 1851-1911 England and Wales; 1851-1901 Scotland

Variable: PARTRELA

Variable label: Relationship between partners in partnership

Type: Numeric

Max length: 2

Description: Same code applies to all partners in same PARTID; deduced from RELA code if in same household and using names where possible. Relates solely to those who are identified as partners and not to rest of household. Based solely on census information so that detailed analysis would potentially reveal more in some cases. Note: see WP 18 for definitions

blank	Unknown relationships (usually where not all partners are identified)
1	Father-sons, and/or MALE grandchildren, great grandchildren
2	Father-daughters, and/or FEMALE grandchildren, great grandchildren
3	Mother-sons, and/or MALE grandchildren, great grandchildren
4	Mother-daughters, and/or FEMALE grandchildren, great grandchildren
5	Husband & wife together; also 2 spouses with all genders of children, grandchildren, great grandchildren (excluded from 1-4 and 12-13)
6	Brother-brothers
7	Sister-sisters
8	Extended family (other than or additional to children, grandchildren, great grandchildren)
9	Extended family with others (as 8, but with non-family members as well)
10	No obvious family connection for any partners
11	Brothers-sisters
12	Father with sons AND daughters and/or mixed gender grandchildren, great grandchildren
13	Mother with sons AND daughters and/or mixed gender grandchildren, great grandchildren
14	Parent (gender unknown) with any children (used only for 1871 to cover codes 1-4, and 12-13 where full population data not available)

Coverage: 1851-1911 England and Wales; 1851-1901 Scotland

Variable: HHDFIRMREF

Variable label: Identifier of employers running co-resident household firms.

Type: Numeric

Max length: 1

Description: 1 for employers resident in the same household as workers in the same ENTOCCODE. These employers are the household reference persons for these firms. For 1851-81 only available for extracted not reconstructed employers.

Coverage: 1851-61, 1881-1911 England and Wales

Variable: HHDTYPE

Variable label: Household firm type code

Type: Numeric

Max length: 1

Description: Code for the relationship of the co-resident workers linked to the household reference person.

1	Son only
2	Daughter only
3	Brother only
4	Sister only
5	Other family only
6	Mixed family
7	Boarder or lodger only
8	Mixed non-family
9	Family & non-family

Coverage: 1851-61, 1881-1911 England and Wales

Variable: 71AGE

Variable label: Age

Type: Numeric

Max length: 4

Description: Age transcribed by S&N from CEBs.

Coverage: 1871 England and Wales only

Variable: 71SEX

Variable label: Gender

Type: Character

Max length: 1

Description: Gender as provided by S&N based on first name.

Coverage: 1871 England and Wales only

Variable: 71RELAT

Variable label: Relationship to household head

Type: Character

Max length: 50

Description: Relationship to head of family, transcribed by S&N from CEBs. This is the source string from which the variable RELA10 is derived for 1871

Coverage: 1871 England and Wales only

Variable: 71BPPARISH

Variable label: Parish of birth

Type: Character

Max length: 100

Description: Birth parish, transcribed by S&N from CEBs.

Coverage: 1871 England and Wales only

Variable: 71CNTI

Variable label: County of birth code

Type: Character

Max length: 3

Description: Birth county, transcribed by S&N from CEBs, then coded by BBCE according to I-CeM CNTI codes. For codes see E. Higgs, C. Jones, K. Schürer and A. Wilkinson, *Integrated Census Microdata (I-CeM) Guide*, 2nd edition, Colchester, 2013), pp. 201-11.

Coverage: 1871 England and Wales only

Acknowledgements

Working Papers and Data Downloads - further sources and information on BBCE:

Working paper series: ESRC project ES/M010953: *'Drivers of Entrepreneurship and Small Business',* University of Cambridge, Department of Geography and Cambridge Group for the History of Population and Social Structure. See for updates www.bbce.uk

WP 1: Bennett, Robert J., Smith Harry J., van Lieshout, Carry, and Newton, Gill (2017) *Drivers of Entrepreneurship and Small Businesses: Project overview and database design.* https://doi.org/10.17863/CAM.9508

WP 2: Bennett, Robert J., Smith Harry J. and van Lieshout, Carry (2017) *Employers and the self-employed in the censuses 1851-1911: The census as a source for identifying entrepreneurs, business numbers and size distribution.* https://doi.org/10.17863/CAM.9640

WP 3: van Lieshout, Carry, Bennett, Robert J., Smith, Harry J. and Newton, Gill (2017) *Identifying businesses and entrepreneurs in the Censuses 1851-1881.* https://doi.org/10.17863/CAM.9639

WP 4: Smith, Harry J., Bennett, Robert J., and van Lieshout, Carry (2017) *Extracting entrepreneurs from the Censuses, 1891-1911.* https://doi.org/10.17863/CAM.9638

WP 5: Bennett, Robert J., Smith Harry J., van Lieshout, Carry, and Newton, Gill (2017) *Business sectors, occupations and aggregations of census data 1851-1911.* https://doi.org/10.17863/CAM.9874

- Data download of classification file

WP 6: Smith, Harry J. and Bennett, Robert J. (2017) *Urban-Rural Classification using Census data, 1851-1911.* https://doi.org/10.17863/CAM.15763

- Downloadable classification E&W
- Downloadable classification Scotland
- Published paper in *Urban History*

WP 7: Smith, Harry, Bennett, Robert J., and Radicic, Dragana (2017) *Classification of towns in 1891 using factor analysis.* https://doi.org/10.17863/CAM.15767

- Published paper in *Urban History*

WP 8: Bennett, Robert J., Smith, Harry, and Radicic, Dragana (2017) *Classification of occupations for economically active: Factor analysis of Registration Sub-Districts (RSDs) in 1891.* https://doi.org/10.17863/CAM.15764

- Downloadable classification by RSD

WP 9: Bennett, Robert, J., Montebruno, Piero, Smith, Harry, and van Lieshout, Carry (2018) *Reconstructing entrepreneurship and business numbers for censuses 1851-81.* https://doi.org/10.17863/CAM.37738

- Downloadable decisions for NUM method
- Downloadable file of workers from NUM method E&W

- Data download of farm cut-offs

WP 9.2: Bennett, Robert, J., Montebruno, Piero, Smith, Harry, and van Lieshout, Carry (2019) *Reconstructing business proprietor responses for censuses 1851-81: a tailored logit cut-off method.* https://doi.org/10.17863/CAM.37738

- Downloadable decisions for IND method
- Downloadable file of workers form IND method E&W

WP 10: Bennett, Robert, J., Smith, Harry and Radicic, Dragana (2018) *Classification of environments of entrepreneurship: Factor analysis of Registration Sub-Districts (RSDs) in 1891.* https://doi.org/10.17863/CAM.26386

- Downloadable classification by RSD

WP 11: Montebruno, Piero (2018) *Adjustment Weights 1891-1911: Weights to adjust entrepreneur numbers for non-response and misallocation bias in Censuses 1891-1911.* https://doi.org/10.17863/CAM.26378

- Data set of Adjustment Weights for the 1891-1911

WP 12: van Lieshout, Carry, Day, Joseph, Montebruno, Piero and Bennett Robert J. (2018) *Extraction of data on Entrepreneurs from the 1871 Census to supplement I-CeM.* https://doi.org/10.17863/CAM.27488

WP 13: van Lieshout, Carry, Bennett, Robert J. and Smith Harry (2019) *Extracted data on employers and farmers compared with published tables in the Census General Reports, 1851-1881.* https://doi.org/10.17863/CAM.37165

WP 14: van Lieshout, Carry, Bennett Robert J. and Montebruno, Piero (2019) *Company Directors: Directory and Census record linkage.* https://doi.org/10.17863/CAM.37166

WP 15: Bennett, Robert, J., Montebruno, Piero, Smith, Harry and van Lieshout, Carry (2019) *Entrepreneurial discrete choice: Modelling decisions between self-employment, employer and worker status.* https://doi.org/10.17863/CAM.37312

WP 16: Satchell, M., Bennett, Robert J., Bogart, D. and Shaw-Taylor, L. (2019) *Constructing Parish-level Data and RSD-level Data on Transport Infrastructure in England and Wales 1851-1911.* https://doi.org/10.17863/CAM.37313

WP 17: Satchell, M. and Bennett, Robert J. (2019) *Building a 1911 Historical Land Capacity GIS.* https://doi.org/10.17863/CAM.42285

WP 18: Bennett, Robert, J., Smith, Harry, van Lieshout, Carry and Montebruno, Piero (2019) *Identification of business partnerships in the British population censuses 1851-1911 for BBCE.* https://doi.org/10.17863/CAM.43890

WP 18: Bennett, Robert, J., Smith, Harry, van Lieshout, Carry and Montebruno, Piero (2019) *Identification of business partnerships in the British population censuses 1851-1911 for BBCE.* https://doi.org/10.17863/CAM.43890

WP 19: Montebruno, Piero (2019) *Datasets and guide: downloads for reconstructing British census responses 1851-1881 for the BBCE.* https://doi.org/10.17863/CAM.42285

WP 20: Smith, Harry, van Lieshout, Carry, Montebruno, Piero and Bennett, Robert, J. (2019) *Preparing Scottish census data in I-CeM for the British Business Census of Entrepreneurs (BBCE).* https://doi.org/10.17863/CAM.44963

- Downloadable decisions for IND method Scotland
- Downloadable file of workers for IND method Scotland
- Data download of farm cut-offs

WP 21: van Lieshout, Carry, Bennett, Robert, J., and Smith, Harry (2019) *Additional codes and people in the British Business Census of Entrepreneurs (BBCE) not available through I-CeM.* https://doi.org/10.17863/CAM.45322

- Data download of additional people in BBCE with their I-CeM codes
- Data download of corrected codes for I-CeM for age, sex, and marital status

WP 22: Bennett, Robert, J. (2020) *Employers and self-employed in the census 1921-2011 and alignment with BBCE: Entrepreneurs, business numbers and size distribution.*