

Company Directors: Directory and Census Record Linkage, 1881-1911

Carry van Lieshout, Robert J. Bennett, and Piero Monteburuno

cv313@cam.ac.uk rjb7@cam.ac.uk pfm27@cam.ac.uk

Working Paper 14:
Working paper series from ESRC project ES/M010953:
Drivers of Entrepreneurship and Small Businesses

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

February 2019

Comments are welcomed on this paper: contact the authors as above.

© Carry van Lieshout, Robert J. Bennett, and Piero Monteburuno, University of Cambridge, members of the Cambridge Group for the History of Population and Social Structure assert their legal and moral rights to be identified as the authors of this paper; it may be referenced provided full acknowledgement is made: *Cite* (Harvard format):

Van Lieshout, Carry, Bennett, Robert J., and Monteburuno, Piero (2019) *Company Directors: Directory and Census Record Linkage, 1881-1911*. Working Paper 14: ESRC project ES/M010953: 'Drivers of Entrepreneurship and Small Businesses', University of Cambridge, Department of Geography and Cambridge Group for the History of Population and Social Structure.

Keywords: Entrepreneurship, Directors, Census

JEL Codes: G34, L26, L25, D13, D22

Company Directors: Directory and Census Record Linkage, 1881-1911

Carry van Lieshout, Robert J. Bennett, and Piero Monteburro

Working Paper 14: ESRC project ES/M010953: Drivers of Entrepreneurship and Small Businesses, University of Cambridge.

1. Introduction

Company directors are a major type of business entrepreneur that became more numerous in the late nineteenth century. Together with sole proprietors, employers, and partnerships they were the individuals who led developments of nineteenth century industry and commerce. Any study of business for this period must include them. They are discussed here as one element of the ESRC project to construct an aligned, complete, quality-controlled, and consistent database that includes as far as possible *all* business proprietors for 1851-1911 for data deposit at UKDA. That project mainly uses data from the census to develop the database to include non-corporate business proprietors. This paper discusses how external information on company directors can be used to enrich these census records, and vice versa, to allow directors and their businesses to be included in a full analysis of business proprietors for the 1851-1911 period. This effort is required because, although the census gives comprehensive coverage of non-corporate business proprietors, it does not provide sufficient coverage of directors and their businesses. Normally only a proportion of directors are identified explicitly in census returns as ‘director’, and almost none give any information about the company(s) they direct. This is to be expected since the census had no explicit instruction for directors to record themselves or their businesses.

In order to satisfy the aims of the ESRC project to construct a full database of business proprietors, a substantial exercise of data enrichment was undertaken to bring directors to the level of non-corporate proprietors. The main input for this enrichment which gives the most systematic coverage is the *Directory of Directors* (DoD). This paper discusses how this directory has been used, assesses the level of coverage achieved, and describes how the

director information has been integrated with the main UKDA data deposit for the ‘British Business Census of Entrepreneurs 1851-1911’. Because the DoD began in 1879, it is available only from the 1881 census date, and the paper is restricted to the enrichment of the censuses 1881-1911. Future research can be directed to enriching director coverage for the censuses for earlier years.

At the outset it is necessary to clarify the definition of corporate proprietor used. Directors have been adopted as the identified individuals in corporations because they most directly parallel the proprietors of non-corporate enterprises: they are the decision makers immediately responsible for strategy and although they are protected from some liabilities by corporate status, they are the main risk bearers similar to that of non-corporate proprietors. However, directors were not full equivalents of other business proprietors. Incorporated businesses were technically owned by their shareholders, which could vary from a handful of people to tens of thousands of shareholders. Recent literature has contributed a lot to our knowledge of who these shareholders were in this period (see e.g. Rutterford et al., 2011). There is an extensive literature on the issue of entrepreneurial responsibility in incorporated businesses, which takes various positions about the different roles of shareholders, other investors and lenders, directors, and managers (Chandler, 1990; Hannah, 2007; Cheffins, 2008; Foreman-Peck and Hannah, 2012; Campbell and Turner, 2011; Acheson et al., 2015). However, while shareholders, lenders, and managers all bore risks, legally it was the directors who were the main decision makers and who bore the direct responsibility for strategic business decisions and were in most direct control of the company. Directors were also particularly relevant for 1851-1911 as many domestic corporations were relatively small, and held as de facto private companies with the main shareholders and directors the same people.

The information used to identify directors for data enrichment is partly taken from the census itself where many individuals self-identified as director, and partly from the DoD. The census data was derived from various extractions to identify non-corporate business proprietors and directors using the electronic versions of the census available in the Integrated Census Microdata (I-CeM) deposited at UK Data Archive (UKDA),¹ as discussed in WPs 1-4. The

¹ Higgs, Edward and Schürer, Kevin (University of Essex) (2014) *The Integrated Census Microdata (I-CeM)* UKDA, SN-7481, derived by FindMyPast using a variety of original FMP transcriptions. Version 2 of I-CeM

DoD data was linked to these census records using nominal record linkage, using a semi-automated method that cast a wide net of possible matches and then manually refined them, resulting in an overall linkage of 36 percent.

This Working Paper first describes the data. Section 3 describes how the directors were matched to the census. Section 4 discusses how their companies were coded to provide sector and locational data. Section 5 describes how directors' roles in their companies were coded. An overview of the project and how directors fit into its analysis is given in WP 1. Preliminary analysis of the characteristics of directors compared to non-corporate proprietors is given in Bennett et al. (2019). Further information on the project can be found in other Working Papers listed at the end of this paper. The directors that are record-linked to census data are part of the main UKDA data deposit for the 'British Business Census of Entrepreneurs 1851-1911'. The full database of all DoD directors and coding of their companies is a separate UKDA data deposit.²

2. The *Directory of Directors*

The *Directory of Directors* (DoD) is an annual 'list of the directors of the joint stock companies of the United Kingdom', compiled by Thomas Skinner, who was also the compiler of the Stock Exchange Year Book (SEYB). In its preface, Skinner described the DoD as intended to be used as a companion to the SEYB. Skinner's full method is unclear: the 1912 preface stated that the DoD was compiled from 'the particulars published by the Companies, or other equally authoritative sources' (Skinner, *DoD*, 1912, vi). However, our data processing revealed that the DoD contained companies that were not included in the SEYB, even in cases where the company's directorship was a director's only listing. Around a third of companies in the DoD that were missing in the SEYB in 1912 were found in the *Red Book of Commerce or Who's Who in Business* (Whitaker, 1912), so it is likely that this was an additional source. However, the Red Book only started appearing in 1906, and even after that date there were still companies in the DoD that appeared in neither the SEYB nor

includes a range of valuable additional inputs from colleagues at Campop; see Schürer, K., Higgs, E., Reid, A.M., Garrett, E.M. (2016) *Integrated Census Microdata V.2 (I-CeM.2)*.

² Database deposit is scheduled for late 2019.

the Red Book. The other sources that Skinner probably used was the registration details of companies when first registered and personal contacts to company promoters and lenders in the City. As a well-connected individual he would have had ready access to these sources; and indeed many directors or their administrators may have contacted him to ensure inclusion in the DoD for the later years when it was a well-established source. The method Skinner used to compile the directory was to write to each of his subjects in order to obtain or check information. This makes the DoD a particularly valuable source for census comparison since it should give an accurate and close residential address that can be matched with an individual's census record. Skinner's process of checking information, however, clearly asked directors to provide information on all their directorships. This is a very useful feature since it results in the DoD including many additional directorships and companies not listed on the Stock Exchange; although it also results in inclusion of directorships of some non-corporate bodies which have to be cleaned from the listings. These features of the DoD are valuable since they make it a more complete coverage of companies than available from any other large scale directory source: including companies public and private, listed and unlisted.

The DoD comprises a list of directors and the companies in which they held directorships. It was first published in 1879 and then appeared annually, meaning it can be linked for the censuses 1881 to 1911. Because the DoD appeared early each year with a late addendum, it was essentially a record of the directors for the previous year. Hence, for alignment with the census undertaken in March-April, the following year is taken as the best and most complete comparison; i.e. 1882 used for 1881.

The DoD lists director names in alphabetical order. Entries stated a director's title and name, with either a personal address, or their business, in which case there followed a business address. These addresses varied in their level of detail. For instance, Mr Francis S. Chapman's 1882 entry listed him at '36, Stanhope Gardens, South Kensington, S.W.', whereas Mr Edward Chatfield's address was limited to 'Farnborough, Kent'. When a director's occupation or business was given, the address was usually that of their business, such as the entry for 'Mr Charles Cheston of Cheston & Sons, solicitors, 1, Great Winchester Street, E.C'. In some cases, only the address of the company of which they were director was provided: Mr William Henry Child was a director of the Chelsea Water Works Company at 35 Commercial Road, Pimlico, S.W.

After a director's personal information followed a list of the directorships held. The companies were listed in alphabetical order, which means there is no way in determining which companies were more important to the individual than others, meaning we have no corporate equivalent to the census definition of 'main' occupation. However, the directorship information also included a director's role which allows some measure of higher orders of involvement. The vast majority of roles were 'directors', but there were also considerable number of chairmen, managing directors, and roles on companies' local or other boards. All this information was transcribed and coded.

Since the initial stage of the 'British Business Census of Entrepreneurs' only included England and Wales, it was not attempted to match all directors to the census. Later research will expand linkage to include Scotland. Table 1 shows the number of directors listed in each DoD and the location of their home address if available, and business address if not. Over 80 percent of directors were based in England and Wales, although this declined from 84 percent in 1881 to 81 percent in 1911. The proportion based in Ireland and Scotland remained stable over time, but the proportion of directors based abroad rose from 1 to 6 percent over this period. This mainly concerned directors of large foreign or colonial companies that were listed in London, and some of them also provided a business address in the City of London.

Year	Total Directors	England & Wales	Scotland	Ireland	Foreign
1881	8,572	7,243	911	306	112
1891	12,598	10,473	1,229	519	377
1901	18,727	15,415	1,975	619	718
1911	22,175	18,010	2,189	753	1,223
Total	62,072	51,141	6,304	2,197	2,430

Table 1. Number of directors in the DoD broken down by location of address.

3. Matching the directors to the census

In the first instance, only directors in England and Wales were linked to the census. A special licence to use the I-CeM names and addresses was obtained from UKDA to allow linkage. There were several inconsistencies between the way the census recorded names and

addresses, and how they were listed in the DoD. First, the DoD supplied the names either with first and middle name fully written out, first name written out and middle name initial, only initials with the surname, or titles (such as ‘the Duke of Abercorn’). The census on the other hand mainly provided first name written out, middle name initialled, and surname. Some enumerators wrote out the middle name, and some people only had initials, but this was rare. Titled people were enumerated under their personal name, i.e. the Duke of Abercorn appeared in the census as ‘James Hamilton, marquis and magistrate’. Secondly, the DoD supplied addresses as either ‘number, street name, neighbourhood, city, county’, with London addresses also featuring a postal district, or as ‘house name, town, county’ – the latter more commonly when the address was a hall, mansion, or a business address such as paintworks. The census address data on the other hand usually provided a number and street, with all other information being part of the census spatial administration units, such as parish, registration sub-district (RSD), registration district (RD), and county. Hence, an address that in the DoD was listed as Liverpool, for instance, could be in the census RD West Derby. Finally, as mentioned previously, not all of the DoD addresses were precise, with non-urban addresses in particular only giving a village and the county, such as the Kent example above. However, even some urban addresses were noted as just ‘Nottingham’, or ‘Sheffield’. These differences mean that straight text string linkage by algorithm was impossible, even when applying fuzzy matching.

In order to link the data, therefore, both the names and the addresses were pre-processed and edited. As the format of first and middle name varied so much, for both census and DoD the first name initial was generated for matching. Surnames were more regular; however, due to miss-keying in both the transcription of the census and the transcription of the DoD and slight variances in formatting in double or hyphenated surnames, a conversion to Soundex was preferred in the first instance. Soundex is a phonetic algorithm consisting of one letter and a set of three numbers that codes names by sound, allowing similar names, such as Smith and Smythe, or Johnson and Jonson, to be matched despite minor differences in spelling.³ Soundex was developed to control for transcription of verbal material. It has been criticised for record-linkage, but for the purposes used here, when controlled through spatial blocking

³ Soundex is based on English phonetics. While it does not always convert foreign names well, this error would occur in both the DoD and the census sample meaning the same code was generated for both, and they could still be matched.

and other choice criteria, and when closely integrated with intensive clerical scrutiny, it proved a robust tool to eliminate typos and minor spelling differences from the datasets. It was more appropriate to this application as well as requiring less processing time than other text-string comparison tools, such as Jaro-Winkler. All census surnames and all director surnames were converted to Soundex.

An initial trial used an existing gazetteer and a GIS to generate detailed locational data for the 1882 DoD directors, creating six-digit X and Y codes. These could be generated for census people as well, based on the centre of their parish. Matching people on the first two digits of the X and Y data meant that they were located in the same 10 km² area. This range provides a means for spatial filtering to break the data into blocks adapted to each individual's detailed location. This is important to achieving good matches by reducing the number of false positives. A full match between the first name initial, surname Soundex, and the same locational area then provided a wide range of possible matches, which were manually narrowed down based on all other information. A trial on a set of northern directors in the 1882 DoD allowed a third of possible directors to be matched.

The pre-processing of locations was very time-consuming, and the method did not work well in larger urban areas that spanned more than 10 km², such as London, home to a large proportion of directors. For this reason, a second trial used the urban classification system adopted for the rest of the ESRC project (Smith et al., 2018; see also WP 6). This used a second form of spatial data blocks to assign all urban directors in the DoD to their city code, to be linked with all people in the parishes associated with that urban coding in the census. Non-urban directors were coded to their county, and matched within spatial blocks for all the non-urban (urban classification codes 2-4) people in the census for that county. As Table 2 shows, the vast majority of directors were urban. This method also allowed a third of directors to be matched, although London had a worse success rate. This is likely to be due to the fact that a large number of London DoD addresses were business addresses rather than home addresses, meaning that the directors would not have been present at their DoD address on census night. Conversely, non-urban addresses had a much better success rate of close to 50 percent, as these were predominantly home addresses.

	E&W total	London	Other Urban	Non-Urban
1881	7,243	3,404 (47%)	2,446 (34%)	1,393 (19%)
1891	10,473	4,953 (47%)	3,928 (38%)	1,592 (15%)
1901	15,415	7,081 (46%)	6,078 (39%)	2,256 (15%)
1911	18,010	7,903 (44%)	7,084 (39%)	3,023 (17%)

Table 2: England and Wales directors in London, other cities, and rural locations, with their percentage of the total in brackets

In order speed up the manual checking of possible matches, a first match was made on exact surname and initial, within the spatial unit specified above, with a second run using Soundex on the remainder only. The vast majority of matches were made using this method. Once it was completed, a separate match was performed on all titled directors, who were compared to an extraction of census people using either a title, or an occupational description that identified them as lord, baron, or peer. Finally, a reverse matching process used the first name initial and Soundex system to link people who were enumerated as a director in the census but not yet linked, to a director in the DoD regardless of location. This picked up people who were not at home on census night, as well as directors who had a Scottish, Irish or foreign address in the DoD but were actually in England or Wales on census night, but only if they described themselves as a director in the census.

After generating a large number of possible matches using the methods described above, the correct match was chosen manually. The manual checks compared full first name, middle name (if available), surname, and address, and used DoD title, business or first listed company descriptors to compare with the census occupation. A hierarchy of matching was created to reflect accountability of each match:

Match 1: The top-ranking matches provided either an exact match on first name, surname, and full street address, or first name, surname, and identification as a ‘director’ in the census occupation.

Match 2: The second-tier matches were as Match 1, but with two people of the same name living at the same address, meaning that a DoD director could be linked to either a father or a son in the census. While both options have been preserved, the

final data has linked all DoD directors to the oldest of the two – while this might introduce age bias it follows the trend for directors in general to be older and only applied to only a few people (1-2 percent of the matches).

Match 3: Third- and fourth-tier matches did not match the exact street address, usually because it was not available in the DoD, but were still positively identifiable within the possible matches using exact name and location (city if urban, county if not), and either matching census occupation to a director's title or occupation (MP, J.P., colonel, etc.), an unusual name, or a matching middle name. Third-tier matches were identified within this sample based on at least three criteria.

Match 4: was the same as Match 3 but using matching on fewer than three criteria; it was only applied in cases where clerical scrutiny judged them close enough to be accurate matches.

As shown in Table 3, the majority of matches were Match 1; the majority of the rest used Match 3. It should be kept in mind that these do not necessarily measure the quality of a match, only the strength of information on which the match was made. For instance, a director match with a very unusual name or a unique title was still allocated Match 3 if the address did not match up exactly, which was often the case in DoD entries where the address information was limited.

	1881	1891	1901	1911
Number Matched	2,704	3,348	5,107	7,041
% Matched	37	32	33	39
% Match 1	53.5	53.9	62.9	68.0
% Match 2	1.3	2.3	1.4	1.2
% Match 3	32.6	29.5	23.5	20.0
% Match 4	11.4	6.8	4.4	2.8
% Titled	0.2	0.5	1.0	0.2
% Reverse match	1.0	4.1	6.8	7.8

Table 3. Number, percentage, and accountability of match for all years (England and Wales matches only).

Table 3 shows the final results of the matching process and the type of matches made. The increase in type 1 matches is due to both better recording of exact addresses in later DoDs, and more directors being enumerated as director in the census. In 1881 only 162 of the 2,704 directors linked to the census identified themselves as director in their census occupation, while in 1911 this was 2,410 out of 7,041, or in other words, the proportion grew from 6 percent in 1881, to 12 percent in 1891, 25 percent in 1901, and 34 percent of DoD directors in 1911 who identified themselves as directors in the census. This is also visible in the increase in reverse matches: this is simply the result of a larger amount of people available for matching in later years. This better census enumeration of directors probably reflects rising familiarity of both respondents and enumerators with the terminology of director; though in 1911 it may also reflect the use of the original householders' returns in the census records rather than the Census Enumerators Books (CEBs). Reverse matching increased the overall percentage of matches, which increased from 32 percent in 1891 to 39 percent in 1911. 1881 bucked this trend, but this is mainly due to the results of the pilot studies included here, which were slightly more fruitful but much more labour-intensive. The overall 36 percent match rate is very good for nominal record linkage considering the limited data available, the variable quality of this information, and the high confidence level of the matches achieved. Most record linkage relies heavily on age as a discriminating match criterion, which is not available in DoD; if it had been recorded it would have been of considerable help to increase match rates as it would have reduced the range of false positives occurring with common names that had poor addresses which had to be excluded in the matching here because there was not enough information to narrow them down.

4. Coding the directors' companies

Table 4 lists the number of directorships held by all directors (matched and unmatched). The total number of companies is less easy to measure, due to errors in transcription, different ways in describing the same company, and directorships in multiple local branches or divisions of a single company. While effort has gone into consolidating company names as much as possible, a full editing of company names would have required extensive research beyond the scope of the current objectives. It is hoped that future researchers will wish to use the database deposit to increase the accuracy of the corporate database now provided. All

directorships, both those from matched directors and unmatched directors, have been coded, although priority for checking was given to companies belonging to matched directors.

Year	Directors	Directorships	Companies
1881	8,572	14,681	ca. 3,100
1891	12,598	22,723	ca. 5,500
1901	18,727	32,432	ca. 8,400
1911	22,175	40,544	ca. 11,200

Table 4. Directors, directorships, and companies by year

To provide a comparator against the non-corporate proprietors derived from the census in the rest of the ‘British Business Census of Entrepreneurs 1851-1911’ database it is important to have similar identification of the business sector in which the company operated. It was also important to be able to separate companies with predominantly domestic activity (within the UK, within England and Wales, or Scotland) from those primarily operating abroad (and indeed from those that were entirely foreign businesses that are included in the DoD only because they have a UK Stock Exchange listing). These codings allow sector comparisons between corporate and non-corporate proprietors, and on a common basis to be able to evaluate important research questions about differences between individuals involved entirely with the domestic economy, and those involved with ex-UK activity. The use of the criterion of domestic and non-domestic location of main activity differs from most previous analyses of companies but is specifically designed to provide a valid comparator against the domestic non-corporate sector. For both sector and location the companies were coded to their ‘main’ activity, recognising that in larger and more complex businesses this will be at best an approximate identification.

Both sector and location are difficult to define precisely in some cases given the information not readily available; however, it is believed that a reasonably accurate final coding has been achieved using a four-stage strategy:

Stage 1. This used the company name, which in many cases provides a direct and accurate identification for either sector, location, or both; e.g. the Madras Railway

could be coded directly to Railways as a sector and Colonial as its location of its main activity.

Stage 2. Those that were uncertain after the first stage were looked up in the SEYB by a research assistant. This provided information on most businesses for which names alone were insufficient or uncertain. Only a small proportion of companies could not be found; this ranged from 5-7 percent in the 1881-1901 censuses, to 12 percent in 1911. This residual for the missing 1911 companies was subsequently then checked in the *Red Book of Commerce*, in which many more were found; this indicates that that this was probably used as a source by Skinner. However, even if a SEYB or *Red Book* listing was found, it sometimes did not give sufficient information on sector or location. Hence this group was then entered into Stage 3.

Stage 3. Substantial additional efforts were made by the authors to identify the remaining unknown companies in a range of different sources such as *Grace's Guide to British Industrial History*, for mines and quarries in *Mineral Statistics*, and using a wide range of online sources, with a particular focus on companies belonging to matched directors. This was extremely successful in reducing the residual and also correcting some errors from the previous stages.

Stage 4. A final check was passed to external experts who made a quick scrutiny to spot obvious mis-codings or omissions that could be readily filled. This provided information on only a few companies but was an important reality check glaring errors.

In the final database, only a small percentage (between 0.4 and 1.1 percent, depending on year) of companies remained un-coded in terms of sector or location. However, it is clear that in some cases the coding may have residual errors and this is an area where further inputs from other researchers will be valuable in updating the database deposit in the future. It is also accepted that in some cases of complex business with multiple operations in different sectors, and/or operating in a range of different locations, any simple coding will have limitations. Future researchers can improve on these codings in the database.

4.1 Sector

For the sector, the main goal was to match the coding used for the non-corporate business owners from the census: the EA 17 codes for economically active (see WP 5). This has 13 business sectors (the remaining 4 being non-business codes). However, in order to reflect the nature of British corporate business during this period several changes were made to EA17. This created a new Company Coding (CC) which replicates the same codes used as EA, but includes a few additions to reflect the large numbers of businesses in some company sectors. There were five modifications to disaggregate the EA 17 codes. First, the utilities sector (water, gas and electricity supply) mainly existed in corporate form and had been mainly included with manufacturing in EA 17 (sector 4). This received its own CC 18 sector – this is most comparable to the more disaggregated EA 51 subsector 12 (see WP 5). Second, communications, such as telegraphs, were separated into their own CC 21 sector. Third coal was split from general mining as CC 19; and fourth, railways were separated from other transport as CC 22, in order to reflect the size and importance of these sectors for corporations. Fifth, steel and chemical manufacturing was split from other manufacturing as CC 20, since these were major corporate fields generally representing manufacturing on a different scale from most of the businesses included in EA17 sector 4. Table 5 shows the complete list of CC, subsectors, as well as details of the companies included. The numbering system is identical to the EA 17 codes, extended to cover the five disaggregated sectors CC 18 – 22. Because of the small company numbers involved, the CC codes combine the EA 17 classification for sectors 8 (Professional & business services) and 9 (Professional & personal services) under CC 9; and combine EA 17 sector 11 (Food retailing) under CC 6; and sectors 14-17 of EA 17 are non-business economic activities (domestic service, public administration, military, etc.)

CC	CC Sector	CC Subsector	Details
1	Farming/estate work	Coffee tea and rubber	Incl. plantations (sugar, cotton)
1	Farming/estate work	Farming fishing	Also some land if not distinguishable
2	Mines and quarries (non-coal)	Mines - metal	Gold and silver, lead, tin etc.
2	Mines and quarries (non-coal)	Mines - oil	Some might also be refineries
2	Mines and quarries (non-coal)	Mines - other	Nitrate, potash, combined mines
2	Mines and quarries (non-coal)	Mines - precious stones	Mainly diamonds, some other gems

2	Mines and quarries (non-coal)	Mines - stone, brick, clay, slate	Mainly quarries, includes cement
2	Mines and quarries (non-coal)	Mines - unknown	
3	Construction	Construction	Incl. house building
4	Manufacturing	Mf - cycle, motor, automobile	
4	Manufacturing	Mf - engineering/transport	Shipbuilding, waggon mf.; also engineering in general
4	Manufacturing	Mf - leather/footwear/apparel	
4	Manufacturing	Mf - metal goods	[a few armaments if not included elsewhere]
4	Manufacturing	Mf - other	Incl. ice mf.
4	Manufacturing	Mf - paper	Pulp etc.
4	Manufacturing	Mf - print/press/publish	Newspapers, presses, publishers; one author
4	Manufacturing	Mf - textiles	
4	Manufacturing	Mf - timber/furniture	
5	Maker dealer	Maker dealer	Incl. photographers, jeweller / goldsmiths, ironmongers. Mainly named companies
6	Retail and wholesale	Retail - co-op	
6	Retail and wholesale	Retail - other	
6	Retail and wholesale	Retail - warehouse	Incl. cold storage
7	Transport	Canals and docks	
7	Transport	Shipping, haulage & carriage	Mostly sea transport, a few land incl. taxis; incl. oil shipment etc.; other
7	Transport	Tramways and omnibus	Incl. a few road maintenance, and the Glasgow and Woolwich subways
9	Personal and business services	Services - entertainment	Incl. theatres, piers, assembly rooms, rinks, public pools
9	Personal and business services	Services - medical	Mainly hospitals, sanatorium
9	Personal and business services	Services - other	Incl. business services, associations & clubs, auctioneers, and personal services such as laundries
9	Personal and business services	Services - schools	
10	Agricultural produce and dealing	Agricultural produce and dealing	Incl. breweries, drinks, tobacco manufacture, sugar, distillery, dairy, bakers, oil & flour mills
12	Refreshment	Lodging and refreshment	Lodging, hotels, restaurants, hydropathic spas; may include some wine merchants
13	Finance and commerce	Banks	Incl. building socs.
13	Finance and commerce	Financial, land and investment	Mainly from SEYB category which includes some mine, construction, farm, agric.

			processing; classified to other CC as far as possible; other
13	Finance and commerce	Insurance	Incl. friendly socs
18	Utilities and Municipal	Cemeteries	
18	Utilities and Municipal	Electric supply and lighting	
18	Utilities and Municipal	Gas	Where distinguishable
18	Utilities and Municipal	Government	Incl. corporations and board of works
18	Utilities and Municipal	Market and tolls	Incl. Exchange buildings; a few other
18	Utilities and Municipal	Waterworks	Where distinguishable
19	Coal mining	Mines - coal	Collieries, coal mines, some in commerce in SEYB but reclassified as far as possible if mainly extracting
20	Steel and chemical mf	Mf - chemical	Incl. dye, explosives, drugs mf. etc.
20	Steel and chemical mf	Mf - Coke iron steel	Incl. 'coal and iron' if clearly not a mine
21	Communications	Communications	Cables and telephones; also a few advertising, and bill posters
22	Railway	Railway	Incl. double function companies ('railway and docks', 'railway and mines') if clear main function is railway OR if unclear and railway mentioned first
0	Unknown	Unknown	

Table 5. Company sector codes (16 CC categories, plus unknown).

4.2 Location

The companies were coded on the location of their main activities. In some cases this was obvious from the company name, but the majority was sourced from information in the SEYB and online searches of other sources. Coverage of this varied; for instance, the 1881 SEYB did not always provide the location, so more were missing for that year after Stage 1 and 2 of processing. Where no other better locational information was available, a simplified coding system was used to code into E&W and non-E&W; and also into UK and non-UK. The main location categories are listed in Table 6.

Location
England and Wales
Scotland
Ireland
Not UK
Channel Islands and Isle of Man
Colonial (British)
European
USA
Foreign (other)
International (multiple territories; can include UK)
Unknown

Table 6. Locations used in company coding

For some analysis purposes this was simplified to a smaller set, shown in Table 7.

Location
England and Wales
Rest of UK and islands: Scotland, Ireland, Channel Islands and Isle of Man
Colonial (British)
Foreign: European, USA, Not UK
International (multiple territories)
Unknown

Table 7. Simplified locations used in company coding

4.3 Listed

An additional code was based on whether a company was ever-listed in the SEYB or not. This code was based on whether the company was listed in the SEYB in any of the years studied – which means that some companies coded as listed in 1881 may not actually have been listed until later in the period. This is a simplification that could be refined by other researchers if desired. Companies not found in the SEYB were considered unlisted. Most of

the group of around 5,000 companies that were coded at Stage 1 without consulting the SEYB, such as the Madras Railway example above, were left as unknown in terms of listing. This again could be refined by other researchers if desired.

4.4 Reducing director sectors and locations to single codes

The DoD lists the companies with which a director was associated in alphabetical order with no other data. Hence, there is no way in determining which companies were more important to the individual than others. Many directors had portfolios of two or more directorships. But unlike the non-corporate proprietors where the census definition of ‘main’ occupation gives a way to identify the most important sector, there is no way to order these by main or primary (unless detailed research is used on each director’s activities). Nevertheless, for many purposes of analysis it is important to assign directors to a single sector or location code. The full database retains all the information on all companies. But a ‘ranked’ coding is also constructed for the most frequent company types. This was achieved on the basis of the **frequency of sectors or locations** as follows:

- Any unknown sectors/locations ignored for this classification
- All directors with just one sector CC or location code: coded to that company’s characteristics.
- All other directors with more than one known company code: frequency of each CC sector and location code calculated; greatest frequency gives the first ‘ranked’ code; this process continued for second most frequent; any equal frequencies assigned at random.

In addition the interaction between sector and location was coded:

- Most frequent sector – its most frequent location (if more than one); if any equal frequencies, then location assigned at random.

The basis of the sector codes was the 16 CC categories in Table 5. In practice this assignment was relatively straightforward since most directors were sectorally specialised: they operated in sectors that were mostly the same. Also only a small number had very numerous directorships and hence multiple categories. As a result, in almost all cases the classification was direct with no ambiguities.

The result is data codes:

- Most frequent sector
- Second most frequent sector
- Most frequent location
- Second most frequent location
- Most frequent location of the most frequent sector
- Most frequent location of the second most frequent sector

Further ranked frequencies are possible from the database as required, though most directors are fully described by the first and second ranked frequencies.

5. Coding director roles and other attributes

5.1 Roles

In addition to the companies, the DoD contains information on the directors' roles in their companies. The vast majority were just termed 'directors', but there were also chairmen, managing directors, council members, trustees, and many other role titles. Table 8 lists the consolidated roles and the director titles as listed in the DoD that have been included. Other titles given in DoD have either been cleaned out as spurious, or recoded as indicated under the description column of the table. Roles of directors were assigned in a hierarchical manner in downwards order of Table 8: so, for instance, if someone was a chairman and director of a local board they were coded as chairman, although in practice there may not have been any actual difference in power of control between a general director and a chairman (who in many cases could be a nominal title difference or was someone who took no active role in the company management and decisions but was brought in to give status to the business – as with many titled individuals). Presidents, which mainly related to non- England and Wales companies, were generally coded under chairmen as indicated in the table, but in England and Wales were coded as presidents in a category with other honorary roles. Certain companies, such as investment trusts and listed schools, only had trustees or members of the council; since these acted in a similar role to directors they were coded as such. Ex-officio positions of directors and chairmen were treated as regular positions.

Category	N	Description
Managing director (MD)	5,173	The person normally with most control of company: MD; general manager and director; partner; proprietor; 'executive directors'; 'on board and managing'; any other combination of 'managing' and director, board, or chairman.
Managing director: joint or local	1,304	Like MD but less powerful: joint MD; assistant MD; 'MD in Western Australia'; European MD; majority are 'joint'.
Chairman	15,564	Directors who are nominally senior, but not necessarily more powerful than rest of the board. Includes Governor (e.g. of Bank of Scotland), and President of Scottish, US, Canadian, French and other European companies.
Vice Chairman: deputy/local board	3,031	Directors who chair subsidiary company boards or committee; e.g. 'Chair of the Glasgow Board' or 'chair of the finance committee'. Includes Joint Chairman; deputy-governor; Vice-President of Scottish, US, Canadian, French and other European companies.
Director	79,697	Includes 'board'; 'member of board'; member of the council, committee members of General Committee, Committee of Management, Advisory Committee, extraordinary directors, baronial directors, technical director, sole director, acting director. Includes trustees where companies just has trustees (e.g. investment trusts).
Director local board	4,543	Includes members of special boards and committees, but mostly location-based boards. Sometime ambiguous, as may be full directors taking on an extra role, or of lower status where only on a separate board in a specific location; e.g. London Committee, Scottish Board, Nottingham local board.
President/VP/honorary	144	Mix of honorary appointments, including president and VP of England and Wales and colonial companies; Hon, directors, and hon. Chairmen.
Representative	289	A representative for a stakeholder/group of stakeholder on the board, sometimes a specific 'trustee'; e.g. Representative of the Manchester Corporation; Trustee representative of debenture holders; any other 'trustees'.
Other	635	'Concerned in', and other non-director roles. The majority are agents, secretaries, and managers, but this also includes a few auditor, treasurer, engineer, auctioneer, etc. Some of these might be spurious and might be excluded.
Total	110,380	

Table 8. Director roles.

Due to the DoD formatting and phrasing, it was not always clear whether a role/company combination was part of a directors' occupational descriptor, or was a form of directorship.

Clear spurious entries that were erroneously transcribed, such as ‘Dean of St Paul’, have been removed. Some other roles have been conserved, and can be included or excluded by researchers depending on the analysis. For instance, ‘representatives’ were included as directors on boards but may not have had voting or any controlling powers in all companies. They could be interesting for network analysis, but they would need to be excluded for a corporate control analysis. Similarly, it was unclear if some of the people described as ‘secretary’ or ‘treasurer’ took on this role in their capacity as director, or were just hired secretaries, in which case they would be spurious.

5.2 Titled directors

Directors were also coded to a separate category where they had a high status title. Many such individuals were invited to take roles in companies because of their aristocratic or political status and connections. They form an important sub-category that can be used for network analysis and other assessments of how company directors operated. Directors were coded using the title they gave in their personal description in the DoD. The completeness of these titles depends on how far they were used by the respondents to Skinner’s request for information, or Skinner used the information from his other sources. It is believed that this coding should be fairly complete, since most such individuals were proud of their titles and used them as part of their normal address and business activity in response to such enquiries. This was period where such titles and the correctness of addressing individuals was a normal act of everyday business life, with guides such as *Debrett’s* devoted to setting out how individuals with various levels of title should be addressed. Titles were coded to categories for: Aristocratic; Political (MPs); Army; and non-titled. Other titles provided in DoD, such as J.P., Hon., Q.C., Dr., Rev., etc. were left as non-titled.

5.3 Other attributes

The DoD gives no further information on directors that can be used on a systematic basis. However, those directors that were linked to the census have the full range of their census responses available. This is one of the great benefits of record linkage since it gives the following data on each director:

- Age

- Marital status
- Relationships with the rest of the family and household where present on census night
- Number of servants employed
- Other occupational information (where given)
- Local parish or district within which present (usually resident) which can be linked to other spatial data
- Other rank and title data that may be of value in a few cases.

Gender is also given in the census, but this is in almost all cases no more than a confirmation that can be derived from the forenames given in the DoD; though it is critical to confirming that a few of the directors identified were female.

6. Conclusion.

This paper describes how data derived from the *Directory of Directors* can be used to link to, and enrich, census records to work towards a complete coverage of all business proprietors for 1881-1911. The paper has shown how directors can be matched to the census, how their companies were coded to provide sector and locational codes, and how directors' roles in the companies were coded. The method adopts a matching approach combining automatic searching on names and key attributes, with blocking by locations, and with considerable clerical intervention. The final accepted matches have to reach a high level of confidence, with multiple positives leading to a decision not to accept the match. The final match rate achieved averages 36 percent across the four census years.

The results of the matched data have been added to the full database of all business proprietors identified in the censuses for this period as part of the main UKDA data deposit of the 'British Business Census of Entrepreneurs 1851-1911'. It is accepted that there are limitations to the matches that have been achieved; and also to the accuracy of some of the sector and location codes. The database deposit allows other researchers, who can deploy additional information and resources, to add to and improve the coding and to seek further matches for the rest of the DoD where a sufficiently certain positive match has not been identified by the methods used here. It is hoped that the database and this paper open the way for continuous improvement of the director and company database as an ongoing resource for research in this field.

Acknowledgements.

We record our thanks to Walter Jansson for checking DoD entries against the *SEYB* and *Red Book*. Les Hannah, Janette Rutterford, and Naomi Lamoreaux gave helpful advice on the coding of director roles; Les Hannah also very helpfully did a quick check on final company codes at Stage 4 detailed above, and also advised on the sector and location codes. Any residual errors and omission remain our own.

This research has been supported by the ESRC under project grant ES/M010953: **Drivers of Entrepreneurship and Small Businesses**. Piloting of the research and data cleaning for 1881 draws from Leverhulme Trust grant RG66385: **The long-term evolution of Small and Medium-Sized Enterprises (SMEs)**, with the support of Gill Newton.

The census database derives from K. Schürer, E. Higgs, A.M. Reid, E.M Garrett, *Integrated Census Microdata, 1851-1911, version V. 2 (I-CeM.2)*, (2016) [data collection]. UK Data Service, SN: 7481, <http://dx.doi.org/10.5255/UKDA-SN-7481-1>, enhanced, see: E. Higgs, C. Jones, K. Schürer and A. Wilkinson, *Integrated Census Microdata (I-CeM) Guide*, 2nd ed. (Colchester: Department of History, University of Essex).

References

- Acheson, G.C., Campbell, G., Turner, J. and Vanteeva, N. (2015) Corporate ownership and control in Victorian Britain, *Economic History Review*, 68, 911–36.
- Bennett, Robert J., Smith, Harry, van Lieshout, Carry, Montebruno, Piero, and Newton, Gill (2019) *The Age of Entrepreneurship: Business proprietors, self-employment and corporations since 1851*, Routledge, London.
- Campbell, G. and Turner, J. D. (2011) Substitutes for legal protection: corporate governance and dividends in Victorian Britain, *Economic History Review*, 64(2), 571–97.
- Chandler, A.D. (1990) *Scale and Scope: The Dynamics of Industrial Capitalism*, London: Belknap Press.
- Cheffins, B.R. (2008) *Corporate ownership and control: British business transformed*, Oxford: Oxford University Press.
- Foreman-Peck, J. and Hannah, L. (2012) Extreme Divorce: The Managerial Revolution in UK Companies before 1914, *Economic History Review*, 65(4), 1217-1238.
- Hannah, L. (2007) The ‘Divorce’ of Ownership from Control from 1900 Onwards: Re-Calibrating Imagined Global Trends, *Business History*, 49(4), 404-438..

- Rutterford, J., Green, D.R., Maltby, J. and Owens, A. (2011) Who comprised the nation of shareholders? Gender and investment in Great Britain, c. 1870-1935, *Economic History Review* 64(1), 157-187.
- Schürer, Kevin and Woollard, Matthew (2000) 1881 *Census for England and Wales, the Channel Islands and the Isle of Man (Enhanced Version)*, UKDA, University of Essex, SN-4177.
- Schürer, Kevin, Higgs, Edward, Reid, Alice M., Garrett, Eilidh M. (2016) *Integrated Census Microdata V.2 (I-CeM.2)* [data collection].
- Smith, H., Bennett, R.J. and Radicic, D. (2018) Towns in Victorian England and Wales: a new classification, *Urban History*, 45(4), 568-594.

Other Working Papers:

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.

- 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: <https://doi.org/10.17863/CAM.9874>

- 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>
- 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>
- 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>
- WP 9: Bennett, Robert, J., Montebruno, Piero, Smith, Harry, and van Lieshout, Carry (2018) *Reconstructing entrepreneurship and business numbers for censuses 1851-81*.
- 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>
- 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>
Adjustment weights: <https://doi.org/10.17863/CAM.26376>
- 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*.
- WP 14: van Lieshout, Carry, Bennett Robert J. and Montebruno, Piero (2019) *Company Directors: Directory and Census record linkage*.
- 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*.

Full list of all current Working Papers available at:

<https://www.campop.geog.cam.ac.uk/research/projects/driversofentrepreneurship>