# Changes in Victorian entrepreneurship in England and Wales 1851-1911: Methodology and business population estimates

***Bennett, Robert J.,[1] Smith, Harry,[2] Montebruno, Piero,[3] and van Lieshout, Carry[4]***

[1] University of Cambridge, Department of Geography & Cambridge Group for the History of Population and Social Structure, Downing Place, Cambridge, CB2 3EN, UK rjb7@cam.ac.uk

[2] Kingston University, Department of Geography, Geology and the Environment, Kingston upon Thames, KT1 1LQ, UK harry.smith@kingston.ac.uk

[3] London School of Economics, Spatial Economics Research Centre, Houghton St, London WC2 2AE, UK P.F.Montebruno-Bondi@lse.ac.uk

[4] Open University, Department of Geography, Milton Keynes, MK7 6AA carry.van-lieshout@open.ac.uk

**Accepted final version to be published in *Business History* 2021**

**Keywords:** Entrepreneurship, Employers, Self-employment, Small businesses, Census

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

***Acknowledgments:***

This research was supported by ESRC project grant ES/M010953: 'Drivers of Entrepreneurship and Small Businesses'. Data for 1881 draw from Leverhulme Trust grant RG66385: 'The long-term evolution of Small and Medium-Sized Enterprises (SMEs)'. Coding the 1871 census data was supported by the ESRC and Isaac Newton Trust research grant 17.07(d): 'Business Employers in 1871'.

The database used for 1851-1911 derives from a Campop updated version of Schürer, K. and Higgs, E. (2014) *The Integrated Census Microdata (I-CeM)*. Essex: UKDS, SN-7481. https://beta.ukdataservice.ac.uk/datacatalogue/doi/?id=7481. We are grateful to Kevin Schürer, Alice Reid and others at Campop for advice and permission to use this version. The figures were drawn by Phil Stickler. We are grateful for referee comments and various conference feedback on earlier versions of the text.

# Changes in Victorian entrepreneurship in England and Wales 1851-1911: Methodology and business population estimates

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

## 1. Introduction

As noted by numerous commentators, we have lacked adequate statistics of business proprietorship in nineteenth- and early twentieth-century Britain. Hannah (2007: 415) likened this period to a 'statistical dark age' while Jeremy (1998: 331) concluded that secular comparisons are not to be trusted. There has previously been no 'truth' data available to understand the scale of the business population and changes of its composition. This has prevented reliable overviews of business history for the period. However, the release of an electronic version of the individual-level manuscript census data for 1851-1911 allows the data gap to be filled using the *original* census responses. Each census asked business proprietors to identify themselves in some way. From 1891 until modern censuses the question asked people to define themselves as employers, own-account proprietors or workers. The potential of these censuses for the 1891-1911 period has been demonstrated in Bennett. et al. (2018a). An initial methodology for covering the earlier censuses for 1851-81, which identified employers and only partially covered own-account, has been used in Bennett et al. (2019a). This paper extends that 1851-81 methodology, presents alternatives, and tests robustness of the estimates.

The paper has two aims: first, to develop and test alternative methodologies for 1851-81 to join up with 1891-1911, to estimate the whole population of proprietors for the entire period 1851-1911. The second aim is to use these new estimates to reassess key elements of previous literature. The paper is primarily methodological, but it also seeks to show the value of the data by addressing three previous research debates that have been stymied by lack of full national data: the so-called 'declinism' of the Victorian period; the evolution of Victorian entrepreneurship into a 'U'-shaped trough over the twentieth century with the associated onset of a 'climacteric' around 1900; and reassessing compositional effects on sector markets and gender participation.

A key contribution of the paper is making available a means to examine the entire business population, not just large-scale firms: to include what Payne (1988: 22) described as the 'regiments of the anonymous'. Smaller firms have been subject to little analysis at a scale beyond case studies. Yet, as Hannah (2007; 2014) observes, small businesses were remarkably persistent, while Marshall (1919: 314), thought that for this period they were 'the representative firm in most industries and trades', a belief echoed by Pollard (1965: 233). Whilst studies of larger firms have characterised much British business history (Landes, 1969; Crouzet, 1985; Wardley, 1999; Hannah, 1983, 2014; Jeremy, 1991; Nicholas, 1999, 2000), such studies have significant dangers of selection bias towards the large, successful and those with surviving records. As Nye (1991) argues, if all struggling, failed and would-be entrepreneurs were included, the probability of success would be better understood. The paper shows the advantages of whole-population data for overcoming such biases.

The paper first summarises the three main debates addressed. It then details how the 1851-1881 censuses collected information, and how methodologies can be developed for identifying the population of proprietors. The resulting proprietor data is deposited at the UK Data Service (UKDS) as *The British Business Census of Entrepreneurs* (BBCE) (Bennett et al., 2020). The following section of the paper briefly addresses the three debates by examining trends in entrepreneur numbers, rates, and composition. The paper then evaluates the robustness of estimates by comparing with the main whole-population comparator available: national and local trade directories. The discussion is confined to England and Wales but can be extended using the full BBCE data to include Scotland.

## 2. How new data can give new insights to debates on the British business population

The new information on business proprietors that is developed in this paper provides a valuable resource that can contribute to a wide range of debates in business history. Because of the challenges of using the census data the paper is primarily methodological. But it also illustrates the value of the data by showing how it opens up new insights into three key debates. The first is reassessment of claims about the decline of British entrepreneurism over the Victorian period. Explanations for Britain's relative economic decline after 1870 have frequently been argued to derive from a decline in entrepreneurship (for critical reviews see Edgerton, 1997; Tomlinson, 2009). Some argue that business proprietors became focused on property ownership and the trappings of gentility, at the expense of the 'entrepreneurial

ideal', as Perkin put it (Perkin, 1969: 436; see also Wiener, 1981; Colman and Macleod, 1986). Others have argued that British entrepreneurs failed because they were too conservative, remaining in control of family firms too long, which limited their growth potential (Rubinstein, 2006; Chandler, 1990). Yet others argue that many proprietors were adequate managers, but failed to press for improvements of the institutional environment which would have allowed Britain to compete in the global economy of the twentieth century (Lazonick, 1983), or were too focused on short term and easy paths (Landes, 1969).

Given previous data limitations it is not surprising that it has been difficult to assess the truth of these claims. Inevitably the sources used, based on aggregate data, case studies of a handful of firms, or recourse to literary sources pick up only strands of the larger whole. It has not been previously possible to consider the general population of firms, or whether the behaviour of case study firms was common. Even amongst the largest firms, only a handful have left archives of the kind that allow analysis of whether or not their proprietors abandoned the 'entrepreneurial idea', suffered from sub-standard management, or failed to recognise the changing global economy. In this sense previous authors have been unable to adequately understand the selection biases and characteristics of case studies within the behaviour of the population as a whole. This has inevitably led to arguments that are unsustainable. For example, how can the allure of gentility have led people away from the 'entrepreneurial ideal' when the new data show that small and medium-sized firms were the vast majority, increased in numbers over the period rather than declined, with increasing rates of involvement by the population, and were mostly run by individuals who were unlikely to come into contact with the institutions that encouraged such a culture, and were more likely to be affected by the changing opportunities of waged labour than aspiring to 'gentility'?

The second debate addressed here is related to 'declinism' and has been important in modern research on small firms and business policy: that a 'U'-shaped distribution developed over the mid-twentieth century, with declines in small business and self-employment numbers occurring sometime early in the century and then being reversed only by recent growth. From this, various policy conclusions have been drawn as to causes. In Britain the declining limb of the 'U'-shape appears to have been first noted by Clark (1957). It has been confirmed as a trend in most advanced economies and has been used by Wennekers et al. (2010) and Carree et al. (2007) to suggest that across many countries entrepreneurship changed as a result of shifts in the sector composition of economic development: primarily, that as agricultural

proprietorship declined this shed labour that fuelled expansion of other sectors. However, perhaps more important in Britain was the contemporaneous recognition in the Macmillan Report (1931), echoed in the benchmark Bolton Report (1971), that small businesses were suffering from the effects of industrial concentration into large firms which starved small firms of access to capital and other resources. Previously there has been no detailed research possible on how these twentieth century trends originated, when the precise timing of downturn occurred, and how important were changes in sector composition such as farming. The trends can now be properly assessed with the data this paper develops.

A discussion in economics related to this 'U'-shaped 'slowdown' in British economic growth is sometimes referred to as a 'climacteric', as reviewed by Feinstein (1997: xi). One of the earliest contributions by Bowley (1920: 26) noted that growth of average money incomes was slow or negative after 1900. Lewis (1978: 128-34) argued that Britain was either unable or too slow to develop remedies because of intensified strength of competition from Germany and the USA, and Phelps Brown and Handfield-Jones (1952: 122-8) found that decline occurred in most manufacturing and agriculture, ascribing it to a slowdown of growth of industrial production compared to money incomes 'about the turn of the century'.

However, like the 'U'-shape, Feinstein (1997: xi-xii) noted it was unresolved how far the slowdown occurred at all, actually occurred from around 1900, or was sufficiently marked to merit a title such as 'climacteric'. McCloskey (1970) suggested that although there was some slowdown this was more an adjustment of growth rates to the resources available and the technological opportunities rather than stagnation. Feinstein's (1972) data on aggregate sector GDP suggested compositional change: decline in real rate of growth in industrial production from the 1870s offset by growth in services until about 1899, after which continued growth of the UK economy began to rely increasingly on offsetting a national slowdown by income for abroad. This indicates that the 1901 census should mark a turning point in the domestic economy towards a 'U'-shaped distribution for the rest of the twentieth century. However, the debate has been left largely unresolved since the 1980s because the available aggregate data was limited. The new data presented here allow insights into the extent and form of these changes that has not been previously possible.

Sector composition and participation by different types of people, the third theme assessed here, are important fields to which estimates of the full proprietor population estimates

contribute a range of new insights. The UK census is limited in the detail it offers on industry and business structures (it is unfortunate that the census administrators did not follow US counterparts who introduced questions on capital and motive power into their censuses from the 1830s). But the UK census is good at indicating sector and demographic composition. The analysis here uses 13 aggregate sectors.[1] Sectors are generally accepted as the main nexus of market competition and adjustment, and hence compositional change, as used by Marshall (1919), Clapham (1932), Kirzner (1973: 89-101) and Ekelund and Hébert (1983). They also fit into wider narratives regarding the shifting structure of modern industrial economies particularly between agriculture, industry and services (Kuznets 1966; Lee 1984; Shaw-Taylor and Wrigley 2014; Broadberry 2014; Tomlinson 2016). Sectors show the specific fields in which larger firms gained increasing advantages from economies of scale, mechanisation and managerial developments, squeezing out smaller proprietors, as argued by Chandler (1962, 1990) and Taylor (1911). But Chandlerian processes were far from uniform. For example, employment figures in agriculture fell but entrepreneur numbers did not. Having information on proprietor numbers gives new scope to assess these long-standing debates about productivity, occupational structure and firm organization. In addition, data on interactions with changing household structure, female labour force participation, and composition of proprietorship by sex can address important debates in contemporary business history about the extent of female entrepreneurship (Kay, 2009; Aston, 2016) and evolving social structures (Smith et al., 2020). Assessment of female participation has been bedevilled by poor recording in historical records. The UK census has its own limitations for assessing female participation (Higgs, 1987; Anderson, 1999), but these are radically reduced by using the *original* records rather than published census tables that clerically screened much women's work (Higgs and Wilkinson, 2016; van Lieshout et al., 2019; You, 2020). The census is generally as good as, or better than, other sources for identifying women as proprietors.

## 3. The population census as a source for business history

The original manuscripts of the UK population census that are now digitally available are the records processed centrally by the census administrators: the General Register Office (GRO).

[1] The detailed definition of the 13 sector categories is based on census occupation codes (Bennett et al., 2017), which also gives a cross-index to other classifications, such as HISCO, Feinstein, or SIC which can be applied to the BBCE data deposit if users prefer; downloads of definitions at https://www.bbce.uk/

The surviving records 1851-1901 are the Census Enumerators' Books (CEBs), and for 1911 from the original householder returns. These records were passed to The National Archives (TNA) where they were scanned and transcribed by various commercial genealogy suppliers. The transcripts used for most years derive from FindMyPast (FMP), and for 1881 from the Genealogical Society of Utah (GSU). These have been processed and coded to produce a census database deposit now available at UKDS: *The Integrated Census Microdata (I-CeM)*.[2] The data in this paper derive from I-CeM, but are enhanced by the authors in the BBCE deposit to identify and code proprietors, re-code occupations to main business activity, and infill major truncations in FMP transcriptions for 1851 and 1861 using the original CEB records and an additional genealogy database from S&N: The Genealogist.[3] No data for 1871 is yet in I-CeM, but this has been input into BBCE from S&N,[4] although this is only for employers so that 1871 data informs only part of the analysis below.

The later censuses (1891-1911) included an explicit question on 'employment status' which identifies *all* individuals as we require: as employers or own account proprietors, and also explicitly identifies workers. This question has continued into modern censuses in a similar format, although historically it had some flaws which require some re-weighting, as described elsewhere (Bennett et al., 2018a).

In the earlier censuses (1851-81) the question on employment status was more narrowly focused on employers. In 1861 the census asked:

> In TRADES, MANUFACTURES, or other Business, the Employer must, in all cases, be distinguished. – *Example: 'Carpenter – Master, employing 6 men and 2 boys;'* inserting always the number of persons of the trade in their employ, if any, on April 8th [the date of the census]

Similar instructions were used for other years 1851-81. However, the GRO published few table from the data for non-farmers; only a summary for 1851, selective tables for 1891-1911, and nothing for 1861-81. This deficiency limited Clapham (1932: 35) to very brief discussion of the published tables.

---

[2] Schürer and Higgs (2014), Higgs et al. (2015); an earlier e-version for 1881 is also used: Schürer and Woollard (2000).
[3] See Bennett et al. (2020). Note that about 3.7 per cent of 1861 records do not survive at TNA. The 1861 results have been up-weighted to maintain comparability of the aggregates.
[4] van Lieshout et al. (2018).

Census responses to these questions created long occupation descriptors. For example, Robert Shaw of Colne, Lancashire was returned in 1881 as 'Magistrate Cotton Spinner and Manufacturer employing about 1200 workpeople & farming 94 acres, employing 2 men'.[5] This string did not explicitly state 'employer', but that status is indicated by the stated employees. All strings have to be searched for this type of content for any employees, and statements of 'master', 'proprietor', 'owner', 'partner', etc.[6] These strings were searched algorithmically and, following extensive hand checking, this identified large numbers of proprietors. However, not all employers answered fully, and it was inevitable from the unsatisfactory question design that few own-account proprietors responded.

However, by using the original responses in the digital records and modern methods of post-response supplementation it is possible to use the characteristics of those proprietors who did respond in order to supplement the data on those who did not, as well as using insights form the much fuller later census responses. The supplementation method developed below follows modern census post-response editing (Lyberg and Kasprzyk, 1997: 355-8), which was first applied in the 1950s to correct biases in the US census (US Bureau of Census, 1950; Jabine and Tepping, 1973). In the UK this was first applied to the occupation question used in this paper in the 1961 census, and is now standard (ONS, 2019).[7] Similar methods are now routinely applied to a wide range of censuses internationally and to other surveys (see Rubin, 1987, 1996; Peress, 2010; Zhang et al., 2013; Weeks, et al., 2018). Post-survey editing has been little used for historic censuses on the scale required here; for example little attempt has yet been made to adjust and scale-up surviving sample responses from the US Census of Manufactures to give national estimates. Supplementation adjustments are particularly necessary for the format of the proprietor question used in British censuses 1851-81. Modern analyses show that this format has two of the highest prevalences for non-response: open reporting (Conrad et al., 2016: 77-80), and targeting the self-employed who find difficulty in accurate reporting for the smallest establishments with multi-attribute activity (Martin et al., 1994: Tables 1 and 2).

[5] TNA, RG11 4163/91, 1881 Manuscript Census Enumerators Book.
[6] See Bennett and Newton (2015) for full search terms. 'Director' identified many corporate proprietors. These were not isolated as a separate category in supplementation, but are included in estimates through their status as employers, own account, and some as workers. Detailed analysis of directors is given in other publications.
[7] In 1961 a post-enumeration survey gave the first estimates of census biases and used correction weights for some published occupational tables: see Census of Great Britain (1961: xiv-xx).

## 4. Supplementation methodology

Supplementation to identify non-responses aims to provide a continuous series of data that can track proprietor numbers on a consistent basis over the entire period 1851-1911. The methodology identifies response/non-response characteristics of proprietors compared to workers from the later censuses 1891-1911 (as reported in Bennett et al., 2018a) to supplement responses in 1851-81. Although entrepreneurial characteristics may have changed, the characteristics of those proprietors who responded had similar response probabilities over time; e.g. lawyers and many professions were less likely to respond fully to question in 1891-1911, and this was similar in 1851-81. More generally, the non-response categories in the earlier censuses closely mirror those of the later censuses.

This is due to the design characteristics of the main census questions, which changed little over time. The least likely to give full responses were those who were *not* head of the household, especially relatives of the family of the head, because they were not replying in the census directly - it was the head who replied for them. Also less likely to respond fully were individuals living in remoter rural areas, in the densest parts of major cities (locations that were more difficult to survey), and those in occupational sectors such as professions, mining and retailing where multiple census instructions for these categories throughout the period confused respondents (see Bennett et al., 2019a: 64-73). The result is that largely consistent flaws allow the full response patterns available in the later census (by demographic characteristics, sector, and location type) to be used to estimate probable proprietor or worker status in the earlier censuses. It should be emphasised that this type of approach is essential because there is no alternative 'truth' data for the whole population collected at this scale, although useful large-scale material is available from trade directories, which is used to test robustness below.

We use the supplementation approach summarised by Rubin (1987, 1988), that accurate covariates of proprietor/worker status are observed for all individuals in the census, which is itself a valid full population coverage not a survey, and the values we want to estimate for proprietor status occur in the occupational descriptors for all respondents and non-respondents. Our aim is a supplementation that allows users of our data to get the results from analysis of the incomplete data that would be valid if there were no missing responses, based on the same model posited for the missing data as for the complete data (Rubin, 1996: 476).

The method is a single imputation technique, but tested against alternatives using other assumptions, tests of residuals, and against the only other available large-scale source of directories. The supplementation follows five stages.[8]

First, the respondent categories 1851-81 were identified directly where possible from their occupation strings, such as Robert Shaw above. Several kinds of employers can be extracted, termed here 'extraction Groups'. These Groups were tagged to a status as either employer, own account, or worker as follows:

**Group 1:** tagged as employer if stated employees;

**Group 2:** farmers stating 'employer' but with no stated employees, tagged as own account; non-farm 'employers' with no employees identified by supplementation.

**Group 3:** anyone stating 'master' but giving no employees, tagged as own account.

**Group 4:** 'farmer' not stating employees or acres, tagged as worker (unless they had other business activities in a portfolio).

**Group 5:** 'farmer' stating acres but no employees, tagged as worker if they had under two acres (the normal size for a smallholder working on other farms).[9] If stating over two acres, tagged as employers if the acreage was large enough normally to require workers in that location, otherwise tagged as own account.[10]

**Group 6:** 'owners' or 'proprietors' of assets such as mines/quarries, ships/barges (but excluding land/house owner), tagged as employers or own account based on their strings where possible, otherwise by supplementation.

These Groups tag actual responses by proprietor status for *all* farmers and many other proprietors directly from their census descriptor strings without supplementation (just under 400,000, and about 40 per cent of the final estimated total). Farmers are fully tagged and do not need supplementation because an additional census question for farm acres gives sufficient direct information.

[8] A full account of the stages given in Bennett et al. (2018b; 2019b); extended to Scotland in Smith et al. (2021)
[9] See PP (1888), PP (1895).
[10] Group 5 farmers were assigned as employers where acreage was larger than the average cut-off normally requiring employees to work a farm taking account of the local farming conditions (such as land quality, climate, and degree of access) in that RSD; see Montebruno et al. (2019). A 'mid-aggregate' was used, following methods to separate employers and own account in farming where farm size is a truncated lognormal distribution: Allanson (1992), Lund and Price (1998).

The second stage of the method was to tag the remaining population where the responses were definitely workers (e.g. labourers, servants etc.) or non-economically active. This was implemented for a fine mesh of 844 census occupational categories, or sub-occodes.[11] This filtering directly established that 75 per cent of the economically active *correctly* did not respond to the census question as 'employers' or own account (6-8 million people); they were excluded from proprietor non-response supplementation.

The third stage developed a 'preferred method' of supplementation for the remaining non-respondents (Method 1). This was a logit regression estimated from the 1891 census response categories of proprietors and workers applied to the earlier censuses. This estimates the probability of individuals over 1851-81 being in similarly defined response categories using the 1891 logit coefficients. The logit used a wide range of variables to distinguish response categories that differentiate between proprietors and workers (see Appendix). Initially this was done for all proprietors (employers and own account together vs. workers). A similar logit was then applied to separate employer response characteristics from own account. Each logit estimate was developed separately for 844 occupational categories. This method is 'preferred' since it is the most direct probabilistic method to distinguish category responses. Further details of the estimates and comparison with alternatives are given in the Appendix.

The fourth supplementation stage was to test of robustness by comparing four other alternative supplementations for each of 844 sub-occodes. The preferred logit using 1891 response characteristics is Method 1. Method 2 was a similar logit regression but using 1901 data, with the same covariates. This tests sensitivity of the 1891 responses against 1901. Method 3 was non-statistical, using 1891 extrapolation ratios of respondents to non-respondents for employers, own account, and workers for 1851-81 in each 844 category. This assumes simple 'organic growth' which is plausible in some situations since many proprietors were actually the *same* people between census years, others were the *same business* run by new proprietors, and new entrants often tend to increase in proportion to increases in population. Method 4 was identical to Method 3 but used the 1901 respondent ratios. Method 5 was a 'tailored logit' that exploits the range of probabilities estimated by Method 1. This 'tailors' each respondent/non-respondent probability by selecting logit cut-offs to match the

[11] The 844 sub-occodes split the 797 census occodes give in I-CeM into sub-categories that reflect concentrations of proprietors/non-proprietors in 1891-1901 using their occupational strings; see example in Appendix Figure A.1. The I-CeM occodes are given in Higgs et al. (2015); the full list of filtered categories and sub-occodes is given in Bennett et al. (2018b).

probability ranges for each 844 occupation, whereas Method 1 averages the cut-offs. The method is summarised in the Appendix. It is expected to be less accurate for estimating total non-response numbers, but should identify specific individuals more accurately.

The fifth stage of supplementation compares the preferred Method 1 with Methods 2-4, and with contemporary sources and existing secondary literature, as summarised in the Appendix. This allows substitution between estimation methods as an 'intelligence-led' approach; i.e. suing secondary sources to guide choices. This is essential given lack of 'truth' data to train the estimator. It is referred to as *Supplement 1* below. Method 5, referred to as *Supplement 2* below, is retained as an independent alternative estimation comparator. This intelligence-led approach follows Rubin's (1988: 81) stricture that well-designed single imputation techniques are perfectly acceptable if checked against alternatives. To allow replication, all the decisions for each sub-occode are available as supplementary materials accompanying Bennett et al., 2018b, 2019b; the actual people identified by Supplementation 1 and 2 are given in the BBCE data deposit as EMPSTATUS_NUM and EMPSTATUS_IND.

***Internal checks***

The alternative methods are compared in Figure 1: separately for all proprietors, employers, and own account. This offers internal checks. External checks using other data sources are already built into the intelligence-led approach, but are supplemented by comparisons with trade directories later in the paper. The intelligence-led final estimate (shown as Supplement 1 in the figure) derives from the logit estimates for 1891 and 1901 (Methods 1 and 2). These are shown in the figure for rounded (R) and unrounded (UR) alternatives, depending on how decimal probabilities are counted; with the UR including all small decimal values. The figure also shows the ratio extrapolations (Methods 3 and 5). The tailored cut-off method is Supplement 2 (individual estimates). Additionally, the figure shows the number in 'extracted' Groups 1-6 directly tagged from census responses before supplementation. The Figure gives aggregate numbers, but the actual estimates are for each of the 844 sub-occodes.

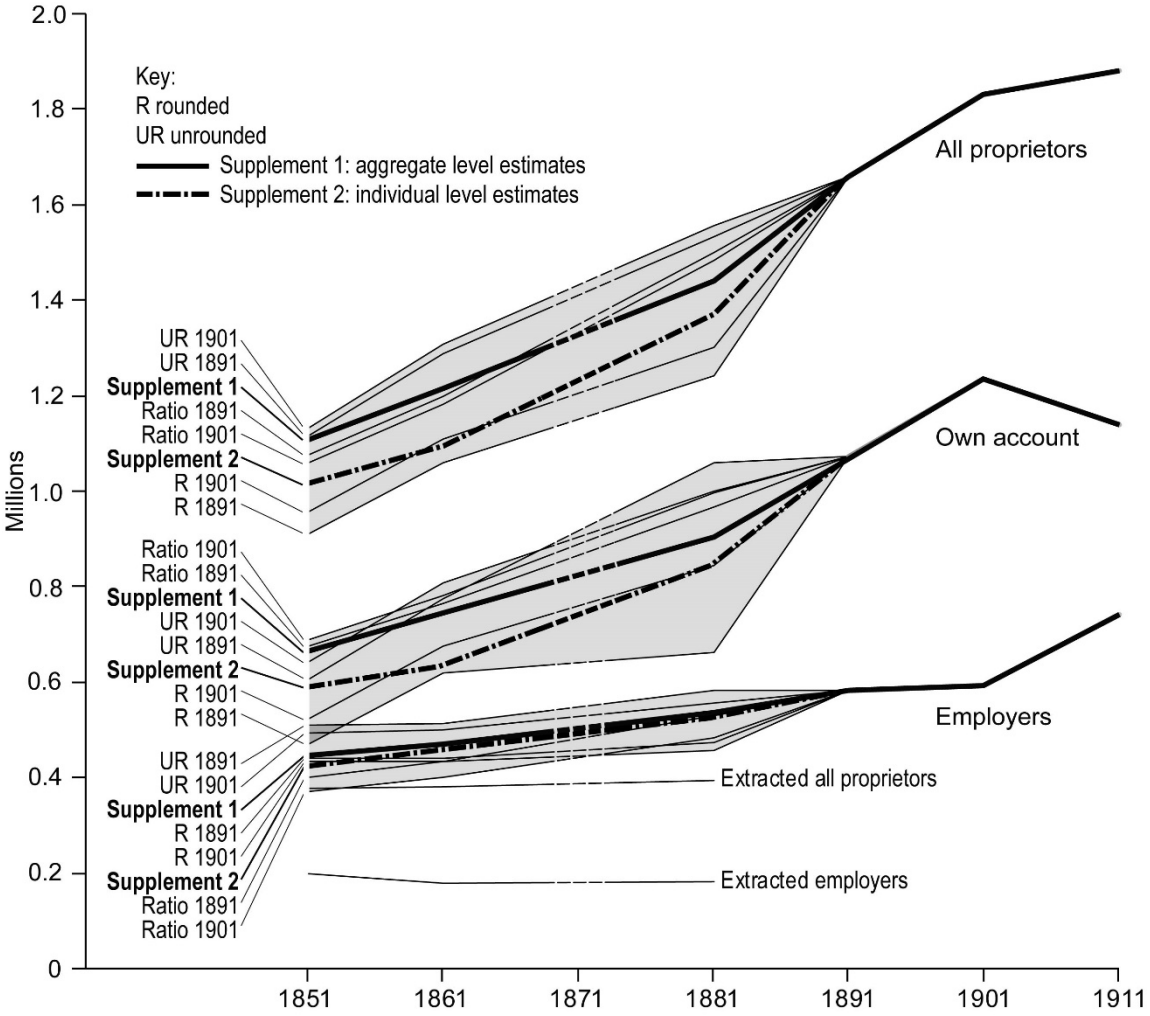


**Figure 1.** Comparison of supplementation estimates of proprietor numbers 1851-81, joined to 1891-1911 data, showing the estimation range for alternatives (R: rounded; UR: unrounded; and extrapolation ratios). The chosen supplementation methods in bold Supplement 1: intelligence-led; Supplement 2: tailored logit. The tagged census responses are shown at the foot for 'extracted' 1851-81 (full 1871 data unavailable).

Although the figure appears complex its key features stand out. First, and most important, the main trends are very similar whatever alternative estimate is used. Hence, although precise numbers will carry a level of uncertainty, which can be quite large in some of the 844 sub-occodes, in aggregate the trends for the period of a general increase and then sharp decline after 1901 have a fairly narrow range of possibilities; and this is the basis for the interpretations which we draw out later.

Second, reliance on the tagged census extractions alone results in substantial under-estimation of proprietor numbers. Responding employers were about one half of the fully supplemented numbers in 1851, and for all proprietors it was about a third. This under-estimation, without adjustment, would result in a huge leap in numbers between 1881 and 1891 which is implausible. Third, despite the limitations, the actual extracted census responses capture a slower rate of increase in own-account proprietors and slight decrease in employers 1851-61 which is largely confirmed by the supplementation estimates.

Fourth, the choice rules between methods used for the intelligence-led approach in Supplement 1 lead to a convergence in most cases towards the median for total numbers between the range of estimates, except in 1851 and 1861 for own account (and hence the total), where the estimates are towards the top of the range of alternatives. Fifth, the range between rounded estimates (R: lower bound) for the logit, and those with many low decimal probabilities (UR: upper bound), is wide for own account, but narrow for employers, as expected since the employer extraction Groups are more complete in the census. However, the choice of rounding is one of the most sensitive aspects: how far we accept very small indications of proprietor characteristics (UR) from the logit, as opposed to requiring a probability that exceeds 0.5 (R). This effect can be clearly seen for the grocers' example in Appendix Figure A.1; it strongly favours using rounded estimates which ignore small probabilities of proprietorship, except in maker-dealer categories. Finally, the trends evident for alternative comparator estimates based on tailored cut-offs (Supplement 2) are very similar to Supplement 1, providing a level of internal consistency check. They are 5-10 per cent lower overall than the intelligence-led aggregate estimates, almost entirely deriving from own account with low probabilities of proprietorship, reflecting the higher level of uncertainty for these respondents.

Taken together these features indicate that, while there are inevitable uncertainties in the supplementation, alternatives based on different plausible assumptions about developments over the period give similar general trends.[12] The main differences between methods are for

[12] Pilot experiments with other alternatives also confirm the supplementation estimates are comparable to results that can be achieved by various forms of machine learning (Montebruno et al., 2020). Perhaps the best improvements in future will lie with combining logit, machine learning and secondary evidence such as directories as piloted later in this paper.

own account, within which the range is most strongly affected by the supplementation chosen for maker-dealers.[13]

## 5. Overview of sector trends, rates, and gender

The primary purpose of this article is methodological, to establish estimates of the proprietor population for 1851-81 that can link with 1891-1911. Here brief commentary on the long-running debates about Victorian entrepreneurship extends and updates the discussion in Bennett et al (2019a: 94-8). The paper then turns to external robustness checks.

### 5.1 Aggregate numbers

The new estimates of proprietor numbers 1851-81 are given in Table 1 together with the 95 per cent bounds of the estimates, expanding the internal checks outlined in the Appendix and above. The supplementation gives 'point' estimates of the total-count for each sex in each sector; the bounds give the confidence interval within which the point estimate in each sector should fall.[14] An important point is that the confidence bounds are very narrow, which reflects that many census responses were accurate (the tagged extracted), most supplementation is applied to individuals who are well-defined by their response characteristics resulting in a narrow range of possibilities (as indicate for the 872 sub-occode 'grocers' in Figure A.1), and the data size is large, which results from the narrow standard errors with high z-values for the logit (see Appendix Table A.1).

[13] Further checks and tests in Bennett et al., (2018b, 2019b). The BBCE database deposit gives the alternatives so that other researchers can opt between them.

[14] The confidence intervals are calculated by

$$\hat{\theta} \pm z_{1-\alpha/2,n-1}\{\hat{V}(\hat{\theta})\}$$

where $\hat{\theta}$ is the total-count estimator and $\hat{V}(\hat{\theta})$ is the estimator of the variance of the total-count estimator, with, $z_{1-\alpha/2}$ a two-sided z statistic corresponding to a significance level of $\alpha$=0.05 with n − 1 degrees of freedom.

| MALE | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|
| **Sector** | **1851** | | | **1861** | | | **1881** | | |
| | **Point** | **lb** | **ub** | **Point** | **lb** | **ub** | **Point** | **lb** | **ub** |
| Farming/estate work | 208,481 | 207,715 | 209,247 | 205,920 | 205,151 | 206,689 | 214,675 | 213,866 | 215,484 |
| Mining and quarrying | 5,566 | 5,420 | 5,712 | 7,174 | 7,009 | 7,339 | 7,337 | 7,170 | 7,504 |
| Construction | 69,054 | 68,562 | 69,546 | 69,747 | 69,252 | 70,242 | 95,669 | 95,091 | 96,247 |
| Manufacturing | 72,196 | 71,694 | 72,698 | 75,015 | 74,504 | 75,526 | 99,202 | 98,615 | 99,789 |
| Maker-dealer | 144,891 | 144,218 | 145,564 | 133,231 | 132,577 | 133,885 | 157,401 | 156,685 | 158,117 |
| Retail | 34,885 | 34,527 | 35,243 | 41,317 | 40,929 | 41,705 | 59,250 | 58,787 | 59,713 |
| Transport | 19,282 | 19,013 | 19,551 | 25,763 | 25,453 | 26,073 | 33,434 | 33,081 | 33,787 |
| Professional services | 20,349 | 20,073 | 20,625 | 24,196 | 23,896 | 24,496 | 39,583 | 39,201 | 39,965 |
| Personal services | 28,590 | 28,265 | 28,915 | 30,171 | 29,837 | 30,505 | 53,485 | 53,044 | 53,926 |
| Agricultural produce processing & dealing | 16,424 | 16,175 | 16,673 | 16,261 | 16,014 | 16,508 | 21,573 | 21,288 | 21,858 |
| Food sales | 95,466 | 94,899 | 96,033 | 110,995 | 110,388 | 111,602 | 158,323 | 157,605 | 159,041 |
| Refreshment | 48,533 | 48,115 | 48,951 | 56,833 | 56,382 | 57,284 | 71,299 | 70,794 | 71,804 |
| Finance and commerce | 13,805 | 13,577 | 14,033 | 17,033 | 16,780 | 17,286 | 24,272 | 23,970 | 24,574 |
| **Total** | **777,522** | | | **813,656** | | | **1,035,503** | | |

| FEMALE | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|
| **Sector** | **1851** | | | **1861** | | | **1881** | | |
| | **Point** | **lb** | **ub** | **Point** | **lb** | **ub** | **Point** | **lb** | **ub** |
| Farming/estate work | 21,158 | 20,883 | 21,433 | 20,370 | 20,099 | 20,641 | 19,955 | 19,684 | 20,226 |
| Mining and quarrying | 689 | 638 | 740 | 796 | 741 | 851 | 610 | 562 | 658 |
| Construction | 1,719 | 1,638 | 1,800 | 2,123 | 2,033 | 2,213 | 1,793 | 1,710 | 1,876 |
| Manufacturing | 27,646 | 27,335 | 27,957 | 28,326 | 28,010 | 28,642 | 27,481 | 27,166 | 27,796 |
| Maker-dealer | 132,458 | 131,914 | 133,002 | 145,880 | 145,309 | 146,451 | 189,068 | 188,417 | 189,719 |

| | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|
| Retail | 15,026 | 14,792 | 15,260 | 14,744 | 14,511 | 14,977 | 22,203 | 21,918 | 22,488 |
| Transport | 1,218 | 1,150 | 1,286 | 1,324 | 1,253 | 1,395 | 1,160 | 1,093 | 1,227 |
| Professional services | 1,417 | 1,343 | 1,491 | 1,609 | 1,531 | 1,687 | 2,393 | 2,297 | 2,489 |
| Personal services | 55,305 | 54,886 | 55,724 | 65,644 | 65,191 | 66,097 | 89,716 | 89,190 | 90,242 |
| Agricultural produce processing & dealing | 2,455 | 2,358 | 2,552 | 1,971 | 1,884 | 2,058 | 1,997 | 1,910 | 2,084 |
| Food sales | 32,609 | 32,274 | 32,944 | 36,787 | 36,431 | 37,143 | 51,816 | 51,396 | 52,236 |
| Refreshment | 24,652 | 24,356 | 24,948 | 28,263 | 27,947 | 28,579 | 43,168 | 42,781 | 43,555 |
| Finance and commerce | 985 | 924 | 1,046 | 1,161 | 1,094 | 1,228 | 2,402 | 2,306 | 2,498 |
| **Total** | **317,337** | | | **348,998** | | | **453,762** | | |

**Table 1.** Total proprietor numbers (000s) 1851-1881 for 13 sectors. The point estimates are the tagged extracted plus supplemented. Confidence interval given between the lb (lower bound) and ub (upper bound); full 1871 data not available.

The estimates in Table 1 are a key output from this paper. They allow the early census estimates of proprietor numbers to be joined up with the later years (as given in Bennett et al., 2018a). The mid-point estimates are shown for the aggregate as Supplement 1 in Figure 1. Over the whole of 1851-1911 total entrepreneur numbers increased steadily, although expansion slowed 1901-11. But within this, employers increased more slowly throughout the period, and most slowly compared to all proprietors between 1861 and 1901, indicating that increasing own-account numbers mainly drove trends of *total* numbers. After 1901, however, the trends show a major reversal: own account fell absolutely and employer numbers began to increase more rapidly but were insufficient to prevent the growth rate for all proprietors markedly slowing down. It is this point which indicates a turning point and the beginning of any 'U'-shape: before 1901 there is no indication of a 'decline' in aggregate entrepreneurship, but after 1901 there was a re-balancing from own-account proprietors towards those employing others.

This downturn was already evident in the estimates for 1891-1911 given by Bennett et al. (2018a). The new estimates show that this downturn was a major reversal of long-established growth rather than a short-term response. Further confirmation of the significance of the 1901-11 downturn is provided by comparison with the 1921 published census data. Own-account numbers, which declined by 8 per cent 1901-11, declined by a further 12 per cent 1911-21; whilst total proprietor numbers increased by a modest 3 per cent 1901-11, they declined by 13 per cent 1911-21.[15] A turning point of proprietor numbers is therefore an early twentieth-century phenomenon not a Victorian decline. This timing fits with the economic arguments about a 'climacteric' about 1900, and confirms that it was indeed very significant, though the mainly own-account composition of this change is very different from previous interpretations that mainly focus on agriculture and manufacturing. It demonstrates the declinist narrative, when it addressed overall numbers, placed changes too early: it was clearly a post-1900 phenomenon, rather than occurring in the nineteenth century.

The sector estimates for 1851-81 given in Table 1, joined up with the later censuses, are shown in Figure 2. This allows the key compositional adjustments in sector markets to be examined. Within the steady increase of numbers up to 1901, maker-dealers were the largest group, the largest contributors to growth, and the largest contributors to dynamics with the

[15] Comparison with published 1921 census tables, corrected for seasonal workers affected by 1921 June census date (see Bennett 2019a: Table 4.2).

largest decline in numbers after 1901. Their turning point in 1901 suggests changes in competitive conditions as maker-dealers experienced increasing competition from specialist manufacturing firms that were gaining advantages from economies of scale, as argued by Chandler (1962, 1990) and Taylor (1911). The retail aspect of maker-dealing was also affected by competition from specialist retailers (which Figure 2 shows had the most rapid growth), and multiple stores with branches across the country that displaced independent maker-dealer businesses, notably the expansion of firms like W. H. Smith, International Stores, and shoe chains (Jefferys, 1954; Alexander et al., 2003; Scott, 1994). Similar downturns also affected food sales, and personal services, though less marked than for maker-dealers.

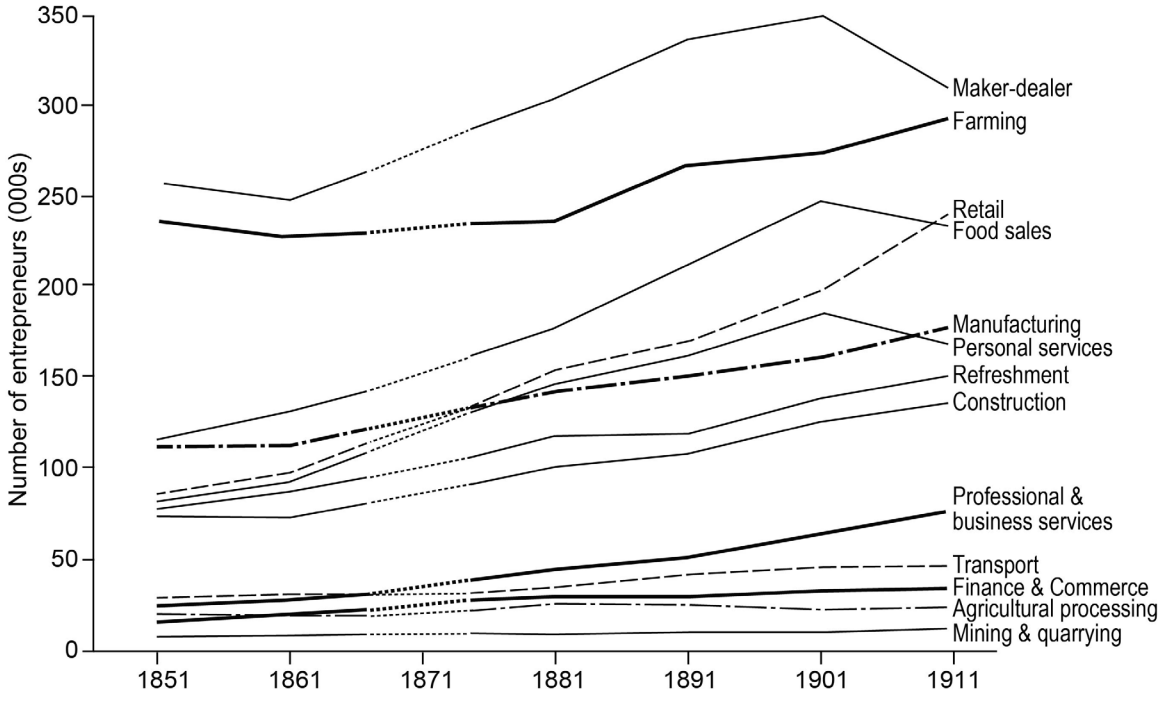


**Figure 2.** Total proprietor numbers (000s), 1851-1911, for 13 sectors, using supplement method 1 for 1851-81.

These changes throw doubt on the compositional inferences drawn by Wennekers et al. (2010) or Clark (1957). Farm numbers did not reduce, even though farm labour steadily declined. Increases in farmer numbers began in 1861-71 and the new estimates indicate that numbers continued to increase. Also manufacturing numbers, though generally growing, declined slightly 1851-61, and increased relatively slowly up to 1901. Any switch between

farming and manufactures did not account for the greatest numerical changes in proprietorship. Looking at proprietor numbers gives an entirely different insight from the trends in occupation numbers that have dominated the previous literature. The largest growth in business numbers until 1901 was in retail and food sales, after which it was construction, professional and business services, and personal services. Hence, manufacturing was far from being the most important contributor to entrepreneurial change, either in numbers or growth rate. The suggestions that manufacturing consolidation and managerial changes of industrial entrepreneurship mainly drove business evolution (e.g. Chandler, 1990) were certainly important, but engaged in only one aspect of a much more varied process.

### 5.2 Entrepreneurship rates

It is essential to compare proprietor numbers against the rest of the working population which was growing steadily over the period. Rates here are compared to the total of all economically active, as shown in Table 2. The growth of manufacturing which has dominated the literature absorbed most waged labour in larger factories, with utilities, coal mines and transport undertakings also major employers in large enterprises. Proprietor numbers did not keep pace, resulting in the general rate of entrepreneurship, shown in the last row of the Table, declining slowly until a steeper drop after 1901. This gives weak evidence of a decline in entrepreneurship rates over the Victorian period, but only because numbers slightly lagged behind population growth. As Marshall (1919: 92-3) observed, the nineteenth century had 'an unprecedented combination of advantages [that] enabled businessmen to make money even when they were not throwing themselves with energy', which explains why later declines might more easily occur. But the new evidence of proprietorship rates does not support significant decline in proprietorship until after 1901. This further confirms the interpretation of a climacteric, with 1901-11 showing the sharpest changes of the whole period.

| | **1851** | **1861** | **1871** | **1881** | **1891** | **1901** | **1911** |
|---|---|---|---|---|---|---|---|
| **EMPLOYERS** | | | | | | | |
| Farming/estate work | 9.4 | 10.1 | 10.3 | 10.0 | 11.9 | 12.7 | 12.9 |
| Mining & quarrying | 1.5 | 1.3 | | 0.8 | 0.7 | 0.7 | 0.7 |
| Construction | 8.4 | 8.1 | | 7.7 | 7.8 | 6.0 | 7.0 |
| Manufacturing | 3.9 | 3.3 | | 3.2 | 3.0 | 2.6 | 2.7 |

| | | | | | | | |
|---|---|---|---|---|---|---|---|
| Maker-dealer | 6.9 | 5.9 | | 6.4 | 6.0 | 5.5 | 6.8 |
| Retail | 14.2 | 14.0 | | 11.9 | 11.4 | 10.4 | 10.4 |
| Transport | 2.7 | 2.1 | | 1.7 | 1.6 | 1.3 | 1.5 |
| Prof services | 10.5 | 9.1 | | 6.5 | 5.3 | 4.7 | 4.3 |
| Personal services | 3.7 | 3.5 | | 3.7 | 3.7 | 3.1 | 3.8 |
| Agricultural produce | 17.3 | 15.1 | | 14.4 | 14.0 | 10.8 | 11.4 |
| Food sales | 17.4 | 16.4 | | 14.6 | 12.2 | 10.4 | 12.7 |
| Refreshment | 16.4 | 14.9 | | 10.0 | 9.5 | 7.3 | 13.3 |
| Finance & commerce | 18.9 | 16.9 | | 11.6 | 9.5 | 8.1 | 6.6 |
| **All employers** | **5.9** | **5.5** | | **4.9** | **4.9** | **4.3** | **4.7** |
| **OWN ACCOUNT** | | | | | | | |
| Farming/estate work | 5.4 | 5.0 | 5.3 | 6.9 | 8.5 | 10.4 | 9.2 |
| Mining & quarrying | 1.0 | 0.8 | | 0.5 | 0.5 | 0.5 | 0.5 |
| Construction | 8.9 | 6.5 | | 5.5 | 6.6 | 5.5 | 6.2 |
| Manufacturing | 4.1 | 3.3 | | 3.0 | 2.8 | 2.6 | 2.2 |
| Maker-dealer | 25.6 | 23.0 | | 24.4 | 24.6 | 25.2 | 21.9 |
| Retail | 35.2 | 33.4 | | 31.2 | 28.5 | 26.0 | 23.0 |
| Transport | 4.4 | 3.9 | | 3.5 | 3.3 | 2.8 | 2.1 |
| Prof services | 10.0 | 9.0 | | 7.0 | 6.1 | 6.2 | 4.3 |
| Personal services | 20.4 | 20.0 | | 21.2 | 20.2 | 20.4 | 15.5 |
| Agricultural produce | 8.1 | 7.6 | | 8.8 | 8.3 | 8.7 | 8.8 |
| Food sales | 33.6 | 34.0 | | 30.4 | 29.2 | 30.1 | 23.6 |
| Refreshment | 47.8 | 48.2 | | 41.7 | 34.3 | 34.7 | 24.2 |
| Finance & commerce | 14.4 | 12.6 | | 11.4 | 9.5 | 10.1 | 5.8 |
| **All own account** | **9.3** | **8.6** | | **8.7** | **8.9** | **8.9** | **7.2** |
| **All proprietors** | **15.2** | **14.1** | | **13.5** | **13.8** | **13.2** | **11.8** |

**Table 2.** Entrepreneurship rates: total numbers of employers, own account, and all proprietors as percentage of all economically active, 1851-1911.

Sector rates are divided into the separate elements for employers in the upper half of Table 2, and own account in the lower half. There are some constraints on the interpretation for sectors. For 1871 complete sector data are not available, except for farmers, as included in the table. Also, estimates of sector rates are over-estimated in some cases because some of the

economically active cannot be accurately assigned to sectors (mainly general labourers and clerks): this mainly affects manufacturing, coal mining, some construction, and some finance. Nevertheless the main sector effects remain clear. Differences between sectors indicate the radical contrasts in firm size, with higher rates indicating high proportions of generally smaller businesses (as in retail), and lower rates indicating larger business sizes.

The most important feature of tends in *employer* rates was similarity across many sectors, generally echoing the slow change in the aggregate. Farming was the main exception to the rather stable and then slow decline up to 1901, confirming again that changes in entrepreneurship were not driven by changes in farming, as suggested by Clark (1957), Kuznets (1966), and Wennekers et al. (2010), although shifts in workforce numbers saw a major occupational shift out of farming driven by technological changes and rising demands for labour elsewhere. The other main contrasts for employers from the pattern of stability and then slow decline up to 1901 were the steep decline of finance and commerce where larger firms were making important inroads.

For *own-account* proprietors there were much more marked changes. The general own-account rate, after a decline 1851-61, rose slowly 1861-1901, before falling steeply 1901-1911 to give the fastest rate of change for the period (penultimate row of Table 2). The turning point of 1901 in the general rate of proprietorship (bottom row) was thus chiefly a shift in own-account proprietors relative to employers. This is confirmed by comparison with the 1921 census. Own-account rates led the changes that occurred in the twentieth century and contributed the main part of the ensuing down-turning limb of the 'U'-shape: they declined from 9 to 7 percent 1901-11 and continued to decline to 6 percent in 1921. In comparison total entrepreneurship percentage rates (bottom row) peaked in 1891, reduced slightly in 1901, but fell steeply from 12.3 to 11.8 over 1901-11, and reached 9.6 in 1921.

Trends in own-account rates had only minor sector contrasts, with refreshments and retail showing the strongest declines. This echoes the changes in numbers and ifs one of the largest contrasts with employer rates. It shows one of the largest organisational changes of the period to be the displacement of the single person shop or refreshment hostelry by larger establishments employing assistants, and by national chains with many branches of shops and pubs, in many cases as corporate businesses. For many sectors, however, own-account rates remained stable up to 1901, after which a large reductions in rates were widespread, though

the main contributors to this downturn were maker-dealers, retail, food sales, personal services, and refreshments.

### 5.3 Gender

Sectors interact with other compositional effect such as sex. A key finding from the new data is that there are much higher numbers of female proprietors than found in previous studies (e.g. Kay, 2009; Aston and Di Martino, 2017), despite limitations of the census data. It is also clear that a significant aspect of aggregate and sector changes derived from changing female participation. As already shown in Table 1, women made up over 20 per cent of all proprietors in six sectors: manufacturing, maker-dealers, retail, personal services, food sales, and refreshments. Table 1 also shows that despite higher uncertainties for female census recording, the bounds of the supplementation estimates in these categories were fairly narrow.

Table 3 shows the rate of change in number of male and female proprietors abstracted from Table 1 and joined up with 1891-1911. Changes are shown for two periods: the nineteenth century 1851-1901, and then the period of downturn 1901-11. The rate of growth of female proprietorship over 1851-1901 was higher for women than men: 62 per cent compared to 54 per cent (bottom row). However, over 1901-11 the pattern was dramatically different: male proprietorship continued to increase by 6 per cent, but women's proprietorship decreased by 11 per cent. Thus the decline of female proprietorship was a dominant part of the downturn after 1901.

| Sector | Male | | Female | |
|---|---|---|---|---|
| | 1851-01 | 1901-11 | 1851-01 | 1901-11 |
| Farming/estate work | 12.4 | 6.5 | -8.6 | 0.2 |
| Mining & quarrying | 44.9 | 24.2 | -57.3 | 102.6 |
| Construction | 65.0 | 8.5 | -23.5 | 30.9 |
| Manufacturing | 51.5 | 5.8 | -37.6 | 29.5 |
| Maker-dealer | 6.4 | -5.8 | 64.3 | -19.8 |

| | | | | |
|---|---|---|---|---|
| Retail | 131.4 | 16.7 | 83.9 | 16.2 |
| Transport | 103.1 | -3.8 | -9.2 | 8.3 |
| Prof & bus services | 117.1 | 21.8 | -21.7 | 150.6 |
| Personal services | 156.6 | 13.1 | 85.4 | -23.7 |
| Agricultural produce | 7.8 | -0.1 | -22.8 | 19.2 |
| Food sales | 102.3 | 2.0 | 147.9 | -22.2 |
| Refreshment | 54.9 | -9.3 | 102.6 | 10.7 |
| Finance & commerce | 128.0 | -10.1 | -8.7 | 21.8 |
| Total | 54.1 | 5.7 | 62.5 | -11.2 |

**Table 3.** Percentage change in number of male and female proprietors by sector 1851-1911.

The sector composition of this change over 1851-1901 was also remarkable. Whilst the concentration of women as proprietors was highest in six sectors, the change of these sectors was substantial. Five of the six main female sectors experienced growth of over 60 percent: maker-dealers, retail, personal services, food sales, and refreshments. This was almost all own-account growth since the employer proportion in these sectors was relatively small and mostly stable. In contrast, over the same period, only one male sector growth rate exceeded 36 per cent (which was for retail) and only four other sectors exceeded ten percent (farming, construction, manufacturing, and professional services). Hence, women were a key driver of changes in Victorian proprietor growth 1851-1901. Not only were most sectors of female entrepreneurship growing faster, but apart from retail, these were primarily in entirely different sectors from male-dominated industries. There were eight sectors with reductions in female participation 1851-1901, but apart from manufacturing, women were a small minority in all so that these reductions had limited effects on reducing the overall rate of female proprietorship.

Sector composition by sex was also important in the 1901-11 downturn. The three main female sectors (maker-dealers, personal services, and food sales) lost 80,200 proprietors over 1901-11, whilst the female aggregate lost only 61,400. The maker-dealers alone accounted for one half of all these losses. In comparison, male proprietorship was slower changing and more stable in the downturn over 1901-11, with maker-dealers a major part of losses but one

third that of women. Hence, the aggregate dynamics for the whole period were strongly affected by the growth of female participation, then its marked downturn after 1901.

However, despite the strong sector contrasts, comparison of the pervasiveness of the downturn over 1901-11 in Figure 3 shows that aggregate changes by sex were very similar. The rate of entrepreneurship was lower at younger ages for women, but higher for older. This remained similar between 1891 and 1901, but for 1911 there was a systematic drop in rates across all ages for both sexes, except for the very youngest men (under 22). But the female decline 1901-11 was roughly twice that for men across all age groups.[16] The lack of cohort effects over the 1901-11 downturn indicates systemic changes not distinctive exits by older or other age groups.

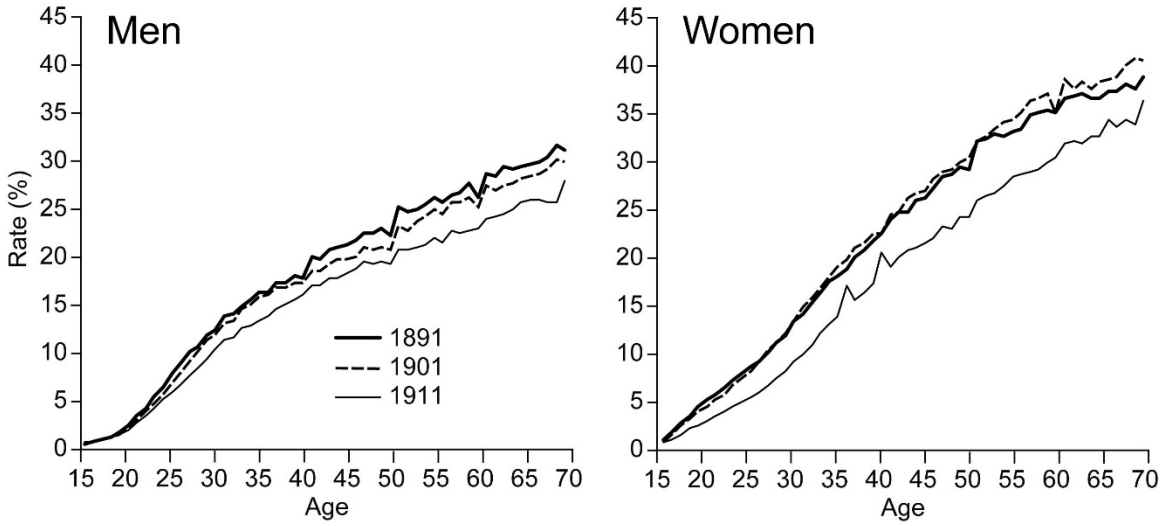


**Figure 3.** Male and female entrepreneurship rates 1891-1911 by age, as a percentage of economically active by sex.

## 6. Comparison with Trade Directories

It is important to use external checks to confirm the methodology that the new estimates are comparable to those in other sources. However, no other source has a consistent definition and collection regime for business numbers at the national scale required over the period

[16] Figure 3 also confirms that any effect of the mobilisation of males into the military that occurred prior to 1914, or changes in the incidence of retirement, was negligible over this period. Analysis of age and family structures for women is taken further in (van Lieshout et al., 2019).

examined; as noted at the outset, there is no national 'truth' data. Most alternatives, such as Factory Returns and Board of Trade reports cover only a fraction of proprietors, are sectorally imbalanced, ignore the smallest firms, and have inconsistent coverage over time. Directories, however, come closest to full coverage and their value has been demonstrated in business history for local studies. Two sets of comparisons are made here: national and local.

National trade directories exist in a little-used series compiled by Kelly which aggregated their local directories for the whole country from 1869. These act as a good test of the aggregate census numbers since they cover all sectors, range over the important period of 1871 to 1911, and bridge the difficult divide between the later censuses, and early censuses (1851-81) where supplementation has to be used. Hence, they provide a uniquely valuable resource for robustness checks. The Kelly directories were initially (1869) compiled for seven broad sectors; three more began by 1877, with the last added for textiles in 1880.[17] The total entries in England and Wales rise from over 900,000 in 1881, to over 1.3 million in 1911. This is about 70 per cent of the total proprietor numbers estimated here. Their entries were checked by the compilers against alternative sources, of which the published census, professional directories, Board of Trade, Inland Revenue, Factory Returns, and specialist sources are all quoted in the prefaces. Inconsistencies of coverage mainly result from expanding over time the range of sectors included, which can be overcome by aligning specific sub-sectors that were consistently covered.

The national directories are used here in two ways. First, to compare the total counts across all directory entries; and second to compare a sample of sub-sectors that can be aligned with census coding.[18] The comparison of total counts is shown in Figure 4 for nine of the directories, and shippers contained in the directory for *Merchants, Manufacturers and Shippers*.[19] Dates are shown for the best fits available to census dates leading to some irregularity of comparison points.

[17] See references for the directories under *The Post Office* (for nine directories), and *Kelly* (for two directories); the titles and coverage of the directories are eclectic with varied sub-sectors.
[18] This is approximate: precise comparison of counts is impossible because directories generally list businesses, while the census gives proprietors, partners and directors, leading to generally larger counts than directories.
[19] The *Post Office Merchants, Manufacturers and Shippers* directory contains many foreign agents and brokers which were often only a UK agency. Only shippers and merchants are consistent and well-defined entries in this directory comparable with the census.

The figure confirms four important features of the general trends evident from census estimates. First, the steady increase in business numbers up to the early 1900s is similar to the census in most of the directory categories, and especially its steepness of growth in some sectors over 1870-80 (notably builders and engineers, which both experienced radical change in business organisation over the period: builders from the emergence of large scale developer-builders, and engineers through its rapid professionalization and increased demand for their skills).

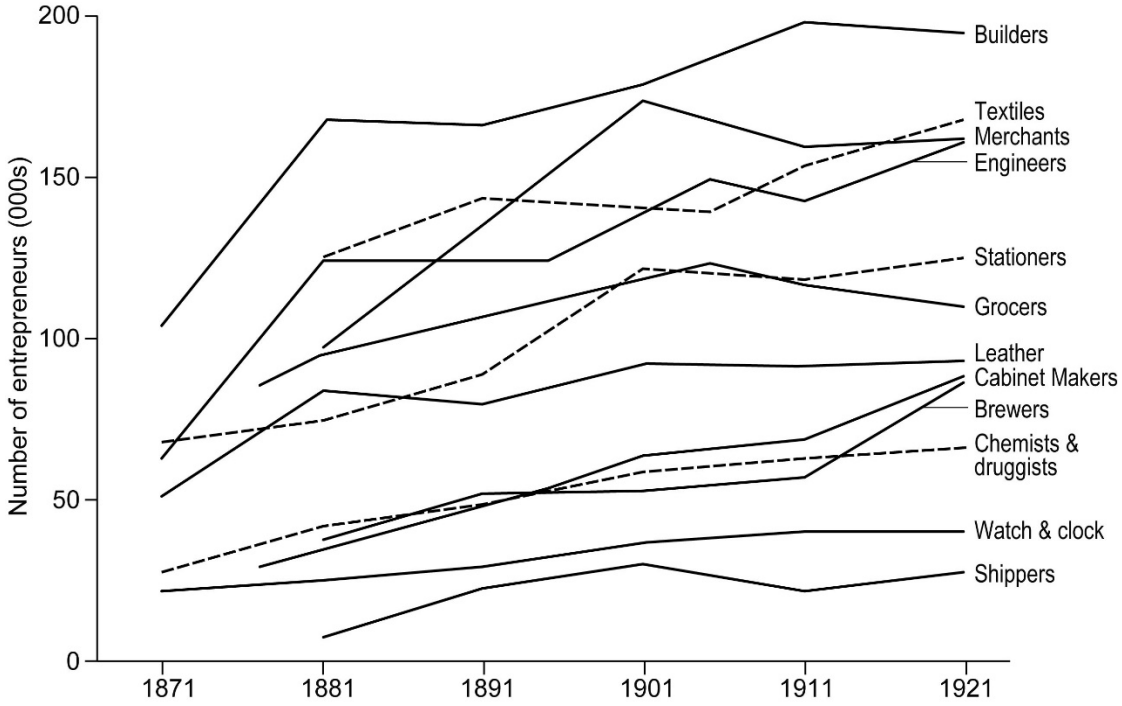


**Figure 4.** Trends of counts of all businesses in England and Wales listed in ten *Post Office* and *Kelly* sector directories for 1869-1922.

Second, the directories confirm the declines after 1900: for 1901-11 for stationers (which includes printers, newsagents, publishers and retailers) and shippers, and for different edition dates for 1905-11 for engineers, and grocers (which included many manufacturers of colour, oil, dyes, etc., and dry salters). Third, the effect of a slowdown due to the trade depression in the 1890s, which is partially evident in the census (compare Figures 1 and 2), is more clearly seen, especially for builders and engineers. Indeed, this slowdown is noted in several directory prefaces. Fourth, the directory comparisons carried forward to 1921 confirm the

emergence of the declining limb of the twentieth century 'U'-shape, mainly for builders, and grocers; as well as levelling of numbers for chemists, and watch & clock makers.

Unfortunately, these tends cannot be reliably converted to rates because precise matching to the rest of the economically active is impossible given the eclectic directory coverage. However, it is clear that in all sectors the upturn in numbers was too low to counteract the rapid expansion of the economically active. Hence, the levelling of directory numbers after 1901-11 confirms the climacteric downturn and initiation of the twentieth-century long decline in entrepreneurship rates.

A second comparison is from a sample of 23 directory sub-sectors chosen because they approximately align to specific census occodes and represent contrasting sectors. This comparison focuses on employers rather than own account since directories cover the latter less fully. The comparisons quoted in Table 4 focus on checking three aspects: first, that the early census data trends are comparable to the directory counts (these are shown in the first set of columns comparing changes 1871-81);[20] second, that the generally increasing trend carries across the divide between the two sets of census questions (shown for the second set of columns comparing 1881 and 1901); and third, how the downturn over 1901-11 is reflected in 1911 directories in different sectors.

| **Sector (I-CeM Occode)** | **Change 1871-81 (%)** | | **Change 1881-1901 (%)** | | **Change 1901-11 (%)** | |
|---|---|---|---|---|---|---|
| | **Census** | **Directory** | **Census** | **Directory** | **Census** | **Directory** |
| 43. Dentists | 45.0 | 40.2 | 118.0 | 58.5 | 96.6 | 4.7 |
| 44. Veterinary surgeons | 25.5 | 8.3 | 33.1 | 26.7 | -2.4 | -12.7 |
| 65. Civil engineers | 76.6 | 97.3 | -1.0 | -26.6 | -31.4 | -6.5 |
| 75. Photographers | 84 | 33.0 | 101.4 | 58.5 | 28.7 | 17.9 |
| 220-3 Quarry owners | 69.4 | 264.6 | -26.3 | -7.0 | 24.7 | -19.0 |
| 244. Steel manufactures | 21.2 | -7.5 | 145.5 | 11.5 | 7.2 | 4.9 |
| 258. Millwrights | 9.4 | 38.6 | -22.5 | -39.2 | -0.4 | -36.4 |
| 259. Iron founders | -27.2 | 38.7 | 168.4 | -21.6 | -0.17 | -25.3 |
| 260. Brass founders | 58.7 | 60.6 | 26.8 | -8.7 | -11.1 | -24.2 |

[20] The data deposit for 1871 infills I-CeM from S&N. The infill may be imperfect and underestimate employer numbers in 1871, hence exaggerating some growth estimates 1871-81.

| | | | | | | |
|---|---|---|---|---|---|---|
| 305. Nail manufactures | 2.0 | 35.6 | -56.3 | -12.0 | -24.2 | -55.6 |
| 369. Ironmongers | 70.3 | 83.9 | 45.9 | 22.7 | -0.9 | -5.5 |
| 388. Photographic apparatus makers | 105.0 | 27.1 | 159.4 | 213.3 | 11.4 | -31.4 |
| 405. Builders | 108.5 | 137.7 | 18.7 | 7.6 | 9.8 | 3.5 |
| 437. Cabinet Makers | 12.7 | 45.7 | 107.2 | 8.5 | -4.0 | -4.7 |
| 440. Upholsterers | 4.8 | 45.1 | 13.2 | 19.4 | -4.7 | -9.8 |
| 462. Timber merchants & wood dealers | 65.2 | 79.6 | 65.8 | 7.4 | -5.2 | -9.5 |
| 480. Manufacturing Chemists | -35.9 | 105.5 | 63.0 | -7.6 | -0.8 | -16.6 |
| 482. Chemists & druggists | 168.8 | 1.8 | 26.8 | 7.1 | 16.2 | 9.0 |
| 506. Tanners | 78.2 | 21.2 | -38.8 | -13.7 | -21.7 | -27.3 |
| 507. Curriers | 119.6 | 65.9 | -49.0 | -45.7 | -6.3 | -19.5 |
| 526. Cardboard box makers | 562.5 | 193 | 164.7 | 66.5 | 57.0 | 12.1 |
| 699. Ginger beer & mineral water manufacturers | 405.4 | 58.8 | -13.0 | 60.3 | -3.5 | -5.3 |
| 722. Wine and spirit merchants | 168.0 | n/a | 289.0 | 0.3 | -20.8 | -0.8 |
| Mean Change | 92.3 | 67.4 | 58.2 | 16.7 | 4.9 | -11.2 |
| Number of sectors (and %) with same trend | 19 (86%) | | 19 (83%) | | 21 (91%) | |

**Table 4.** Comparison of percentage change in counts for census *employers*, and directories, 1871-1911 for 23 sample sectors identified by I-CeM Occode number. Sources: 1871-81 comparison uses extracted tagged employers only; 1881-1901 uses the supplemented data for 1881.

Table 4 generally confirms the census analysis: the general upward trend 1871-81 evident from the census in most sectors, and the start of a downturn in business numbers in the early 1900s. The directories also show that the divide between the early and later census questions is satisfactorily bridged by the supplemented data. Over 80-90 per cent of the sample sectors show the same direction of trend between the census and directories for each period, although there are certainly differences in magnitude: the census generally records larger changes (mainly as a result of directories' poorer coverage of rural areas). The directories are also valuable in indicating that some of declines may have set in earlier than the census indicates,

suggesting that the 1890s trade depression may have affected some retailing, wholesaling and manufacturers more severely, as well as some professions. Hence, while complex, the directory comparisons generally confirm the same trends as the census estimates.

In addition to national directories, local directories can be used to give more precise external checks of supplementation by matching actual individuals. This is piloted here for two contrasted locations in 1881 which both had farming and rural industries as well as commerce and manufacturing: Devizes in Wiltshire, a typical small market town which was prosperous but had a stable population of about 9,000 over the entire period, using *Gillman's Devizes Public Register and Business Directory*; and Idle in Yorkshire, a township and small market centre between Bradford and Shipley with large manufacturers of woollens and worsteds, stone quarrying, and general trades with growing population (11,600 in 1851, reaching 13,400 in 1881), using *White's Clothing District Directory*. Both were well-connected to a major canal and had rail access, but Devizes was a more self-contained centre for local markets, while Idle was part of the complex and dispersed industrial and textile development typical of the West Riding. The contrasts allow comparisons across rural and urban, stable and growing, with different sector mixes.

Table 5 compares the directory content with the census after 1881 supplementation. The comparison of columns 1 plus 5, with 7, indicate that the numbers identified in the directories and census are close, with some discrepancies explained by a small number listed in directories being inactive or not proprietors in the census, although both sources were dated 1881. The directories do contain additional traders not identifiable as individuals in the census (about 8 per cent of the directory list), some of whom relate to companies that had their partners or directors elsewhere. On the other hand, the census had definite local employers or own-account proprietors not listed in the directory. Although only a small percentage, some were significant businesses, especially in Idle, where proprietors resident in Idle were not listed in the directory, whilst others with businesses elsewhere were. This is indicative of inconsistent treatment in directories.

| TD | Traders in TD also active in census (1) | Traders in TD inactive in census (2) | Extra TD traders not in census (3) | Listed in TD; workers or managers in census (4) | Employers & masters in census not listed in TD (5) | Total all TD listings (cols 1-4) (6) | Valid proprietors in TD (cols 1,3 (7) |
|---|---|---|---|---|---|---|---|
| **Devizes** | | | | | | | |
| Male | 305 | 4 | 9 | 10 | 4 | 328 | 318 |
| Female | 91 | 7 | 1 | 2 | 5 | 101 | 99 |
| Total | 396 | 11 | 10 (17)* | 12 | 9 | 429 | 417 |
| **Idle** | | | | | | | |
| Male | 388 | 9 | 30 | 11 | 16 | 438 | 418 |
| Female | 48 | - | 5 | - | 2 | 53 | 53 |
| Total | 436 | 9 | 36 (41)* | 11 | 18 | 492 | 481 |

* plus 7 companies Devizes, and 6 companies Idle without director/partner names (in brackets)

**Table 5.** Trade directory listings for all proprietors in Devizes and Idle 1881; note sex in directories is sometimes impossible to infer resulting in potential for overestimating males (more precise in Devizes than Idle).

Table 6 compares the census with the directories, before and after supplementation. It confirms that the supplemented census generally gives fuller coverage of proprietors, especially own account. However, before supplementation those extracted from census responses as employers or masters shown on the left side were about 20 per cent of the directory counts (compare Table 5); whereas after supplementation, on the right of Table 6, estimates of the total proprietor population of 498 in Devizes and 609 in Idle, are 19 and 27 per cent higher than directories, respectively. This is mainly because directories include fewer own account traders, especially women. Census supplementation also gives separate estimates for employers, own account and workers which are not available from directories., and is particularly valuable in identifying a much larger number of female proprietors than recorded in the directories. Women are almost all the additional individuals identified in Devizes, and half of the additional in Idle.

| Census | Census extracted responses | | | Census after supplementation | | | |
|---|---|---|---|---|---|---|---|
| | Employers* | Own account | Total | Employers* | Own account | Workers | Proprietors |
| **Devizes** | | | | | | | |
| Male | 55 | 14 | 69 | 154 | 157 | 1,741 | 311 |
| Female | 2 | 1 | 3 | 20 | 167 | 955 | 187 |
| Total | 57 | 15 | 72 | 174 | 324 | 2,696 | 498 |
| **Idle** | | | | | | | |
| Male | 98 | 14 | 112 | 178 | 285 | 3,215 | 463 |
| Female | 4 | 2 | 6 | 25 | 121 | 2,088 | 146 |
| Total | 102 | 16 | 118 | 203 | 406 | 5,203 | 609 |

* includes farmers

**Table 6.** Census listing of proprietors in Devizes and Idle 1881: employers extracted and tagged, from intelligence-led Supplementation 1.

More detailed comparisons demonstrate the challenges of supplementation. For ‘grocers’, for example, numbering 20 in Devizes and 69 in Idle, all census extracted employers matched individuals listed as ‘grocer’ in the directory, as did all but 2 and 3, respectively, of supplemented individuals. A further 6 census-identified grocers in Devizes and 24 in Idle were not shown in directories, of which 4 and 19 were estimated by supplementation (mainly partners not shown in directories). This shows good matching and demonstrates the advantages of sub-dividing occupations, as shown for grocers in Appendix Figure A.1. There is also close matching for farmers, most retailers and many maker-dealers, such as: bakers, confectioners, butchers, fruiterers, saddlers, victuallers, physicians, hairdressers, drapers, tailors, and cabinet makers. There is perfect matching of Devizes coal merchants, ironmongers, music teachers, solicitors, watchmakers, waste dealers, and other small trades, and near-perfect for Idle in the same trades.

For other sectors, the use of the same occupational descriptors by both proprietors and workers limits precise supplementation infills. For example, the numerical counts were similar for the generic descriptor ‘blacksmith’ between directories and the census, but only one person in each place was the same. Similarly many building trades such as carpenters, painters, plasterers have good numerical matches but poor identification of individuals

(though plumbers were well matched in the supplemented). The largest challenge was in Idle where hundreds of ‘weavers’ and ‘spinners’ in wool and worsted were not differentiated between workers and proprietors; three spinners and two weavers identified in the directory were matched in the supplementation but only one employer extracted in Group 1. Similarly the large ‘stone mining’, ‘quarrying’ and ‘stone dressing’ trade in Idle has five 14 directory entries matched, which include all extracted. However, the directory ignores others known from local histories to be operating as small proprietors in a large and famous Idle industry at the time, which were estimated as 26 proprietors in the census supplementation.

An important aspect is that all-female categories listed in directories are exactly matched for 19 ‘dressmakers’, and 7 ‘milliners’ in Devizes, and 5 dressmakers, and 5 of 6 milliners in Idle. But the census also indicates a further 51 and 5 in Devizes, and 51 and 10 in Idle, respectively, who were proprietors, as well as differentiating the numerous workers. Hence, although there are imperfections, the local directory comparisons confirm good estimation by census extractions and supplementation, the wider coverage offered by the census especially of own account and women, but also the limitations for identifying proprietor status of specific individuals who have generalised occupational descriptors.

## 7. Assessment and conclusion

This paper has sought to develop a methodology to supplement the census responses for employers and own-account proprietors in the early British censuses (1851-81) to align them with later censuses. This provides an entry point for many areas of subsequent research by business historians. The fundamental contribution of the paper is methodological: to develop estimates of the whole population of proprietors as a benchmark resource. A major part of the paper has discussed what the original census responses provide, and the steps that are required to render these into consistent and robust estimates of the proprietor population. Whilst there are limitations that derive from the format of the original census process, the estimates have fairly narrow confidence bounds and small differences between alternatives, especially for employers. Moreover, the employer supplementation can be confirmed as robust against alternative large-scale estimates available from national directories, and the turning point of 1901-11 is also confirmed by comparison against the 1921 census. A sample of local directories in 1881 also confirms the robustness of the supplemented estimates of

proprietor *numbers* and that generally greater numbers of proprietors are identified in the census than in directories, especially women. However, the matching for specific *individuals* varies in quality, mainly weaker where generic descriptors like blacksmith, weaver, spinner or stone quarryman are used. However, many own account and female supplementations, which are expected to be difficult to estimate, match surprisingly well.

Given the imperfect design of the census questions there remain limitations. It is impossible to have totally reliable supplements for responses that were never given at the census date. However, the paper overcomes most selection biases in much previous research and allows initiation of a new agenda for business research using whole-population data on proprietors. The alternative supplementations together with the tagged extractions of the original census responses are available in the BBCE data deposit; all supplementation decisions are provided at https://www.bbce.uk/. This will allow other researchers to replicate and explore improvements

The analysis of the new data demonstrates a strong contrast between the long-term steady increase in employer numbers, and the rise and then decline of own account. The developments had strong compositional effects by sector and gender, with the sector distribution showing that the own-account rise and then decline was mainly driven by numerical changes of men in maker-dealing, refreshments, and food sales, and by changes of women's rate of own account proprietorship participation.

Many different analyses can be developed from these new data. We have focused on three debates. We demonstrate that the Victorian period does not support the declinist contention: there was no major decline in *proprietor numbers* until 1901. The new estimates do show a slow decline of employer *rates of entrepreneurship*, with brief reversal in 1911. However, the decline is trivial compared to what happened over 1901-21. The new estimates place decline later than previously claimed in the declinist literature, but accord better with the arguments about a 'climacteric' around 1901. The new estimates now allow historians to focus on actual proprietors and sectoral change, rather than relying on economic aggregates, case studies, or patchy literary evidence to generalise about motivations.

Since our analysis allows the first reasonably robust estimates of long-term entrepreneurship rates, it also allows inference about the origin of declines in these rates. Rather than declining

entrepreneurship, it is now clear that the changes over 1851-1911 derived primarily from changes in the participation of the economically active in the waged labour market, and hence changes in industrial organisation towards concentration in larger firms. Extractions of workforce numbers directly from the employers' census returns confirm previous knowledge (e.g. Gatrell, 1977; Hannah, 1983, 2007), as discussed more fully elsewhere (Bennett, 2019a: 113-6). They show mean firm size increasing from 7.3 to 11.8 over 1851-81, with firms with over 500 employees increasing from only about 160 in 1851, accounting for 163,000 workers, to about 430 in 1881 employing over 450,000 workers. Those with over 50 employees increased from about 2,900 with 429,000 workers in 1851, to 5,800 firms with over 1.1m workers in 1881. Hence, the actions of only a few hundred or few thousand large-firm proprietors who were expanding had most impact. In contrast, the much more numerous firm-size classes between 5 and 25 employees failed to expand at a rate equivalent to the economically active, resulting in their proportion of firms and workforce declining over the period. Hence, it was not the major firms and their entrepreneurs that were conservative or focused outside the world of business, as suggested by Perkin (1969), Wiener (1981) and Rubinstein (2006). Rather it was the slower development of small-medium firms. This was more a result of increasing large-firm dominance, which choked off the pipeline of smaller businesses, than entrepreneurial failure. This issue of firm-size shifts deserves further attention in future research.

What is clear is that the increasing dominance by large firms impacted the waged labour market. Partly this was a result purely of increasing numbers of firms that were very large that could absorb the expanding population. But in terms of effects on entrepreneurship, as indicated in the rate estimates, there were also major effects on incentives. Increasingly, waged employment in many of the numerically largest sectors (retail, manufacturing, agricultural produce) offered comparable or better pay than the profits to be earned in small and marginal firms (especially own-account proprietorship), and also usually gave greater stability and lower risk. This increased the incentives towards being a worker. The resulting very rapid expansion of the worker population appears to be the key factor that drove down the overall entrepreneurship rate at the end of the nineteenth century, and also the chief cause of the deepening 'U'-shaped distribution of entrepreneurship rates through the first half of the twentieth century.

These changes also interplayed with expansion of the corporate sector which accounted for the overwhelming majority of the largest firms by the 1920s (Hannah, 1983, 2014). This paper mainly engages with non-corporate proprietors because the census questions were imperfect at eliciting replies for corporate directors or managers and have to be supplemented in a different way, investigated paper in future research. However, the effect of growth of corporations can be inferred from Feinstein's (1972) estimates. Non-corporate output continued to grow up to 1911, but was a rapidly *decreasing share* of total output compared to corporations.[21] Given the trends in proprietor numbers evident in this paper, it is clear that the slowdown 1901-11 coincided with the rapid rise in corporate numbers and incomes which changed fundamentally the balance of incentives towards incorporation, and had critical influences on incentives to take waged posts as foremen or managers in larger firms rather than struggling as a proprietor on one's own.

Many of the trends that we have highlighted have been previously observed, but our estimates now begin to measure the changes more precisely and unravel the sector composition. We can also date the tipping point when entrepreneurship rates began a steep decline to the period mainly after 1901. This confirms the starting point of the 'U'-shaped pattern of the twentieth century. Clearly further research is required on the details: especially relative corporate and non-corporate roles, the dynamics of different parts of the firm-size distribution, and the relation between employers and own account in different sectors. The methodological developments presented here open the way for future researchers to explore these and other questions in a way not previously possible.

**References**

Alexander, A., Shaw, G. and Hodson, D. (2003) 'Regional Variations in the Development of Multiple Retailing in England, 1890-1939', in J. Benson and L. Ugolini (eds.), *A Nation of Shopkeepers: Five Centuries of British Retailing*, London: I.B. Tauris, pp. 127-54.

Allanson, P. (1992) 'Farm Size Structure in England and Wales, 1939-89', *Journal of Agricultural Economics*, 43/2 (1992) pp. 137-148.

[21] Feinstein estimates (1972: Tables 23 and 29) for all-UK; further comparison in Bennett et al. (2019a: 129-32).

Amemiya, T. (1981) 'Quantitative response models: a survey', *Journal of Economic Literature*,19, 1483-1536.

Amemiya, T. (1985) *Advanced Econometrics*, Oxford: Blackwell.

Anderson, M. (1999) 'What can the mid-Victorian censuses tell us about variations in married women's employment?', *Local Populations Studies,* 62, 9-30.

Aston, J. (2016) *Female Entrepreneurship in Nineteenth-Century England, Engagement in the Urban Economy*, Basingstoke: Palgrave Macmillan.

Bennett, R.J. and Newton, G. (2015) 'Employers and the 1881 population census of England and Wales', *Local Population Studies*, 94, 29-49.

Bennett, R.J., Smith, H., van Lieshout, C. and Newton, G. (2017) '*Occupations and occupational aggregation of census data 1851-1911',* Working Paper 5: ESRC project ES/M010953, University of Cambridge, https://doi.org/10.17863/CAM.9874

Bennett. R.J., Smith, H. and Montebruno, P. (2018a) 'The Population of Non-Corporate Business Proprietors in England and Wales 1891-1911', *Business History*, 62, 1341-72.

Bennett, R.J., Montebruno, P., Smith, H., and van Lieshout, C. (2018b) '*Reconstructing entrepreneur and business numbers for censuses 1851-81*', Working Paper 9: ESRC project ES/M010953, University of Cambridge, https://doi.org/10.17863/CAM.37738

Bennett, R.J., Smith, H., van Lieshout, C., Montebruno, P., Newton, G. (2019a) *The Age of Entrepreneurship: Business proprietors, self-employment and corporations since 1851.* London: Routledge.

Bennett, R.J., Montebruno, P., Smith, H., and van Lieshout, C. (2019b) '*Reconstructing business proprietor responses for censuses 1851-81: a tailored logit cut-off method'*, Working Paper 9.2: ESRC project ES/M010953, University of Cambridge, https://doi.org/10.17863/CAM.43891

Bennett, R.J., Montebruno, P., Smith, H., and van Lieshout, C. (2019c) '*Entrepreneurial discrete choice: Modelling decisions between self-employment, employer and worker status*'. Working Paper 15: ESRC project ES/M010953, University of Cambridge, https://doi.org/10.17863/CAM.37312

Bennett, R.J., Smith, H., van Lieshout, C., Montebruno, P., Newton, G. (2020) *British Business Census of Entrepreneurs, 1851-1911*. [data collection] UK Data Service. SN: 8600, http://doi.org/10.5255/UKDA-SN-8600-2

Bolton, J.E. (1971) *Small Firms: Report of the Committee of Inquiry on Small Firms,* Cmnd 4811, HMSO, London.

Bowley, A.L. (1920) *The change in in the distribution of the National Income 1880-1913*, Clarendon Press, Oxford.

Broadberry, S. (2014) 'The rise of the service sector', in R. Floud, J. Humphries and P. Johnson (eds.), *The Cambridge Economic History of Modern Britain: Volume II, 1870 to the Present*, Cambridge: Cambridge University Press, 330-61.

Caree, M.A, Stel, A van, Thurik, A.R. and Wennekers, S. (2007) *The Relations Between Economic Development and Business Ownership Revisited*, Discussion Paper TI 2007-022.3. Amsterdam/Rotterdam, Tinbergen Institute.

Census of Great Britain (1961) *Occupation Report*, London: HMSO.

Chandler, A.D. (1962) *Strategy and Structure: Chapters in the History of the Industrial Enterprise*, Cambridge MA: MIT Press.

Chandler, A.D. (1990) *Scale and Scope: The Dynamics of Industrial Capitalism*, London: Belknap Press.

Clapham, J.H. (1932), *An Economic History of Modern Britain, Vol. 2: Free trade and steel 1850-1886*, Cambridge: Cambridge University Press.

Clark, C. (1957) *The Conditions of Economic Progress*, third edition, Macmillan, London.

Colman D.C. and Macleod, C. (1986) Attitudes to new techniques: British businessmen, 1800—1950, *Economic History Review*, 39, 588-611.

Conrad, F.G., Couper, M.P. and Sakshaug, J.W. (2016) 'Classifying open-ended reports: Factors affecting the reliability of occupation codes', *Journal of Official Statistics*, 32, 1, 75-92.

Çrouzet, F. (1985) *The First Industrialists*, Cambridge: Cambridge University Press.

Davidoff, L. and Hall, C. (1997) *Family Fortunes: Men and Women of the English Middle Class 1780-1850*, Revised Edition, Routledge, London.

Edgerton, D. (1997) 'The Decline of Declinism', *Business History Review*, 71/2, 201-6.

Ekelund, R. B. and Hébert, R.F. (1983) *A History of Economic Theory and Method*, Second edition, New York: Mc-Graw Hill.

Feinstein, C.H. (1972) *National Income, Expenditure and Output of the United Kingdom 1855–1965*, Cambridge: Cambridge University Press.

Feinstein, C. (ed.) (1997) *Introduction: Themes in the economic development of the United Kingdom since 1870*, vol. 1., Edward Elgar, Cheltenham.

Gatrell, V.A.C. (1977) 'Labour, Power and the Size of Firms in Lancashire Cotton in the Second Quarter of the Nineteenth Century', *Economic History Review*, 30(1), 95-139.

Hannah, L. (1983) *The Rise of the Corporate Economy*, Second edition, London: Methuen.

Hannah, L. (2007) 'The "Divorce" of Ownership from Control from 1900 Onwards: Re-Calibrating Imagined Global Trends', *Business History*, 49, 4, 404-438.

Hannah, L. (2014) 'Corporations in the US and Europe 1790-1860', *Business History*, 56, 6, 865-99.

Higgs, E. (1987) 'Women, Occupations and Work in the Nineteenth-century Censuses', *History Workshop Journal*, 23, 59-80.

Higgs, E. (2005) *Making Sense of the Census Revisited: Census Records for England and Wales 1801-1901*. London: Institute of Historical Research, National Archives.

Higgs, E., Jones, C., Schürer, K. and Wilkinson, A. (2015), *Integrated Census Microdata (I-CeM) Guide*, 2nd. ed. Colchester: Department of History, University of Essex.

Higgs, E. and Wilkinson, A. (2016) 'Women, Occupations and Work in the Victorian Censuses Revisited', *History Workshop Journal,* 81, 17-38.

Jabine, T.B. and Tepping, B. J. (1973) 'Controlling for the quality of occupation and industry data', *Bulletin de L'Institute International de Statistique*, 50, 360-92.

Jefferys, J.B. (1954) *Retail Trading in Britain, 1850-1950: A study of trends in retailing with special reference to the development of Co-operative, multiple shop and department store methods of trading*, Cambridge: Cambridge University Press.

Jeremy, D.J. (1991) 'The Hundred Largest Employers in the United Kingdom, in Manufacturing and Non-Manufacturing Industries, in 1907, 1935 and 1955', *Business History*, 33, 1, 93-111.

Jeremy, D.J. (1998) *A Business History of Britain, 1900-1990s*, Oxford University Press.

Kay, A.C. (2009) *The Foundations of Female Entrepreneurship: Enterprise, Home and Household in London, c. 1800-1870*, London: Routledge.

*Kelly's Directory of the Cabinet, Furniture* (afterwards *Furnishing) and Upholstery Trades ... etc.*, Kelly & Co., London, vol. 1, 1877 -.

*Kelly's Directory of the Manufacturers of Textile Fabrics ...;* afterwards *Directory of the Textile Industries ...* , Kelly & Co., London, vol. 1 1880 -.

Kirzner, I. M. (1973) *Competition and Entrepreneurship*, Chicago, IL: Chicago University Press.

Kuznets, Simon (1966) *Modern Economic Growth: Rate, Structure and Spread*, New Haven, CT: Yale University Press.

Landes, David S. (1969) *The unbound Prometheus: Technological change and industrial development in West Europe from 1750 to the present*, Cambridge: Cambridge University Press.

Lazonick, W. (1983), 'Industrial Organization and Technical Change: The Decline of the British Cotton Industry', *Business History Review*, 57(2), 195-236.

Lee, C.H. (1984) 'The service sector, regional specialization, and economic growth in the Victorian economy', *Journal of Historical Geography*, 10/2, 139-55.

Lesage, É., Haziza, D., and D'Haultfoeuille, X. (2019) 'A Cautionary Tale on Instrumental Calibration for the Treatment of Nonignorable Unit Nonresponse in Surveys', *Journal of the American Statistical Association*, 114, 526, 906-915. http://dx.doi.org.ezp.lib.cam.ac.uk/10.1080/01621459.2018.1458619

Lewis, W.A. (1978) *Growth and Fluctuations, 1870-1913*, Allen and Unwin, London.

Lund, P. and Price, R. (1998) 'The Measurement of Average Farm Size', *Journal of Agricultural Economics*, 49, 1, 100-110.

Lyberg, L. and Kasprzykm, D. (1997) 'Post-survey processing and operations', in Lyberg, L. Biemer, P.P., Collins, M., de Leeuw, E., Dippo, C., Scwharz, N. and Trewin, D., 353-70, *Survey measurement and process quality*. New York: Wiley.

Macmillan Committee (1931) *Committee on Finance and Industry: Report*, Chair Harold Macmillan, Cmd. 3897. London, HMSO.

Marshall, A., (1919) *Industry and Trade: A Study of industrial Technique and Business Organization; and their Influence on the Conditions of Various Classes and Nations*, London: Macmillan.

Martin, J., Bushnell, D., Campanelli, P. and Thomas, R. (1994) 'A comparison of interviewer and office coding of occupations', *Survey Methods Centre Newsletter*, Social and Community Planning Research, 15, 2, 18-25

McCloskey, D.N. (1970) Did Victorian Britain fail?, *Economic History Review*, 23, 446-59.

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.

Montebruno, P. Bennett, R.J., Smith, H. and van Lieshout, C. (2020) 'Machine learning classification of entrepreneurs in British historical census data', *Information Processing & Management,* 57, 3 https://doi.org/10.1016/j.ipm.2020.102210

Nicholas, T. (1999) 'Clogs to Clogs in Three Generations? Explaining Entrepreneurial Performance in Britain Since 1850', *Journal of Economic History*, 59(3), 688-713.

Nicholas, T. (2000) 'Wealth Making in the Nineteenth- and Early Twentieth Century: The Rubinstein Hypothesis Revisited', *Business History*, 42, 2, 155-68.

Nye, J. V.C. (1991), 'Lucky fools and cautious businessmen: On entrepreneurship and the measurement of entrepreneurial failure', in J.Mokyr ed. *The Vital One: Essays in Honor of Jonathan R.T. Hughes*, Greenwich, CT: Research in Economic History Supplement 6, Research Annual, JAI Press, 131-52.

ONS (2019) 'Methodology: Census', https://www.ons.gov.uk/census/censustransformationprogramme/progressanddevelopment/methodology

Payne, P.L. (1988) *British Entrepreneurship in the Nineteenth Century*, Second Edition, Macmillan, London.

Peress, M (2010) 'Correcting for Survey Nonresponse Using Variable Response Propensity', *Journal of the American Statistical Association*, 105, 492, 1418-30.

Perkin, H. (1969) *The origins of modern English society 1780-1880*, London: Routledge & Kegan Paul.

Phelps Brown, E.H. and Handfield-Jones, S.J. (1952) The Climacteric of the 1890s: A study in the expanding economy, *Oxford Economic Papers*, 4, 266-90.

Pollard, S., (1968) *The Genesis of Modern Management: A Study of the Industrial Revolution in Great Britain*. London: Edward Arnold.

PP (1888) *Report from the Select Committee on Small Holdings; with the proceedings of the committee.* House of Commons Papers 358, XVIII. 413.

PP (1895) Small Holdings Act, 1892. *Return for each County of Number of Petitions under Small Holdings Act, 1892; Amount of Land acquired and Number of Small Holdings provided*. House of Commons Papers, 407, LXXXIV.49.

Rubin D.B. (1987) *Multiple imputation for nonresponses in surveys*, New Jersey: Wiley.

Rubin D.B. (1988) 'An Overview of Multiple Imputation', *Proceedings of the Survey Research Methods Section of the American Statistical Association*, 79-84.

Rubin D.B. (1996) 'Multiple imputation after 18+ years', *Journal of the American Statistical Association*, 91, 434, 473-489

Rubinstein, W.D. (2006) *Men of Property: The Very Wealthy in Britain Since the Industrial Revolution*, second edition, London: Social Affairs Unit.

Schürer, K., Higgs, E. (2014) *Integrated Census Microdata (I-CeM), 1851-1911*. [data collection] UK Data Service. SN: 7481, http://doi.org/10.5255/UKDA-SN-7481-1

Woollard, M., Schürer, K. (2000) *1881 Census for England and Wales, the Channel Islands and the Isle of Man (Enhanced Version)* [data collection] SN: 4177, http://doi.org/10.5255/UKDA-SN-4177-1

Scott, P. (1994), 'Learning to Multiply: The Property Market and the Growth of Multiple Retailing in Britain, 1919-39', *Business History*, 36, 3, 1-28.

Shaw-Taylor, L. and Wrigley, E.A. (2014) 'Occupational structure and population change', in R.Floud, J. Humphries and P. Johnson (eds.), *The Cambridge Economic History of Modern Britain: Volume I, 1700-1870*, Cambridge: Cambridge University Press, 53-88.

Smith, H. Bennett, R.J., van Lieshout, C. and Montebruno, P. (2020) 'Households and entrepreneurship in England and Wales, 1851–1911', *The History of the Family*. https://doi.org/10.1080/1081602X.2020.1796750

Smith, H., Bennett, R.J., van Lieshout, C. and Montebruno, P. (2021) 'Entrepreneurship in Scotland, 1851-1911', *Journal of Scottish Historical Studies*, 41

Taylor, F. W. (1911) *The principles of scientific management*, New York: Harper & Brothers.

*The Post Office* (afterwards *Kelly's) Directory of Chemists and Druggists ... ;* afterwards *Kelly's Directory of the Chemical Industries* …, Kelly & Co., London, vol. 1, 1869 -.

*The Post Office* (afterwards *Kelly's*) *Directory of the Engineers and Iron and Metal (& Motor) Trades ...* ; later *Kelly's Directory of the Engineering, Hardware, Metal and Motor Trades*, Kelly & Co., London, vol. 1, 1870 -.

*The Post Office* (afterwards *Kelly's*) *Directory of the Building Trades ...,* Kelly & Co., London, vol. 1, 1870 -.

*The Post Office* (afterwards *Kelly's*) *Directory of the Grocery & Oil & Color Trades: with which is included the tobacco and provision trades, and wholesale confectioners ...* , Kelly & Co., London, vol. 1, 1870 -.

*The Post Office* (afterwards *Kelly's*) *Directory of the Leather Trades*, Kelly & Co. London vol. 1, 1872-.

*The Post Office* (afterwards *Kelly's*) *Directory of Stationers, Printers, Booksellers, Publishers, and Paper Makers ...* ; Kelly & Co., London, vol. 1, 1872 -.

*The Post Office* (afterwards *Kelly's*) *Directory of the Watch and Clock Trades, Goldsmiths and Jewellers ... ;* afterwards *Directory of the Watch, Clock and Jewellery Trades*, Kelly & Co., :London, vol. 1, 1872 -.

*The Post Office Directory of the Brewers and Maltsters and other Trades connected therewith*; (afterwards *Kelly's*) *Directory of the Wine and Spirit Trades, with which*

*are included Brewers and Maltsters, and Other Trades connected therewith, of England, Scotland and Wales, and also the principal Wine Merchants, etc., on the Continent*, Kelly & Co., London, vol. 1, 1877-.

*The Post Office Directory of the Merchants and Manufacturers of The United Kingdom, being a Guide to the Export, Shipping and Manufacturing Industries*; afterwards *Kelly's Directory of the Merchants, Manufacturers and Shippers of The United Kingdom, Being a Guide to the Export and Import Shipping and Manufacturing Industries*, Kelly and Co., London, vol. 1, 1877 -.

Tomlinson, J. (2009), 'Thrice Denied: "Declinism" as a Recurrent Theme in British History in the Long Nineteenth Century', *Twentieth-Century British History*, 20.2, 227-51.

Tomlinson, J. (2016) 'De-industrialization Not Decline: A New Meta-Narrative for Post-War British History', *Twentieth-Century British History*, 27, 1, 76-99.

US Bureau of the Census (1950) *The post-enumeration survey: 1950*. Washington D.C.: US Government Printing Office.

van Lieshout, C., Day, J. Montebruno, P. and Bennett, R.J. (2018) '*Extraction of data on Entrepreneurs from the 1871 Census to supplement I-CeM*', Working Paper 12: ESRC project ES/M010953, University of Cambridge https://doi.org/10.17863/CAM.27488

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 http://doi.org/10.1080/03071022.2019.1656929

Wardley, P. (1999) 'The emergence of big business: The largest corporate employers of labour in the United Kingdom, Germany and the United States c. 1907', *Business History*, 41, 4, 88-116.

Weeks, A., Fallows, A, Broad, P., Merad, S. and Ashworth, K. (2018) *Non-response Weights for the UK Labour Force Survey? Results from the Census Non-response Link Study*, London: Office of National Statistics.

Wennekers, S., Stel, A. van, Carree, M., and Thurik, R. (2010) *The relationship between entrepreneurship and economic development: is it U-shaped?* EIM Research Reports, Zoetermeer.

Wiener, M.J. (1981) *English Culture and the Decline of the Industrial Spirit, 1850–1914*, Cambridge: Cambridge University Press.

You, Xuesheng (2020) Women's labour force participation in nineteenth-century England and Wales: evidence from the 1881 census enumerators' books. *Economic History Review*, 73, 106-33.

Zhang, L.-C., Thomsen, I. and Kleven, Ø. (2013) 'On the Use of Auxiliary and Paradata for Dealing with Non-sampling Errors in Household Surveys'. *International Statistical Review*, 81, 2, 270-88.

**Appendix:**

***Logit Estimator and alternatives tested***

Table A.1 gives the estimates of the logit model for 1891. This is the probability of a census response as proprietor (employer and own account) compared to worker, after non-economically active removed; base categories are worker, urban areas, male and single, household head. Density measured at RSD level; RELA10 is an aggregation of the I-CeM full list of relationship codes; full definitions and sub-occode estimates reported in (Bennett et al., 2018b, 2019b). Logit weighted to adjust for non-response and misallocation bias in 1891; N of observations 7,109,988; Pseudo $R^2$ 0.456.

| **Estimation of Entrepreneur = 1** | **Coefficient** | **z value** |
|---|---|---|
| Age | 0.135*** | (245.76) |
| Age # Age | -0.00102*** | (-170.21) |
| | | |
| **Density and urban dummies:** | | |
| Density | -0.00821*** | (-139.63) |
| Density # Density | 0.00002*** | (90.38) |
| Urban # Density | 0 | (.) |

| | | |
|---|---|---|
| Hinterland 1 # Density | 0.00648*** | (3.56) |
| Hinterland 2 # Density | 0.00678*** | (4.34) |
| Rural # Density | 0.179*** | (36.09) |
| Urban # Density # Density | 0 | (.) |
| Hinterland 1 # Density # Density | -0.00016 | (-1.88) |
| Hinterland 2 # Density # Density | 0.000011 | (0.66) |
| Rural # Density # Density | -0.00323*** | (-27.54) |
| **Gender and marital status:** | | |
| Male # Single | 0 | (.) |
| Male # Married | -0.100*** | (-16.64) |
| Male # Widowed | -0.00860 | (-1.01) |
| Female # Single | -0.0576*** | (-8.96) |
| Female # Married | 0.144*** | (18.82) |
| Female # Widowed | -0.0430*** | (-5.07) |
| **No. of Servants** | 0.532*** | (156.02) |
| **Household relationships to head (RELA10 codes):** | | |
| 1. Head | 0 | (.) |
| 2. CFU member (conjugal family unit) | -0.825*** | (-139.54) |
| 3. Older generation than CFU | -0.903*** | (-53.14) |
| 4. Siblings | -0.721*** | (-71.67) |
| 5. Other family | -1.065*** | (-77.37) |
| 6. Servants | -3.300*** | (-76.88) |
| 7. Working title (assistant etc.) | -2.829*** | (-80.60) |
| 8. Lodgers/boarders | -1.187*** | (-162.21) |
| 9. Non-household (visitors etc.) | -1.460*** | (-55.05) |
| 10. Unknown RELA | -0.601*** | (-48.34) |

**Table A.1.** Logit estimates for probability of responding as proprietor in 1891; *** p≥0.01; # indicates interaction variables; estimate made for each 844 sub-occode are not reported here; note z-values for the logit are equivalent to t-tests in normal regression.

The logit was preferred to a probit because in these census data there are many variables explaining entrepreneur response that are interdependent and non-additive (between individuals), and the individuals themselves are potentially interdependent (within households). Although often preferred in econometrics, for the probit to be valid the individuals have to be independent drawings, and the factors influencing the binary distinction have to be independent of each other and additive so that in the limit the probability link function is cumulative normal (Amemiya, 1985: 269). Also in modern applications (e.g. Peress, 2010; Lesage et al., 2019) dichotomous classifiers like the logit have been shown to be equivalent to multiple imputation techniques like those favoured by Rubin (1987, 1996). Although the probit can be manipulated and usually yields equivalent results to the logit (Amemiya, 1981), given the data structure the logit is preferable.

The logit was initially tested by using it on the 1891 data to see how far it estimated the actual data correctly. This led to a relatively large set of covariates being adopted (as listed in Table A.1, plus the 844 sub-occode which are not reported). Various more parsimonious estimators were initially attempted. They gave only marginally larger pseudo $R^2$, and tests of randomness of the residuals exhibited a high level of clustering in either urban or remote rural areas. They also gave unsatisfactory estimates in occupational categories which had extensive in-household relationships for employment of family and others, often involving high female participation. These residual relationships reflected the expected interrelationships in the data,[22] and were a reason to favour the logit. They were particularly high for the large categories of lodging-house keepers, dressmakers, and laundry working. In comparison the expanded non-parsimonious logit used the entrepreneur response/non-response characteristics for relationships, gender, marital status, and locational characteristics, and passed the randomisation test that there was no significant clustering of residuals.[23]

The logit is the natural estimator to use as a binary classifier (of being a respondent/non-respondent entrepreneur), followed by the binary separation of respondent employers and own-account. The most direct alternative is to use a multinomial logit that splits the observations into the three categories in one step: as employers, own account, and workers. This was tested, but was not superior to the logit: it had similar though more clustered

[22] Anderson (1999), Davidoff and Hall (1997), Bennett et al, (2019c).
[23] Residuals tests reported in Bennett et al., (2018b).

residuals, lower pseudo $R^2$, but higher standard errors. It should be noted that, without 'truth' data, there are no statistical tests of fit that can be performed on the early censuses, such as the standard likelihood ratio Chow test, so that comparison tests as used in this paper, or randomisation which is not used, are the only viable ways to proceed.

***Decision stages***

The logit was also used initially to test if supplementation was required. If the logit estimates of entrepreneur numbers were the same or lower than the number of extracted entrepreneurs identified directly using extraction Groups 1-6 after data cleaning and coding, then no supplementation was required. If the estimates were larger, the respondents were supplemented, with the 1891 logit (Method 1) preferred. Other methods had to show major differences to replace it using three decision rules. Method 1 was retained if within ±10 per cent of the equivalent ratio extrapolations for a sub-occode; failing this, a 1901 logit was chosen (Method 2). If neither logit was within this margin, then the 1891 extrapolation ratio was used (Method 3), failing which 1901 extrapolation (Method 4). Comparisons at each stage by sub-occode were made against known patterns from contemporary and secondary literature to tune the feasibility of the final choice to permit more extreme changes if justified by previous commentary. This means that the logit met the test of having in most cases the smallest difference from alternatives methods, as well as according with known contemporary and secondary literature (an intelligence-led input); but the logit was replaced by an alternative if it did not accord. The decision stages are recorded in downloads so that users can replicate our results, or test alternatives (see supplementary material at https://www.bbce.uk/)

The logit was used for almost 90 per cent of sub-occodes, with over half using 1891; 14.5 percent used ratio extrapolations as shown in Table A.2. Of the 844 sub-occodes, although only thirty five had no supplementation, most categories needed only small additions. In addition there were two ways to use the logit estimates, each of which gives decimal probability ranging from 0 (not an entrepreneur) to 1 (definitely an entrepreneur). One way of using the logit was to use rounded (R) probabilities up or down from 0.5; an alternative used unrounded (UR) estimates for all decimal probabilities greater than zero (however small that decimal). Table A.2 shows that the UR estimate was often preferred because many non-

respondents had small aspects that indicated they should have responded as proprietors, especially in the large categories of own-account maker-dealers, retailers and food sellers.

| Method | | Sub-occodes using this method (N) | Sub-occodes (%) | Proprietors estimated by this method (%) |
|---|---|---|---|---|
| Actual tagged responses 1881 extracted: un-supplemented | | 35 | 4.9 | 0.5 |
| Method 1: 1891 logit | Rounded (R) | 23 | 3.2 | 14.7 |
| | Unrounded (UR) | 341 | 47.8 | 46.9 |
| Method 2: 1901 logit | Rounded (R) | 27 | 3.8 | 8.1 |
| | Unrounded (UR) | 184 | 25.8 | 19.9 |
| Method 3: Ratio 1891 | | 91 | 12.7 | 9.5 |
| Method 4: Ratio 1901 | | 13 | 1.8 | 0.4 |

**Table A.2.** Methods used for intelligence-led Supplementation 1 for estimating *all proprietors* (employers plus own account) by sub-occode for 1881.

The benefit of splitting some occupations into 844 sub-occodes is illustrated for the case of the census category of ‘grocers and tea dealers’ in 1851. This was split between a new category where their strings contained only specifically ‘grocers’ or ‘tea dealers’, and a residual containing various other occupations coded by the census as grocers as well as any transcription and I-CeM coding errors. As clear from Figure A.1, splitting led to a clear separation between those who were most likely grocer proprietors (occode 872) with probabilities concentrated above 0.8, and the rest (of I-CeM occode 697) whose response was more likely to be grocery workers with probabilities concentrated below 0.25. This figure also illustrates the more general way in which the logit splits probable respondents and non-respondents in a given sub-occode.

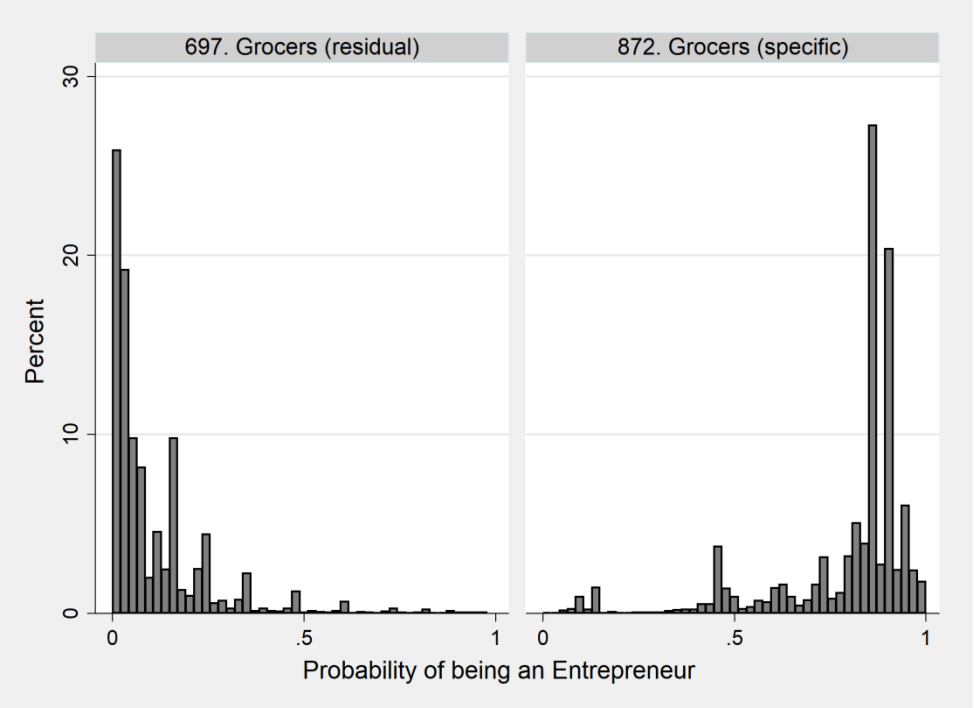


**Figure A.1.** An example of splitting I-CeM occode 697 into sub-occodes for grocers: percent of individuals estimated at each probability level from logit Method 1 for 1851.