SN 5397 – Study Documentation

This document was supplied by the depositor but has been modified by AHDS History

Male occupational data from English parish registers c.1700-1820 (raw data)

Note: Please note that there are two data collections of male occupational data from English parish registers deposited with AHDS History. This collection (SN 5397) contains the raw unprocessed material as it was collected. The other, *Male occupational data from English parish registers c.1700-1820 (processed data)* (SN 5398), contains the processed summary data.

Contents

This data collection consists of 19 files. In most cases each file pertains to one of England's ancient counties.[1] In every case the file contains male occupational data abstracted from virtually all the surviving Anglican baptism registers for that county for the period 1813-20.[2] A list of the counties covered and the number of baptism registers from which occupational data have been collected can be found in Table 1.

Table 1 Data abstractions for 1813-20

County	Number of Rose Act abstractions (1813-20)
Bedfordshire	129
Cambridgeshire	171
Cheshire	132
Cornwall	206
Devonshire	460
Durham	90
Huntingdonshire	95
Lancashire	237
Leicestershire	256
Middlesex	53
Northumberland	110
Oxfordshire	232
Rutland	49
Westmorland	68
Worcestershire	206
City and Ainsty of the City of York	40
Yorkshire East Riding	197
Yorkshire North Riding	230
Yorkshire West Riding	303
Total	*3,264*

[1] However, Yorkshire has been split into the three Ridings and the City and Ainsty of the city of York.
[2] The very small number of missing registers relate to cases where the incumbent has not deposited the records and the relevant archives do not hold microform copies or Bishops' Transcripts.

In the case of the eight ancient counties listed in table 2 the files also contain occupational data culled from Anglican baptism registers for one, and sometimes two, eight-year periods between 1740 and 1799. This represents virtually all the surviving occupational data that met our data quality requirements (described below).[3]

These data are derived from the occupations of fathers recorded in Anglican parish registers at the baptism of a child. From 1813, under the terms of Rose's Act it was a legal requirement to record the fathers' occupation and practice almost universally conformed to the letter of the law. Hence virtually complete geographical coverage of Anglican baptism registers is possible for this period.

Table 2. Eighteenth century data abstractions

County	Parishes with c18th abstracted data
Bedfordshire	63
Cheshire	33
Durham	18
Lancashire	60
City and Ainsty of the City of York	35
Yorkshire East Riding	122
Yorkshire North Riding	96
Yorkshire West Riding	135
Total	*562*

Before 1813 the recording of occupations was not a legal requirement but some incumbents or parish clerks nevertheless chose to record occupations regularly. However, to locate suitable runs of data it was necessary to search all the surviving registers for the whole of the period between 1740 and 1799.[4] Table 2 indicates the number of baptism registers for which suitable runs of occupational data were located and abstracted.

Note

There are two data collections of male occupational data from English parish registers. This collection is of the raw or unprocessed material as it was collected. The other, *Male occupational data from English parish registers c.1750-1820 (processed summary data)*, is of 'processed' material.

Most users simply seeking data on the occupational structure of particular areas will probably wish to use the other collection of processed material. This collection is likely to be of interest only to those who wish to reconstruct the occupational structures of particular areas from the raw data for themselves or specifically to test or probe the way in which we have processed the data.

The processed material consists of counts of occupational descriptors by parish or chapelry whereas the unprocessed material consists of the individual uncounted events. It is important to stress that the raw data do not provide an entirely satisfactory picture of the male occupational structure. To produce the processed data the raw data had to be adjusted in four ways in addition to counting. Firstly, a number of codes used to increase the speed of data collection (for details on field 16 under

[3] Again the small number of exceptions relate to cases where the incumbent has not deposited the records and the relevant archives do not hold microform copies or Bishops' Transcripts.
[4] In the case of Bedfordshire the registers were searched for the period 1690–1799.

Structure below) were stripped out. Secondly, the numbers of events were inflated for those parish registers that had been sampled by the appropriate sampling ratio (for details on field 7 under *Structure* below). Thirdly, years in which occupational recording fell below 95 per cent were rejected (for details see field 16 under *Structure* below). Fourthly, abstractions were standardised to eight years in length by re-weighting them if they were shorter than eight years. These procedures are documented more fully in the documentation accompanying the other collection.

Additionally it should be noted that the data for four counties (Buckinghamshire, Hertfordshire, London and Northamptonshire) exist in the processed form in the other data collection but not in the unprocessed form in this data collection. This is because these datasets, collected early in the project, used a different pen and paper-based form of original data collection.

Provenance

Who created the data and why?

These data were collected as part of a research project run by Dr Leigh Shaw-Taylor and Professor E.A. Wrigley and funded by the Economic and Social Research Council: *Male occupational structure and economic growth in England 1750-1851* (RES-000-23-0131). The data were collected by a number of individuals employed on the project: Joseph Barker, Richard Churchley, Alec Corio, Dr Peter Kitson, Dr Amanda Jones, Dr Victoria Masten, Niraj Modha, Thomas Nutt, Dr Silvia Sovic, Stephen Thompson, Rebecca Tyler, Matthew Ward, Alison Warren and Matthew Westlake. Dr Peter Kitson undertook the quality control of the datasets to ensure that the data met with our requirement that the occupational recording should be 95 per cent or more complete. Dr Kitson then created the 19 datasets in this collection with assistance from Ms Gill Newton from 4,008 original data collection files. All project members were affiliated to the Department of Geography at the University of Cambridge. The aim of this project was to reconstruct the evolution of England's male occupational structure from c.1750 to 1851. The underlying aim was to improve our understanding of the industrial revolution. The results of the project have not, at the time of writing, been published.

Further details can be found on the ESRC website where publications will be listed as they come out: http://www.esrc.ac.uk/. Much more detailed information, including unpublished papers, can be found on the project website:http://www-hpss.geog.cam.ac.uk/research/projects/occupations/

How were the data created?

A major element of the research project was the collection of male occupational data from Anglican baptism registers from the period c.1750-1820. We selected an eight year period because this was on the one hand short enough to be viewed as a single moment in time and on the other hand long enough that any man in a fertile marriage would be likely to appear at least once in the baptism register. In a very few cases the new rules were not implemented immediately. In such cases the eight-year period begins a few years later. In other cases, where one or more years failed a quality test regarding the completeness of occupational recording (described below) extra years were collected.

The coverage is different for the two sub-periods period c.1750-1799 and 1813-1820 as follows:

1 1813-20

From 1st January 1813 it was a legal requirement under Rose's Act of 1812 for the occupations of fathers to be recorded at baptism. It is therefore possible to create parish level datasets of male occupations for every parish in a county from 1813 onwards. For the 19 counties in the datafiles the coverage of parishes and chapelries within the ancient counties is virtually complete (see footnote 2).

2 c.1750-1799

Before 1813 there was no legal requirement to record the fathers' occupation but some registers nevertheless did so. These were located by searching all surviving Anglican baptism registers for the nine counties enumerated in table 2 for the period 1740-1799 for suitable eight-year periods. For these counties virtually all surviving registers have been searched (see footnote 3).

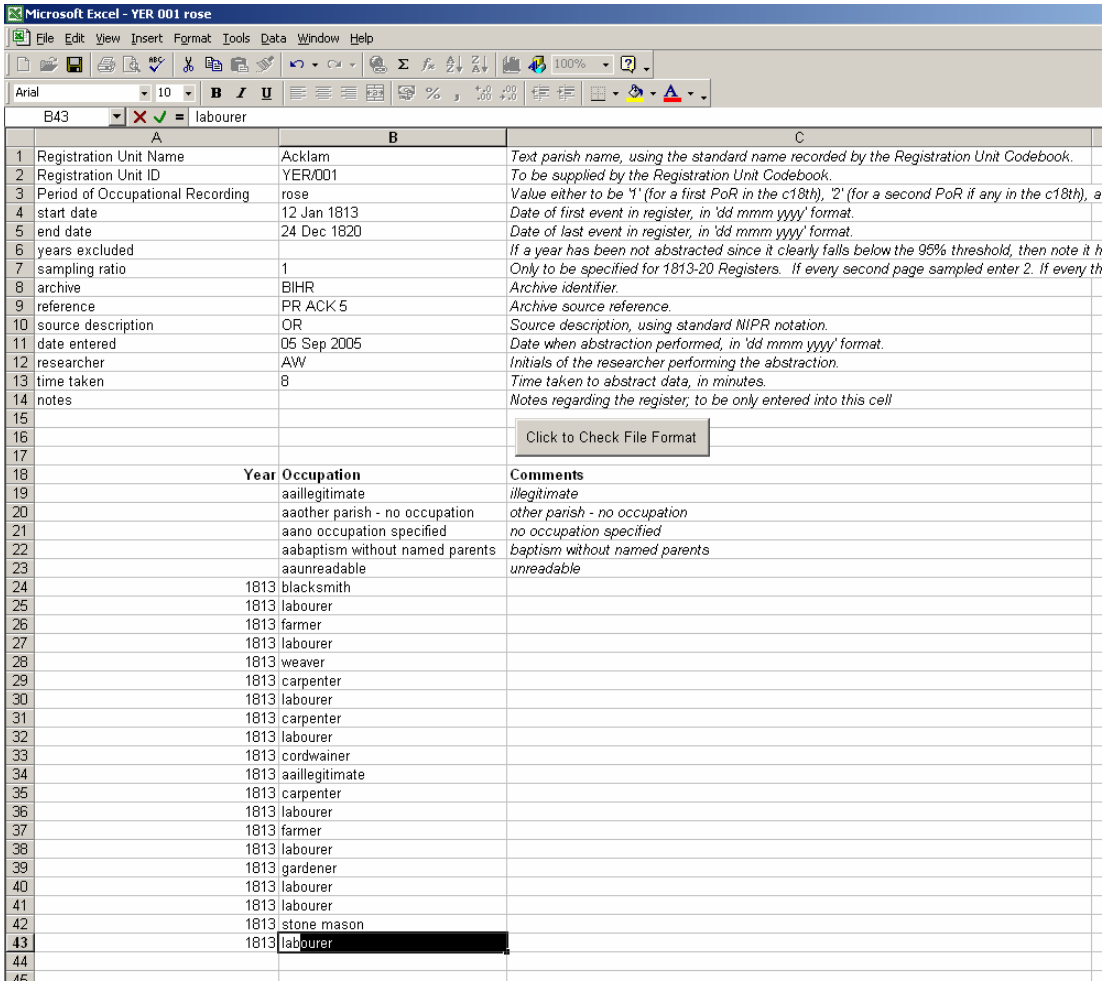
For both periods it was a requirement that occupational recording was 95 per cent or more complete for the period of data abstraction. However, our measure of completeness was not quite straightforward.[5] Firstly, we excluded illegitimate births from our calculation. Secondly we excluded events for which the occupational description was illegible. Thirdly, for the pre 1800 period we discarded cases where the father was from outside the parish or chapelry, and when no occupational data was recorded. Finally, baptisms where no named parents were recorded were also excluded. We rejected runs of data where less than 95 per cent of legible legitimate baptisms of children did not record an occupation. Additionally, we also rejected runs of data which had both less than eight years of data and less than 200 events. To create runs of occupational recording which met these rules sometimes required collecting more than eight years of data and subsequently rejecting individual years. Each run of occupational recording has been termed an abstraction. This data collection comprises all the raw data, including years that were subsequently rejected. The datasets created from it according to these rules form a separate collection.

Before data collection we created a complete list of all parishes and chapelries in each of the 19 ancient counties. These *Registration Units* were assigned a unique identifier and its archival location specified. Research assistants then visited the archives concerned with laptop computers. All the 1813-20 data were then entered direct into an Excel spreadsheet usually from microfilms of the original registers. Searches of the baptism registers for the period 1740-1799 were mostly undertaken using the microfilm or microfiche copies. For data abstraction the microform sources were usually too poor in quality for efficient abstraction. However, in most cases it proved possible to take digital photographs of the original source material. Abstraction then took place in Cambridge with the image on one computer screen and the data entry was performed on another computer, usually a laptop, into Excel files.

Dr Kitson designed the Microsoft Excel data entry form (an example of which is included below) and the research assistants passed the data to him and he then implemented the necessary data quality tests. Dr Kitson with the assistance from Gill Newton then created the county level datasets that form the current collection.

[5] The account which follows is somewhat simplified. In due course we will publish a full account of the data collection rules.

Example of a Parish Register Abstraction Form



	A	B	C
1	Registration Unit Name	Acklam	*Text parish name, using the standard name recorded by the Registration Unit Codebook.*
2	Registration Unit ID	YER/001	*To be supplied by the Registration Unit Codebook.*
3	Period of Occupational Recording	rose	*Value either to be '1' (for a first PoR in the c18th), '2' (for a second PoR if any in the c18th), a*
4	start date	12 Jan 1813	*Date of first event in register, in 'dd mmm yyyy' format.*
5	end date	24 Dec 1820	*Date of last event in register, in 'dd mmm yyyy' format.*
6	years excluded		*If a year has been not abstracted since it clearly falls below the 95% threshold, then note it h*
7	sampling ratio	1	*Only to be specified for 1813-20 Registers. If every second page sampled enter 2. If every th*
8	archive	BIHR	*Archive identifier.*
9	reference	PR ACK 5	*Archive source reference.*
10	source description	OR	*Source description, using standard NIPR notation.*
11	date entered	05 Sep 2005	*Date when abstraction performed, in 'dd mmm yyyy' format.*
12	researcher	AW	*Initials of the researcher performing the abstraction.*
13	time taken	8	*Time taken to abstract data, in minutes.*
14	notes		*Notes regarding the register, to be only entered into this cell*
15			
16			Click to Check File Format
17			
18	**Year**	**Occupation**	**Comments**
19		aaillegitimate	*illegitimate*
20		aaother parish - no occupation	*other parish - no occupation*
21		aano occupation specified	*no occupation specified*
22		aabaptism without named parents	*baptism without named parents*
23		aaunreadable	*unreadable*
24	1813	blacksmith	
25	1813	labourer	
26	1813	farmer	
27	1813	labourer	
28	1813	weaver	
29	1813	carpenter	
30	1813	labourer	
31	1813	carpenter	
32	1813	labourer	
33	1813	cordwainer	
34	1813	aaillegitimate	
35	1813	carpenter	
36	1813	labourer	
37	1813	farmer	
38	1813	labourer	
39	1813	gardener	
40	1813	labourer	
41	1813	labourer	
42	1813	stone mason	
43	1813	labourer	
44			

What sources were used to create the data collection?

The data collection was created from 4,064 Anglican baptism registers. Full source details are provided in the data files as documented below.

Structure

Table 3 indicates, for each county, the name of the data file, the number of records in that file and the size of the file in kilobytes.

Table 3 Data Files

County	Name of file	Number of records	Size of file in Kbytes
Bedfordshire	BED.txt	31,412	4,831
Cambridgeshire	CAM.txt	31,220	3,871
Cheshire	CHR.txt	66,585	8,922
Cornwall	COR.txt	60,745	7,465
Devonshire	DEV.txt	91,617	10,915
Durham	DUR.txt	46,126	6,142
Huntingdonshire	HNT.txt	10,099	1,395
Lancashire	LAN.txt	163,691	21,522
Leicestershire	LEI.txt	37,482	5,079
Middlesex	MID.txt	21,056	2,727

Northumberland	NHB.txt	43,131	6,239
Oxfordshire	OXF.txt	33,523	4,634
Rutland	RUT.txt	4,112	560
Westmorland	WML.txt	11,068	1,474
Worcestershire	WOR.txt	44,152	5,901
City and Ainsty of the City of York	YCA.txt	16,034	2,478
Yorkshire East Riding	YER.txt	49,736	6,460
Yorkshire North Riding	YNR.txt	58,796	7,859
Yorkshire West Riding	YWR.txt	201,663	29,844
Total		*976,168*	

Each file contains 17 fields as follows (the data type is indicated in brackets following the field name):

1 *runame (string)*

This is the name of the entity to which the baptism register belongs. This may be either an Anglican parish or an Anglican chapelry. For ease of reference these have been termed Registration Units. Each name used is unique at the county level but not necessarily so at national level.

2 *ruID (string)*

This is a unique identifier (at national level) for each registering unit. The three-letter prefix refers to the ancient county.

3 *por (string)*

This stands for 'period of occupational recording.' It takes one of three values: '1', '2' or 'rose.' Ideally we wanted one tabulation starting around 1750 (indicated by '1'), one starting around 1780 (indicated by '2') and one for 1813-20 (indicated by 'rose' – which alludes to the operation of Rose's act discussed above). In practice we took the nearest available set of years of good occupational recording and the exact dates are recorded in the next two fields. Bedfordshire is different and for this county alone por 1 runs from 1698 to 1749 while por 2 runs from 1750 to 1799.

4 *dstart (string)*

This records the date of the first observation within the period of occupational recording. This was recorded in the format 'dd mmm yyyy'.

5 *dend (string)*

This records the date of the last observation in the within the period of occupational recording. This was recorded in the format 'dd mmm yyyy'.

6 *xyears (string)*

This records any years in between *dend* and *dstart* that were not abstracted. Years were omitted where it was unlikely that their inclusion would permit the completeness rule to be met.

7 *sr (integer)*

Where Registration Units recorded over 2000 baptism in the period of occupational recording 'por' samples were normally taken. This field records

the sampling ratio ‘sr’. This is the ratio of the number of events in the register to the number of events recorded in the data file. In the vast majority of cases ‘sr’ is 1 because the register was not sampled.

8 *archive (string)*

This is the name of the archive which holds the baptism register or copy of the baptism register from which the data were abstracted.

9 *ref (string)*

This is the archive’s catalogue number for the baptism register or copy of the baptism register

10 *tsource (string)*

This field records the type of source. The following abbreviations were used:

Abbreviation	**Full Text**
BT	Bishops' Transcripts
electronic file	electronic file
MF Bt	microfilmed Bishops' Transcripts
MF OR	microfilmed Original Register
MF OR and OR	microfilmed Original Register and Original Register
Mfc OR	microfiched Original Register
Mfc Ts	microfiched Transcript
OR	Original Register
pc OR	photocopy of Original Register
pg OR	photographed Original Register
ptd	printed version of Original Register
Ts	transcript of Original Register
Ts BT	transcript of Bishops' Transcripts

11 *denter (string)*

This is the date on which the abstraction was performed.

12 *res (string)*

This refers to the initials of the research assistant who collected the data.

13 *ttaken (integer)*

This records the time taken, in minutes, to perform the abstraction. It does not include time taken to order or get hold of the source in the archive just the time to actually abstract the data.

14 *notes (string)*

This field contains notes that pertain to the period of occupational recording.

15 *year (integer)*

This field records the year in which the baptism took place.

16 *occ (string)*

This field records the occupation ascribed to the father. In the case of illegitimate children no father was normally recorded. Such cases can

normally be identified and in such cases 'illegitimate' has been entered in this field. No attempt was made to retain original spellings except in cases where it was not clear to the research assistant either what the correct spelling should be or what the occupation was.

Additionally, six non-occupational text strings can be found in this field, each beginning with 'aa'. The prefix 'aa' was used, in conjunction with the built-in 'Autocomplete' function in Microsoft Excel, to expedite data input.

aaillegitimate

This means that the baptism was judged to be illegitimate. In the overwhelming majority of cases where the birth was illegitimate, no occupation was recorded because no father was recorded. On the rare occasions where occupation was recorded we did not abstract it in order to maintain consistency in the data.

aaother parish - no occupation

Here, no occupation for the father was recorded at baptism but that the father was recorded as being resident in another parish.

aano occupation specified

Here, no occupation was recorded for the father at baptism.

aabaptism without named parents

This indicates that neither parents was named at the baptism. In consequences there was therefore no occupation for the father. In most cases this occurred because the individual being baptised was an adult. Where this individual was male an occupation was occasionally recorded, but for the sake of consistency we did not abstract such occupations. The total number of cases was very small.

aaunreadable

This means that the entry was unreadable.

The reason for recording these items was that we did not wish to use data where the occupational recording was less than 95 per cent complete. However, this was defined not as 95 per cent of all entries but as 95 per cent of those with an occupation or for which no occupation was specified ('aano occupation specified'). In other words entries where the baptism was illegitimate, or where no occupation was specified but the father came from another parish, or where no parents were named or where the entry was unreadable were excluded from the calculation. These particular codes were standardised so that the computer script that performed the quality control tests could readily identify these cases.

17 *com (string)*

This field contains comments that pertain to individual baptisms.