
RESEARCH NOTE

Local Population Studies using Parish Registers: Ideas for Future Research

Andrew Hinde and Paul Tomblin¹

Abstract

This short note discusses possible ideas for future research using parish register data and ways in which local and amateur historians might contribute to a new research agenda. In this, it is an attempt to resurrect and strengthen the links between amateur and professional historians that were integral to the work of the Cambridge Group for the History of Population and Social Structure in the 1960s and 1970s, and which led to the foundation of the journal Local Population Studies. The ideas discussed here are not fully formed, and should be seen as a contribution to a research agenda which is likely to be fluid, open-ended and responsive to initiatives from local and family historians.

E.A. Wrigley and R.S. Schofield's monumental work *The Population History of England, 1541–1871: a Reconstruction*, published in 1981, was dedicated to 'the local population historians of England'.² This was in recognition of their work in supplying the data on counts of baptisms, marriages and burials that underpinned the aggregate analysis of demographic processes on which Wrigley and Schofield's book was based. The journal *Local Population Studies* was founded in 1968, in part as a channel of communication between local, amateur and family historians who typically had detailed and expert knowledge of their local communities, and professional historians and demographers who were more often interested in analysing trends at the regional, national (and sometimes international) scales. *The Population History of England, 1541–1871: a Reconstruction* was based on data from 404 parishes out of the 11,000 or so parishes in England. These data have been rendered machine-readable and the Local Population Studies Society (LPSS) is responsible for making them publicly available.³

1 Andrew Hinde, Department of Social Statistics and Demography, University of Southampton, SOUTHAMPTON SO17 1BJ, Andrew.Hinde@soton.ac.uk. Paul Tomblin, 5 England Crescent, Heanor, Derbyshire DE75 7BE, paultomblin@hotmail.com.

2 E.A. Wrigley and R.S. Schofield, *The Population History of England, 1541–1871: a Reconstruction* (London, 1981), p. v.

3 The data are available in Excel files and in .txt files. They are organised by parish, and grouped into counties. There is also a set of .txt files describing the characteristics of each parish using information from the 1811 and 1831 censuses and other sources relating to earlier centuries. Originally, the dissemination was organised using CD-ROMs containing the data, which were accompanied by a guide written by Roger Schofield: R. Schofield, *Parish Register Aggregate Analyses* (Colchester, 1998). A nominal charge of £5 was made for the CD-ROM and the guide. Dissemination in this format was discontinued some years ago. At present, copies can be (and have recently been) supplied as compressed (zipped) files sent by e-mail. The uncompressed files take up 56.0 Mb (Excel), 24.3 Mb (.txt), and 1.6 Mb (Characteristics), a total of 81.9 Mb. When compressed they occupy 10.6 Mb (Excel), 4.1 Mb (.txt) and 0.5 Mb (Characteristics). They may be obtained in this form from Andrew Hinde at the e-mail address given in footnote 1.

Given this responsibility to disseminate the data for the 404 parishes, LPSS is wondering whether a wider project involving the collation and distribution of parish register data could be built upon this. This paper is an attempt to consider what a more ambitious ‘parish register project’ might look like. In thinking about this, LPSS is acutely aware that it was originally conceived as a society of amateur and professional historians working together. Any project which builds on the data for the 404 parishes should promote this foundational aim of the Society. Accordingly, this paper is a collaboration between a ‘professional’ historian with a background in demography and statistics and an ‘amateur’, with a background in family history.

This paper is not designed to be a worked-out description of a proposed project. Our aim is to provoke discussion. One of our conclusions is that any project based on parish register data will need to be flexible and responsive to the needs of local and amateur historians.

In the next section we describe the availability of parish register data in 2019. We then consider ways in which we might re-invigorate the links between local and family historians, on the one hand, and professional academic historians on the other. Finally, we propose some ideas that a ‘parish register project’ co-ordinated by the Local Population Studies Society might take forward.

Parish register data in 2019

The data for the 404 parishes used in *The Population History of England, 1541–1871: a Reconstruction* were assembled more than 40 years ago. They took the form of counts of baptisms, marriages and burials. They were, of course, derivative in that they were summaries of the information in the original parish registers. During the last 40 years, there have been many developments in the availability of parish register data, and it will be helpful to summarise these and, by doing so, briefly describe the parish register data landscape in 2019.

Parish register data come in a variety of forms.

- (1) The original parish registers are handwritten documents recording the baptisms, marriages and burials in each parish of the Church of England. The earliest of them date back to 1538, but in most parishes they started somewhat later than that. Until 1753 for marriages, and 1812 for baptisms and burials no standard format was required for the records, and they were written in a variety of ways. Most of these original documents survive, and are stored in local archives. These days, researchers only rarely need to consult the originals, as local archives have photocopied them and made them available on microfilm, microfiche or online.
- (2) Over the years many local and family historians have transcribed the parish registers and published them using more modern media. By ‘transcribed’ we mean essentially that the information in the register has been copied verbatim. Some transcriptions have

been published as part of series by local history societies.⁴ Others have not been formally published, but have simply been typed or handwritten and placed on the shelves of local archives. It is almost impossible to know how many such transcriptions there are, but in some counties (for example Suffolk or Nottinghamshire) almost all the registers have been transcribed.⁵ The quality of these transcriptions is somewhat variable, but many of them—and probably most of them are of high quality.⁶

- (3) Some of the transcriptions, but at present probably only a minority, have been input into computer files or databases such as Excel or Access. An advantage of complete transcripts is that they allow users to analyse the data in any way they wish. Disadvantages are that they may require extensive processing before interesting results can be obtained, and that different users may process them in different ways, so that comparability of analyses across parishes is compromised. Related to this, it seems that no consistent format for these files has been followed. In a recent project to assemble as many data base files from burial registers as possible, the Cambridge Group for the History of Population and Social structure asked family history societies to let them have copies of the machine-readable transcripts in their possession. Some 19 county-based societies agreed, but the resulting files were in a wide variety of formats:

[d]ata values for relationships were ... more varied than it was possible to predict from one batch to the next. Some inputs had more than one date field, or several age fields expressed in different ways, or in different states of standardisation. ... Field names were not a reliable guide to field contents, nor were field names standard or predictable, both within batches and from one batch to the next.⁷

- (4) From the original registers, or from transcripts, summary data files have been derived such as the count data for the 404 parishes described above. These are generally in machine-readable format, such as Excel files of the 404 parishes. They consist of monthly counts of baptisms, marriages and burials, usually from the date a parish register began, until (usually) 1812 or 1837, when civil registration started. An extract is

4 See, for example, the Yorkshire Parish Registers Series published by the Yorkshire Archaeological and Historical Society, which now run to more than 180 transcriptions of the parish registers within that fair county. <https://www.yas.org.uk/Publications/Yorkshire-Parish-Registers-Series> [accessed 29 November 2019].

5 Just to take one county as an example, the Nottinghamshire Family History Society explicitly aims to transcribe the registers of every parish in their county. The Thoroton Society has, over the years, also transcribed many Nottinghamshire registers and placed them on the shelves of the local Record Office.

6 One of us (Tomblin) has made transcriptions for several family history societies. At a minimum, the work involved one person's register transcription being checked by someone else and then any differences being then checked by a third person. Some databases of transcriptions take this a stage further in that, if a user believes that a transcription is wrong, it can be reported and the original document is then consulted.

7 G. Newton, *Data Mining Family History Society Burials*, Cambridge Working Papers in Economic and Social History no. 34 (Cambridge, n.d.). It should be said that individual county family history societies often try to ensure that all their 'own' transcript files are in the same format. So there is consistency within counties, but not between counties.

Table 1 Monthly counts of baptisms in the Hampshire parish of Odiham

Year	Jan	Feb	Mar	Apr	May	Jun	Jul	Aug	Sep	Oct	Nov	Dec	Total
1538	5	4	5	4	2	0	1	2	5	4	6	2	40
1539	4	7	8	2	2	1	8	3	2	2	2	4	45
1540	8	3	7	2	4	8	4	2	4	3	8	1	54
1541	2	2	10	4	2	5	2	1	2	3	3	4	40
1542	7	4	6	8	2	2	4	4	3	8	5	2	55
1543	6	4	0	5	3	4	4	6	1	2	2	5	42
1544	5	4	7	3	3	3	4	0	1	5	0	3	38
1545	5	3	3	4	1	3	4	4	6	1	4	4	42
1546	4	6	4	4	6	2	3	5	8	3	3	2	50
1547	2	5	4	6	2	2	7	8	4	4	8	2	54
1548	4	6	3	8	4	4	4	7	1	5	7	4	57
1549	5	4	10	3	2	3	3	4	5	3	6	4	52
1550	5	2	7	5	1	3	2	4	3	1	6	2	41
1551	5	2	3	8	6	1	1	4	6	6	7	3	52
1552	4	7	2	1	2	1	3	2	1	3	4	2	32
1553	3	9	3	3	5	1	2	4	5	3	3	4	45
1554	2	7	5	3	3	5	4	6	6	2	5	4	52
1555	3	7	4	4	3	4	3	2	4	4	4	3	45
1556	3	4	3	1	4	3	5	3	2	5	2	2	37
1557	4	9	2	0	5	2	1	1	0	3	1	1	29
1558	4	1	4	6	2	2	1	2	4	1	4	5	36
1559	4	3	4	2	0	3	1	4	1	4	4	0	30
1560	0	3	1	6	4	16	3	3	2	1	5	2	46
1561	4	3	5	3	3	1	3	0	3	3	5	5	38
1562	3	2	3	3	1	4	5	3	5	4	1	5	39
1563	4	4	2	1	4	2	4	3	3	4	2	3	36
1564	3	2	3	3	2	0	1	6	3	4	4	4	35
1565	6	6	1	6	5	0	5	0	3	4	5	5	46
1566	5	5	1	4	3	0	2	1	2	2	0	3	28
1567	4	2	3	4	5	1	9	2	3	4	5	0	42

given in Table 1 for baptisms in the Hampshire parish of Odiham from 1538 to 1567. Data made available in this form are derivative in that they abstract from the original lists of baptisms, marriages and burials, and the analysis that can be done on the basis of this kind of data is limited, but they have the advantage that comparative studies of many parishes are practicable. It is also relatively straightforward to carry out quality checks on the count data, and the kinds of quality checks that are needed are well known.⁸

- (5) In some cases, transcripts have been input into Excel files or Access databases that include most, or perhaps, all of the information in the original parish registers, but which have re-organised and augmented that information so that it is more amenable to the kind of analysis that population historians would like to undertake. By this we mean that some processing has been carried out, perhaps to link records within the same file, or to link marriages to baptisms, or baptisms to infant deaths. The most well-known examples are those parishes for which family reconstitution has been carried out, but there are many others for which some processing has been done short of a full family reconstitution.⁹ The advantages data in this format are, first, that certain operations can be standardised across files, which increases the comparability of analyses; and, second, that the process of carrying out basic re-organisation and augmentation of the data can be bound up with a quality check, and should thus reveal problems with the quality of the transcriptions, or of the original registers. The disadvantage is that the processing will be time-consuming and will slow down the dissemination of the data.¹⁰

Linking local and family historians and academic research

Research into population history is broadly organised according to two models and is carried out by both amateur and professional historians. The two models are, first, large-scale projects with defined aims and objectives and which, when carried out by professional historians, are usually publicly funded; and, second, research by individuals, whether amateur or professional, often on specific local areas, or answering particular questions.

8 On the quality of parish register data and the checks that can be done to ensure data quality, see Wrigley and Schofield, *Population History of England*, pp. 15–154.

9 On family reconstitution, see G. Newton, 'Recent developments in making family reconstitutions', *Local Population Studies*, 87 (2011), pp. 84–9. Family reconstitution, if fully carried out, is a very time-consuming technique. So far it has only been applied to a limited number of parishes, most famously the 26 parishes analysed in E.A. Wrigley, R.S. Davies, J.E. Oeppen and R.S. Schofield, *English Population History from Family Reconstitution 1581–1837* (Cambridge, 1997). Other parishes have been subjected to family reconstitution, but the total number is unknown as some of the attempts at reconstitution have not resulted in publications.

10 Even at the most optimistic assessment, it is unlikely that the number of parishes in which the population has been fully reconstituted for a reasonable length of time (say 50 years or more) numbers much over 100. Chris Galley (see C. Galley, 'Infant mortality in England, 1538–2000: the parish register period, 1538–1837', *Local Population Studies*, 103 (2019), pp. <<>> describes analyses of infant mortality which have employed various methods of analysis of parish register data short of full family reconstitution.

Figure 1 Modes of research in population history

	Amateur	Professional
Individual research project	Study of population change in a specific parish or group of parishes (e.g. M. Bolton, 'The experience of plague in East Kent, 1636-1638', <i>Local Population Studies</i> , 96 (2016), pp. 9-27.	Study of specific topic, perhaps using data from a particular geographical locality or several localities (e.g. P. Razzell, C. Spence and M. Woollard, 'The evaluation of Bedfordshire burial registration', 1538-1851, <i>Local Population Studies</i> , 84 (2010), pp. 31-54.
Large-scale research project	Collective project organised by a county family history society (e.g. projects to digitise or make available transcripts of registers from every parish in a county)	Major publicly-funded research project (e.g. Cambridge Group for the History of Population and Social Structure's project on 'Migration, mortality and medicalisation: investigating the long-run epidemiological consequences of urbanisation 1600-1945')

Many papers published in *Local Population Studies* report the results of work of this type. It is important to recognise that both amateur and professional historians can participate in work of both types (Figure 1).

The uniqueness of the original parish register project of Wrigley and Schofield was that it involved both amateur (local and family) historians and professional historians working together on a large-scale project with specific aims. That said, the relationship was not one of equals: the time of the local or amateur historians was used instrumentally to provide data: the aggregate analysis of those data was carried out by the professional team.¹¹ Nevertheless, it probably was the high-water mark of collaboration between amateurs and professionals. Since then, the large-scale projects of professional historians have sometimes made use of data gathered by local historians but direct engagement has been lacking. For example, the current Cambridge Group for the History of Population and Social Structure (CAMPOP) project entitled 'Migration, mortality and medicalisation: investigating the long-run epidemiological consequences of urbanisation 1600–1945' uses parish register data assembled by country family history societies. Of course, these data were originally transcribed by individual local historians but the CAMPOP team are not engaging with

11 See J.D. Marshall, *The Tyranny of the Discrete: a Discussion of the Problems of Local History in England*, pp. 35–6: '[t]he Cambridge organizers ... commenced with excellent intentions ... [b]ut it soon became clear that despite their large-scale and exciting use of local data, they were primarily and understandably concerned with questions of national and international demography; the local material was used instrumentally'.

these individuals directly, but with the family history societies.¹² The data they are using are in format (3) listed above, and CAMPOP have written elaborate and ingenious computer code to extract consistent information from the multiplicity of formats in which they have received the data files. It is not clear, though, that the result of this will be a set of parish data that is freely available, let alone useful, to amateur population historians.¹³

Individual academics and smaller projects of professional historians continue to use parish register data, of course, and many of them are prepared to make their data available—at least to other members of the academic community. Quite often, though, academic historians have only collected the data they need to answer their specific research questions. They are not interested in providing a general data collection and transcription service for other historians, amateur or professional.

There is also a feeling among amateur and family historians that there is an enormous gulf between what they are doing and the research that academic historians carry out. Part of this seems to involve a belief that academic population history necessarily involves quantification and ‘statistics’. This is a misconception, but it is one which the academic historical community could do more to dispel. Harder to dispel—because it is nearer to the truth—is the impression that population history as practiced by academics has become increasingly organised as a set of large projects using sophisticated and heavy-duty computing resources.¹⁴

A parish register project

The foregoing comments suggest that future research on parish register data would benefit from a project designed to make easily available in a central location a set of data which could comprise both aggregate ‘count’ data of the form in which the data for the 404 parishes has already been supplied, and data in form (5) above for parishes for which individual local historians have collected data for their own purposes, and who are willing to make their data publicly available.

The project would, as a minimum, make available for download the data for the 404 parishes used in *The Population History of England, 1541–1871*, together with the contextual information about the parishes and a guide to the use of the data in the form of an updated edition of Roger Schofield’s *Parish Register Aggregate Analyses*. However, it is known that

12 See Newton, *Data Mining*.

13 Despite the instrumentality that Marshall complained about, as a collectivity, the local historians of England all benefitted from access to the data for 404 parishes in exchange for the work they put in to aggregate the totals for the parish(es) on which each worked. To be sure, anyone could gain access to these data, so the local historians of England performed a public service as well as a service to the academic historical community. With the current CAMPOP project it is not clear that anyone will have access to the cleaned and checked data other than members of the research team. One reason for this is that the family history societies are acting as gatekeepers and have imposed conditions on exactly what elements of the data CAMPOP can use and analyse.

14 It may be that family historians are interpreting the use of intensive computing power as the use of quantification and ‘statistics’.

count data exist for well over 100 additional parishes, as 542 parishes were used by Ann Kussmaul in her book *A General View of the Rural Economy of England, 1538–1840*.¹⁵ There is also a large collection of parish register data available in county record offices, on the web sites of organisations like FindMyPast, the Genealogist and Ancestry. It is not always easy to extract from these data information in the format of the data for the 404 parishes without looking at transcripts. For example, FindMyPast will allow the number of burials in each parish in each year to be extracted fairly easily without looking at transcripts, but not the number in each *month*. However, where transcripts exist it will not take too long to extract counts. The project would check these counts for data quality, including both accuracy and completeness, and where the data passed the tests, add them to the data base using files of the same format as those already in existence for the 404 parishes. This process would be continuous, in that count data for additional parishes could be added as they become available. Roger Schofield's booklet *Parish Register Aggregate Analyses* could also be updated with details of the additional parishes.

The second part of the project might involve making available data in form (5) above for a smaller number of parishes. These data files would, essentially, be complete transcripts of parish registers including details of individuals' names, but they would be presented (probably in Excel files) in a standardised format, with a consistent set of columns. It would be a challenge at an early stage of the project to decide upon a suitable standardised format. Fortunately, the ability to sort and re-arrange cases in Excel means that there is considerable flexibility about the precise sets of records to be include in a particular worksheet (for example, whether we have separate worksheets for baptisms, marriages and burials). Excel is much less flexible when it comes to re-arranging columns of data, though, which means that some thought should be given to making the column structure of the data as helpful as possible to users. A further question is how much the project should do by way of making links between records in the data files we post online.

Beyond that, the development of the project would depend on the needs and aspirations of those involved. In particular, the Local Population Studies Society (LPSS) would seek a dialogue with family history and local history societies with the aim of involving them in determining the direction of the project.¹⁶ The vision is of a project quite different from the large-scale academic projects with defined objectives, a nominated research team, and a specific time-frame. It is of an open-ended project managed by LPSS, but without particular objectives save to create a repository for parish register data in the form of both counts of events, and full transcripts, presented in a common form and with a guarantee that they have been rigorously checked for data quality. It would explicitly involve local and amateur historians in supplying the data, in re-arranging and augmenting them and in performing the quality checks.

15 A. Kussmaul, *A General View of the Rural Economy of England, 1538–1840* (Cambridge, 1993).

16 Such a dialogue will be essential if the project is to use data assembled by family history societies, as they have a commercial interest in these data.

Local Population Studies using Parish Registers

The project would not, initially, have a substantive research agenda, but would furnish both professional and amateur historians with data through which a substantive agenda of the kind set out by Chris Galley in his recent article on infant mortality could be pursued.¹⁷ In the future it is possible that research outputs and research reports from both amateur and professional historians using the data in the repository could be added to the web site.

17 Galley, 'Infant mortality in England, 1538–2000: the parish register period', pp. <<>>