

## 4491 - Parish Register Aggregate Analyses

REPORT TO ESRC ON 404 DATA

AWARD RES-479-26-4003

AIMS: To reformat and aggregate the whole of the 404 Parish Register Dataset into a national integrated version of the parish event totals.

This project was designed to make the original data which informed the *Population History of England and Wales, 1538–1838* available in a format which was more analytically friendly.

METHOD:

The process was to transform 5,252 worksheets into a three flat-file tables. Each of the four files will be available to users as tab-delimited text files, after re-cataloguing by the AHDS History. Re-cataloguing will not take place until the ESRC have accepted this report, and that these data can be distributed in their current format.

### Events data

The main data, now held in a single table: *events.tab* was rapidly achieved using complex macros implemented within a text editor.

### Mortality crisis data

Data specifically relating to mortality crises (*crises.tab*) were similarly repurposed into a single yet separate table.

### Indexed vital events data

The third main table to be created (*eventsidx.tab*), that which gives 50-year index numbers for vital events, proved most troublesome.

In the original deposit these data had been corrupted, and attempts to obtain correct values from the original depositor have proven fruitless. As an example, the original data for the 50 year index

Years	Jan	Feb	Mar	Apr	May	Jun	Jul	Aug	Sep	Oct	Nov	Dec
1500s	0	0	<b>0 0</b>	0	0	0	0	0	0	0	0	0
1550s	0	0	<b>1218</b>	0	0	0	0	0	0	0	0	0
1600s	64	140	<b>32 99</b>	223	165	32	96	132	127	33	64	
1650s	72	158	<b>57 119</b>	115	119	86	57	74	216	89	43	
1700s	36	78	<b>36 111</b>	214	111	36	36	221	179	0	143	
1750s	45	124	<b>23 211</b>	181	94	113	45	70	68	94	136	
1800s	67	24	<b>89 184</b>	133	115	22	89	23	156	92	200	

It will be clear that for some reason the two data series have been concatenated within a single variable. For all years with the exception of the 1550s we have simply broken these concatenated variable at the space. The remaining figure is a curiosity. It is impossible to

break the figure, and even so, the result would in any case, look implausible. (Furthermore, for this parish, Gatton, Surrey, there are no recorded marriages in the 1550s so the figure here must be erroneous.) *Without recourse to the original data, which is not in our possession, we strongly recommend that this table is not to be used or at least with considerable care.*

Specifically, the problems can be found in:

Easingwold, Yorks NR, Marriage index (1550s);  
Torver, Lancashire, Marriage index (1550s);  
Meonstoke, Hampshire, Burial index (1550s)  
Gatton, Surrey, Marriage index (1550s).

Some formatting decisions were made which may impinge on the manner in which the data can be used. For example, in the original data, periods for which no events were recorded were denoted in the file with an asterisk which covered all months in that year. These were consistently recorded as 'NULL' in the finalized database, and have subsequently been transformed within SPSS.

### **Parish characteristics**

The material which has been subjected to the greatest work is the characteristics table. All material which had previously been in textual format has been coded. Documentation has been provided which describes the alterations made in this process, but some original interpretational material remains elusive. Schofield's table 1 in his *Parish Register Aggregate Analyses: the Population History of England database and introductory guide*, 7 is not fully detailed to describe each variable in depth.

The other important addition to this table is the inclusion of parish identifiers which allow users to easily locate the parish using Kain and Oliver's *Historic Parishes of England and Wales*.

### **Remarks**

The essential events table and the less essential crises tables have been created with little difficulty. The parish characteristics table has been fully converted from its original textual format into a format which allows the characteristics to be linked in a relational database or statistical processing package with relative ease.

The original data deposited by Local Population Studies at the AHDS History was missing the characteristics of a single parish; details which were supplied in paper format by the Cambridge Group rapidly. However, when attempting to retrieve from the Cambridge Group further information relating to the other problems mentioned above we met with a silence. A number of attempts were made to secure the original dataset from the Cambridge Group but these were also unsuccessful. It is for this reason, and this reason alone that there has been a delay in making this report available to the ESRC.

Matthew Woollard  
30 January 2004  
revised 1 April 2005