THE ANALYSIS OF CENSUS-TYPE DOCUMENTS

Valerie Smith

Mrs. Smith was research assistant to the Cambridge Group for the History of Population and Social Structure from 1966 to 1968, specializing in the analysis of early listings of inhabitants. She is now working with Dr. T. Aston on the History of the University of Oxford.

Local population studies usually centre on the parish register, which gives a running record of vital events. The register, however, cannot tell us either how many people were alive in the parish at any one time or how individuals lived together in families or household groups. To answer questions of this kind, for example, whether old people usually lived with their sons and daughters or in lonely isolation, we need a census of the population.

The first official census in England was taken in 1801. Before that date census-type lists were made principally by individuals for a specific purpose. The major exception to this is a collection of lists for the years 1695, 1696 and 1697, which were made as a direct result of an act of parliament imposing a levy on baptisms, marriages and burials and on bachelors and widowers. This act was an attempt to raise money to further the war against France. From the beginning, the act was difficult to administer, and insufficient thought seems to have been given to the machinery necessary for its enforcement; nevertheless, a considerable number of lists are extant and these comprise the largest single group of lists from pre-census times. (1) Although it is unlikely that new lists from this source will be found, since exhaustive enquiries for them have already been made, it is always possible that random lists drawn up by individuals for special purposes will be discovered.

There appear to have been many reasons for the drawing up of these so-called random lists. The majority of them seem to have been compiled by vicars, and all of them must have been compiled by members of the literate minority. It seems often to have been the case that a parish priest, being more than commonly conscientious, felt that it was important for the discharging of his duties to have some record of the souls in his care, and therefore to have listed his
parishioners. An example of such a list is that found in the rector's book for the parish of Clayworth (2) In England such lists as we have of this kind were the result of individual initiative and not of a general order, as was often the case in France. Occasionally a more specific reason for compiling a list is apparent, varying from precautionary measures in case of a feared French invasion of Ardleigh, in Essex, in 1790 to an attempt to ensure that all parishioners were provided with a Bible and Prayer Book.

The quantity of information given in such lists varies as greatly as the reasons for their compilation, though the two factors are not related in any obvious way. The quality is influenced considerably by the method of compilation used. Obviously, there was no question of sending forms to each householder to be filled in, since both the administrative complexity of such an operation and the high level of general illiteracy made this impracticable. A person wishing to list all the inhabitants of a particular village had three possible courses of action open to him. He could rely on his memory and compile the list without stirring from his study; he could assemble all the inhabitants in one place and question them; or he could go round from house to house asking questions. Neither of the first two methods could be expected to give particularly reliable results. If he were to rely on memory he would have no way of checking the accuracy of his list, and would probably either forget some households completely or make mistakes in the detailed composition of households, particularly as regards servants since they were very mobile. To assemble everyone would be difficult and time consuming, and would perhaps give rise to unhelpful resentment, if not to deliberate absenteeism, moreover the sick and the infirm would probably be omitted. The third method, that of house to house questioning would be the most promising. Its most serious drawback would be that it is most unlikely that the operation could be completed in one day, and in a large parish it might take several weeks. This means that the resulting list is not, strictly speaking, a census, since change would have taken place between the beginning and the end of its writing. A list drawn up by this last method is, however, most likely to be of value to the historian, since it is less likely to be biased in its omissions. A list drawn up by the recall method may be biased by including all the prominent households which would be brought to mind easily, and concentrating its exclusions among the obscure. On the other hand, a list which
results from assembling the inhabitants may be biased in favour of those the compiler felt free to command, and may omit the families of high social status. It must be borne in mind that omission of particular households may be deliberate. For reasons which are not clear, some lists deliberately excluded the gentry or gave fewer details about their households. A list of inhabitants of Puddletown, Dorset, omits both the Manor House and the Vicarage; whilst that of Cardington, Bedfordshire, gives minute details of the cottage households but only the most sketchy information about more substantial households. This list has been well analysed by Tranter in Population Studies, and gives a good indication of some of the information which can be recovered from a list of high quality.

One of the most difficult of all problems confronting the historian wishing to analyse a pre-census list is that of deciding whether the list is real or ideal in construction. All the censuses proper are 'real'. A real list includes everyone present when it was taken and omits everyone not present irrespective of whether they were normally resident in that particular place or not. Thus a real list would exclude a servant who was away at market on the day in question and would include a person visiting a relative in the village who normally lived elsewhere. An ideal census describes the population as it normally is; that is, it omits temporary visitors and includes residents who were temporarily absent. It is not usually possible to decide which class a given list falls in absolutely certainly, one can only say which seems more likely. The problem is particularly important if a comparison with a true census is desired, or a comparison between two lists the nature of one of which is known beyond doubt, but it should not be ignored when only one isolated list is being considered.

Some types of list do not even pretend to include the whole population. Of these, the communicants lists, and particularly the Compton census, are important. As their name implies, these list communicant members of the population only. It is just about possible to treat these as modified census type documents for some purposes, but in order to do so two questions have to be answered: one is the age at which a person was normally admitted to Communion. The other is the proportion of persons above this age to those below to be expected in the population as a whole. Since both these questions are complicated, particularly the latter, it is wiser to treat communicant lists like hearth tax returns as a class apart and a specialist study in their own right. (3)
True census type listings are of less value for most purposes if they are known to be, or strongly suspected of being, incomplete, since an incomplete list is not representative. A complete list which does not contain information of high quality can, nevertheless, be of value if used in comparison with other lists. An incomplete list is not capable of comparison with other lists; it follows that it will repay study only if the information it gives is in itself of high quality.

Pre-census lists are rare and their survival is due to accident. It is therefore not possible to think in terms of representative samples (or any type of samples at all), nor is it possible to be over fussy in deciding how good a list must be before it can be used. Even so, it is important that detectable bias be avoided. An incomplete list cannot but be biased if its incompleteness is due to deliberate omission. If it is used, its bias should be clearly declared and its uselessness for comparative purposes acknowledged.

The quality of a list is judged by the amount of information which it is capable of yielding. In general, the more information a list gives the more rewarding it is to work. Since there is immense variation in the style and content of each individual list, it is not possible to classify all types of information exhaustively. A few general principles can, however, be stated. Since a list of inhabitants is the only type of document which gives a picture of the social organisation of a particular community at a particular time, all judgement and all analysis should be in terms of the advantages and limitations of this particular quality. In particular, a list should not be judged by its ability to yield such dynamic information as can best be derived from other sources.

The first major division of types comes between those lists which give information in household units and those which do not. It is never possible to do any work on the structure of the household unless each household is unambiguously identified as a unit. It is very tempting to believe that household divisions can be made by the researcher, and when the list is otherwise promising the temptation becomes almost irresistible. Nevertheless, it must be resisted. To insert household divisions invalidates all work done on the structure of the household since it introduces a circular argument. The following is an (invented) example to illustrate this:
John Marshall
Elizabeth Marshall
Susan Marshall
James Marshall
William Marshall
Ann Jenkins
Robert Scott
Richard Jones
Mary Jones
Rebecca Jones

Such a situation is quite possible on a list which does not give household divisions. In attempting to insert such divisions, it is necessary to have recourse to probability. It is probable that all the Marshalls belong to one household and that a new household begins with Richard Jones. It is, however, only probable. It is by no means impossible that John Marshall is living alone and that Elizabeth Marshall is the widowed head of a new household. The two names, Ann Jenkins and Robert Scott, do not obviously belong anywhere. There are many possibilities e.g. they are both servants of John Marshall, they both live alone, Ann Jenkins is Elizabeth Marshall's Mother who lives with her son-in-law and Robert Scott is either a servant or a solitary or even head of a household in which the Jones family are lodgers. These are only some of the possibilities.

It is theoretically possible to work out a statistical balance of probabilities to assist in placing household divisions. In practice, it is doubtful if there is a sufficient body of data available from such lists as have been analysed to make this feasible. Even if it were possible, it would not be possible to use these imposed household divisions to support or contest any theory without circularity. If an analysis on the usual lines, that is an analysis which aims to discover the social structure of the community by using as its basic unit the household and its organisation, is envisaged, these lists must be left severely alone. They can, however, yield a certain amount of very restricted information: they can reveal the sex balance of the community and occasionally, very rarely, the age structure.

When household divisions are clearly given, the next major dividing line comes between those lists which name every individual and those that do not. It is quite common for a list to give only the name of the householder and the number of persons living in the house, sometimes subdivided into various categories. Such a list may read as follows.
<table>
<thead>
<tr>
<th>Males</th>
<th>Females</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>Henry Coles</td>
<td>2</td>
<td>3</td>
</tr>
<tr>
<td>Emma Finch</td>
<td>1</td>
<td>2</td>
</tr>
<tr>
<td>John Roberts</td>
<td>1</td>
<td>0</td>
</tr>
</tbody>
</table>

A list such as this will yield household size, sex balance, and in some cases, occupational detail. It is possible for a list to give considerably more information than this without naming individuals.

The most detailed lists are always those which name every individual, and these are more interesting in their own right apart from their comparative value. Unfortunately, they also seem to have been less common than the more abbreviated form.

One further important distinction must be recognised. Some lists will state the relationship of each member of the household to the head, whilst others will not. The first type, which may be of the form

<table>
<thead>
<tr>
<th>Peter Cooper</th>
<th>Husbandman</th>
</tr>
</thead>
<tbody>
<tr>
<td>Ann Cooper</td>
<td>his Wife</td>
</tr>
<tr>
<td>Mary Cooper</td>
<td></td>
</tr>
<tr>
<td>Jane Cooper</td>
<td>children</td>
</tr>
<tr>
<td>William Cooper</td>
<td></td>
</tr>
<tr>
<td>Thomas Masters</td>
<td>nephew</td>
</tr>
<tr>
<td>Ellen Rhodes</td>
<td>servant maid</td>
</tr>
<tr>
<td>James Andrews</td>
<td>servant</td>
</tr>
</tbody>
</table>

poses no problems and can be analysed in detail easily. The case where the relationship is not stated is more complicated. In order to exploit the full potential of a list which is of the type

<table>
<thead>
<tr>
<th>William Simmons</th>
<th>Blacksmith</th>
</tr>
</thead>
<tbody>
<tr>
<td>Jane Simmons</td>
<td></td>
</tr>
<tr>
<td>John Simmons</td>
<td></td>
</tr>
<tr>
<td>Jane Simmons</td>
<td></td>
</tr>
<tr>
<td>Edward Simmons</td>
<td></td>
</tr>
<tr>
<td>Peter Smith</td>
<td></td>
</tr>
<tr>
<td>Elizabeth Matthews</td>
<td></td>
</tr>
</tbody>
</table>

certain presumptions must be made. A presumption is not the same thing as a guess and it should only be made when it has the status of a near certainty. In the case above, it is possible to presume that William and Jane Simmons are married and that they have three
children living with them. It is mere speculation to suggest that Peter Smith is a relation of William Simmons. In such a case the use of a parish register will help enormously in removing persons from the category of 'unknown' and in confirming or disproving previously made presumptions.

3

In the rare event of the register being of sufficiently high quality, a full reconstitution should be considered. This however should not be undertaken lightly, since it is an immensely time consuming process. The register should be tested stringently for its suitability for reconstitution (4). Since both lists of inhabitants and registers good enough for family reconstitution are rare in England, the chances of the two coinciding are minimal. If the register is not suitable for reconstitution this does not mean that it cannot yield valuable information in confirming or denying presumptions.

If an inferior register is to be used in this way the researcher will probably find it necessary to make out his own transcript of the list, onto which additional information can be entered and which can then be used as a working copy. Double sheets of lined foolscap are most convenient for this. At this stage it is helpful to give each person a number for ease of cross reference. This should consist of the number of his household on the list and the number of his position in that household. For example, the person whose name came third in the list of members of the twelfth household from the beginning of the list would be numbered 12:3.

When the list has been transcribed, the next stage is to make out a separate card for each person on the list (5 x 3 index cards are recommended). These cards should contain the person's number and name and any other information stated on the list. Presumed information should not be entered. Typical cards might read:

8:1 James Green
     Weaver

8:2 Margaret Green

Even if it appears obvious that Margaret is the wife of James she is not so described on her card.

The dual process of transcription should not be neglected: contrary
to appearance it does in the long run save time. It is most valuable to have a working copy which can be written on freely. The cards serve a dual purpose. They are used for consulting the register, and later they can be sorted into groups and counted. This tends to greater accuracy than scanning the list to record all the members of a particular class such as widows. The gain in accuracy can be so great in the case of a large (over 1,500) list that the compilation of cards should be considered even if there is no question of consulting the register.

In consulting the register the procedure is first to arrange the cards in alphabetical order and then to work back through the register from the date of the list checking each entry against the cards. If the name is identical, the information is copied (in full) from the register on to the card. At this stage no attempt is made to resolve ambiguities of identification, therefore, particularly if the name is a common one, each card may carry more than one possible date of baptism.

It is usual to work back for a period of about seventy years from the date of the list. It is always the case that far more information is recovered from the years immediately before the list than for the earlier years. This is partly the result of migration and also partly the result of the fact that women change their name on marriage, which means that it is impossible to relate a baptism entry to a married or widowed woman unless the marriage is also recorded in the register. It is also true that in many cases the earlier entries are more ambiguous in their application. This being so, it is rarely possible to recover from a parish register sufficient detail about the age structure of a community to make this analysable, since the chance of discovering an age is so much greater for the young than for the old.

When all the information from a register has been copied onto the index cards, the next stage is to resolve as many ambiguities as possible and to discard information which remains ambiguous. An example will best illustrate how this is done. This example has been contrived so as to illustrate as much as possible: it is most rare for so much information to be recovered for any one household in practice.

Suppose that one household in a list made in 1770 contained the following members:
8:1 James Green Weaver
8:2 Margaret Green
8:3 James Green
8:4 Thomasine Green
8:5 Mary Anderson
8:6 John Cartwright

the first working copy might read

8:1 James Green Weaver
8:2 Margaret Green Presumed Wife
8:3 James Green Presumed child
8:4 Thomasine Green Presumed child
8:5 Mary Anderson Possibly servant or lodger
8:6 John Cartwright Possibly servant, apprentice or lodger

after the register had been consulted the cards might read

8:1 James Green Weaver
married 6.8.1766 Margaret Anderson
married 21.4.1759 Jane Smith
Baptised 5.3.1735 son of John and Susan
Baptised 3.9.1732 son of Henry and Mary

8:2 Margaret Green
married 6.8.1766 James Green nee Anderson
Baptised 2.2.1740 d. of Charles and Mary
Baptised 8.5.1715 d. of James and Mary

8:3 James Green
Baptised 9.12.1763 son of James

8:4 Thomasine Green
Baptised 12.10.1769 d. of James
buried 7.11.1768 d. of James and Margaret
baptised 24.5.1767 d. of James

8:5 Mary Anderson
Husband William buried 1.6.1767

8:6 John Cartwright
Baptised 17.11.1755 Son of Robert and Susan

It is clear that James Green has married twice. Although no burial
is recorded for Jane it is certain that she was dead by 1766, since
divorce was unknown. James Green junior must be Jane's son.
The entries for Thomasine make it clear that, in fact, two children
are here involved. A child Thomasine was born in 1767 and
died in 1768, another daughter was born in 1769 and given the same
name. This second Thomasine is the child entered on the list, the
daughter of James and Margaret. Mary Anderson is very probably
a relative of Margaret Green, whose maiden name was Anderson.
The strongest presumption is that she is her mother. The fact that
she is known to be a widow strengthens the probability that she would
be living in her son-in-law's household. Of John Cartwright it is
only known that there are no grounds for presuming him to be
related to the family. This leaves the original presumption unaltered.
His age strengthens the probability of his being a servant or
apprentice. The question of the date of baptism of James Green is
not capable of resolution, both possible dates must therefore be
discarded. The baptism of Margaret can be dated. If the first
date is correct she would be thirty years old at the time of the list
and twenty nine years old at the birth of her daughter Thomasine.
If the second date is correct she would be fifty five years old at the
date of the list, and fifty four at Thomasine's birth. On
physiological grounds, therefore, the second date can be rejected and
therefore the first date accepted. It should be noted that this
implies that if Mary Anderson is her mother she must have married
at least twice, since the name of Margaret's father is not that of her
recently buried husband. It would therefore be safer to go no
further than presuming her to be relative without postulating the
actual relationship.

The use of the parish register has two further possible advantages.
If the list does not give occupational information, it is possible in
some cases to recover this from the register. The register can
also be used to date the list more accurately than the date given.
If, for example, the list is dated 1660 and it is important for the
researcher's purpose to know the month in which it was drawn up,
then the register for 1660 should be consulted with the aim of
discovering the latest baptism before the date of the list and the
earliest burial after it. The date of the list must fall between the
dates of these events. For example, if John Miles was baptised on
27th April and is included in the list, and Ann Jones, baptised on
5th June is not included, and if moreover, Joanna Roberts buried
1st May is not on the list but William Cooper buried 8th June is
included, then the list must have been compiled after 1st May but
before 5th June. More accurate dating is rarely possible since a
certain number of days or weeks were taken up in the preparation of the list, and the list is inconsistent during that period.

This completes the preliminaries and the actual analysis can now be begun with the assurance that a firm foundation has been laid.

4

Whatever the quality of a list the aims and methods of analysis remain the same; only the extent of the analysis varies.

The information recoverable from a list of inhabitants is static; that is, it is a state of affairs at a particular time, unlike most types of demographic information recoverable from local historical sources which involve the changes of various phenomena over a period of time. The process of analysis consists of counting heads in categories. In choosing the most interesting categories, two considerations have been taken into account; these are first the inherent limitations of the nature of the material which effectively rules out specifically demographic information and demands that the researcher concentrates on what may be loosely termed social structural information. Secondly the importance of the household as a unit in pre-industrial times is a deciding factor. The household unit was the labour unit (5). This being so, it is common sense to concentrate on the household, its composition and its variations; and to use the individual as a unit somewhat rarely.

The heads under which analysis is carried out are listed below in approximate order of the probability of the information concerned being recoverable.

1. Distribution of household size.
2. Sex and marital status of household heads.
3. Sex and marital status of total population.
4. Proportion of children in households and in population, distinguishing where possible offspring of conjugal family from other children.
5. Proportion of households containing less than, or more than, two generations.
6. Proportion of servants in population, and size and composition of servant groups in households.
7. Proportion of kin (outside the conjugal family resident in households).

8. Occupational structure of the community.

9. Variations of household size by occupational class.

10. Variations in numbers of dependent children, servants, and resident kin by occupational class.

11. Proportion and distribution of widows and orphans in the community.

12. Age structure of the community.

13. Effect of age and marital status of head of household on size of household.

14. Effect of age of mother on number of resident offspring.

15. Age structure of servant group.

16. Proportion and distribution of lodgers in the population.

For an ideal list all these would be recoverable. In practice, however, it is virtually never possible to recover information on lodgers for lists outside London and out of a total of well over two hundred lists analysed in Cambridge less than ten have given information about the age structure of the community.

To present such information requires the use of some form of table. A general rule is that the simplest table is the best, providing that no information is wasted. It is not practicable to try to squeeze all the information under each head on to one table. The Cambridge group is now using a set of fifty tables, some of which are subdivided. Each table should examine only one variable, lest differences or similarities be masked. Category thirteen therefore would require three tables each using the same breakdown of ages into groups, and each dealing with married, widowed and single heads respectively, the two latter categories being further subdivided by sex. Detailed information about the best type of tables to use has been given by Laslett (6).
The filling in of the tables is by far the simplest part of the whole operation. If cards are used these can be sorted into piles and counted for all tables which involve individuals, such as category three. For tables which do not involve individuals, or for all tables if cards are not being used, the list is scanned and an entry made in the appropriate box of the table. In the first category of information, the list is scanned and a mark for each household entered in the box corresponding to the number of persons in that household.

One of the most important things to remember in this, as in any other area of original work, is that it is never enough for the researcher to be perfectly clear about what he has done and why he has done it. It is most important that the reader also should be aware of exactly what has been done and why. Therefore it cannot be emphasised too strongly that all presumptions should be clearly stated, even those which seem most obvious.

The analysis of lists of inhabitants is a rewarding field of study. It takes comparatively little time, and, if sufficient care is taken, it can yield most interesting results.

NOTES

1. For example: London (see D.V. Glass London Inhabitants within the Walls 1695— London Record Society, 1966); Shrewsbury; Southampton; Kent, Wingham Division.


3. A thorough study of the Compton Census is being undertaken by Miss Ann Whiteman of Lady Margaret Hall, Oxford.

4. E.A. Wrigley, Editor. Introduction to English Historical Demography (1966), chapter 4, discusses tests for suitability and the reconstitution process itself.

5. P. Laslett, The World We Have Lost (1965), chapter 1.