A NEW APPROACH TO THE STUDY OF MARRIAGE HORIZONS

Jeremy Millard

Jeremy Millard is Staff Tutor in the Social Sciences in the East Midlands Region of The Open University

Introduction

There have been relatively few attempts by historical demographers to analyse marriage horizons beyond a simple examination of marriage numbers and percentages. This article aims, by way of a case study of parishes in north Buckinghamshire, to introduce readers of LPS to several more powerful techniques which are well within the grasp of the amateur local historian and which are, indeed, introduced in Leslie Bradley's A Glossary for Local Population Studies. It must be stressed from the outset, however, that the techniques introduced here are not intended to replace the more traditional historian's descriptive and analytical skills, nor supersede the more straightforward use of percentages. On the contrary, these additional techniques must be seen as complementing the more familiar methods. It is generally true that any technique, if properly used, can throw its own unique light on a subject, but, at the same time, its use inevitably involves a series of assumptions, short cuts or over-generalisations. This is why, in one sense, the more techniques used in analysing a particular topic the better, although in practice techniques will be selected, with both their strengths and weaknesses in mind, to do a particular job or fulfill a specific need. This article is therefore a demonstration of how certain techniques can be used when looking at the topic of marriage horizons. It will also be suggested that they are especially powerful in bringing out the underlying trend in a series of data, in other words, in looking for valid generalisations, as well as seeking to explain particular local situations. The value of these generalisations is that it allows us to make comparisons between different places and times.

The marriage distance

Parochial marriage registers constitute a vast and, as yet, largely untapped source of data for the analysis of marriage distances, both spatially and temporally. A marriage distance is defined as the distance between the parish of residence of the groom and the parish of residence of the bride, on the eve of their marriage, as recorded in the register. The site of the parish church is normally used in measuring these distances.
In looking at the marriage distances for a particular parish during a specific time period, it is possible to devise a measure of parish isolation by calculating the percentage of extra-parochial marriages (i.e. marriages in which one partner resided outside the parish). The higher this percentage is the less isolated the parish can be said to be. It is also demonstrable that, in general, parishes with large populations tend, for many reasons, including their better developed transport system, to have more, and more extensive, connections with the outside world; i.e. these parishes would, in these terms, exhibit a low degree of isolation. From a study, completed in 1976, however, it is apparent that there is a tendency for parishes with large populations to have low percentages of extra-parochial marriages. In other words, parishes with large populations, which we assume are less isolated than those with small populations, in fact come out as more isolated when this is measured by the percentage of extra-parochial marriages. This apparent paradox is a useful illustration of the fact that the use of quantitative techniques requires an awareness of the assumptions being made and the implications involved, and a very careful definition of what precisely is being measured. In fact the contradiction can perhaps be explained by the notion that in parishes with a large population any individual seeking a marriage partner within the parish has a greater chance of finding one simply because the population is large. It also tends to be the case that these parishes contain more urbanised populations, and it can be suggested that there is a greater likelihood of intermarriage within this population than with the more 'rural' parishes round about, because any individual tends to seek a mate from amongst the same or similar class or occupational grouping.

It is apparent, therefore, that, when trying to interpret something like the percentage of extra-parochial marriages, it is not possible to do so in any simple cause-and-effect manner, that many other factors also have a greater or lesser impact to make, and that it is constantly necessary to relate any attempt at explanation to the total societal context from which it comes. This lesson is particularly important when using graphical or quantitative techniques to examine an aspect of society, and this is certainly true of the techniques presented in this paper. However, I hope I can demonstrate the analytical value of certain techniques in advancing our understanding of the past, when these problems are borne in mind. Graphical and quantitative techniques are only aids, or tools, to help increase our understanding; they are certainly not panaceas allowing us to circumvent historical problems. In fact their use often throws up a whole new set of problems which must be faced by the historian. If they are faced successfully, however, they can add greatly to our knowledge.

Hypotheses

A study of extra-parochial marriages can also be used to examine spatial interaction. In other words, the distance, direction, and intensity of the inter-connections between different places over time. In this context, the marriage distance is a surrogate measure for spatial interaction. By a surrogate measure is meant something which will stand in, or approx-
imate, for something else for which there is no direct information. For example, historical demographers often use baptisms as a measure of births, and burials as a measure of deaths. They are not quite the same thing and there are, therefore, problems associated with their use. Nevertheless these surrogates are often the only measures available. In fact extra-parochial marriages are more than mere surrogates, because they are themselves one type of spatial interaction, and, in this sense, they are used as indicators of the general pattern of interaction over time.\(^3\)

Given fairly complete and accessible marriage registers, which also record each individual's parish, a study of marriage distances can commence at an early date. These conditions, however, do not normally prevail before 1754. It was in this year that Lord Hardwicke's 1753 'Act for the prevention of clandestine marriages', making it virtually impossible to contract a valid marriage unless it were carried out in a church according to an Anglican ceremony, came into force. The act also provided printed registration forms for the first time and these made provision for an individual's 'parish' to be recorded. From 1754, therefore, despite problems of under- and mis-registration, we can be reasonably certain that there is a representative picture of the marriage patterns of a particular parish. It was not until 1837 (when civil registration of births, marriages and deaths commenced) that male occupational information was explicitly asked for on the registration form, although it was many years before it was also elicited from the bride. Before this date detailed class and occupational analyses can be carried out on a few registers for limited periods, and it is only possible to examine the gross marriage patterns.

The compilation of marriage distance data must be based on sufficiently large samples, in order to ensure that the resulting interaction patterns are likely to reflect reality. For any one parish, this means aggregating data into time-period groupings. The periods I used in north Buckinghamshire were of forty years' duration, and, over the 160-year period from 1754 to 1913, this allowed four sequential patterns to be compared. As outlined above, it was felt that the accuracy and completeness of data is sufficiently reliable after 1754, whereas the socio-economic context (particularly because of transport possibilities) was significantly changed after 1913, for this data to act as a natural terminus. The four 40-year groupings were allocated because the commencement of the second, in 1794, is close to the appearance of the first relatively reliable population data (1801 Census); shortly after the commencement of the third in 1834 the era of railway construction began; and the last period (1874-1913) coincides nicely with the rapid breakdown of rural isolation first identified by Perry in his study of rural Dorest,\(^4\) thus facilitating a comparison with his study. Naturally there are many other possible period groupings. The point needs to be made, however, that they should be chosen both to ensure a sufficient sample size and to reflect the needs of particular localities and study objectives.

The study objectives which I pursued were articulated as a series of six hypotheses. These can be stated as follows:
1. Interaction between places is strongly influenced by distance, i.e. the greater the distance from the parish the less interaction there is with that parish. This phenomenon will be called the distance effect in this article. (Note in much geographical literature the phrase ‘distance decay’ is used.)

2. The distance effect itself decreases over time.

3. A sudden decrease in the distance effect is evident in the last quarter of the nineteenth century (as identified by Perry in Dorset).¹

4. In the context of north Buckinghamshire, there are two distinctive scales of interaction:
   
   (i) up to about twenty kilometres. This is the supposed maximum walking and, later, cycling distance enabling regular face-to-face contact to be maintained; this normally being necessary for a marriage to take place. This local scale of interaction exhibits no directional bias;

   (ii) over twenty kilometres. At this regional scale there is strong directional bias to the south-east and the north-west (i.e. to London and the Midlands). An important channel of this movement being Watling Street, later supplemented by the railway.

5. At the local scale of the interaction there is a tendency for urban parishes to interact more strongly with other urban parishes than with rural parishes. Thus, at this scale, the socio-economic character is another variable, in addition to distance, which affects spatial interaction. In this study, parishes were allocated to an index of agricultural occupation on the basis of the 1831 Census, taken midway through the study period, which gives nine occupational categories for males over twenty years of age on a parish basis (information which is not available before or after this date at the parish level). Of the categories given the major divisions are between persons employed in agriculture, those employed in manufacturing, in retailing and crafts, as capitalists, bankers, professional and other educated men, non-agricultural labourers, others, and servants. Ideally we need to identify ‘urban’ occupations in order to identify ‘urban’ parishes, but there are great problems in doing this. If the aim is to define ‘urban’ in terms of a ‘service’ centre, and not as a location for manufacturing or crafts (which, in any case, often took place within a rural environment in the early nineteenth century), the only occupation listed which we could use would be retailing. Unfortunately, retailing is grouped together with crafts. It appears intuitively sounder, however, to identify agricultural occupations with a ‘rural’ location, and to assume that ‘rural’ and ‘urban’ characteristics are the inverse of one another, so that as one increases the other falls. Thus, I decided to use what I have called the index of agricultural occupation as a surrogate for allocating parishes to either a ‘rural’ or an ‘urban’ category.

The index was obtained by ascertaining the numbers employed as
farmers or labourers in agriculture, and dividing this total by the numbers of those known not to have been employed as such. Those employed as servants or others were completely excluded from the calculation as we do not know whether or not they were employed in an agricultural context. The index thus obtained for each parish, within the local scale of interaction, was used to allocate it to either a ‘rural’ or an ‘urban’ category. Parishes with an index of 0.4 or above were categorised as ‘rural’, and those with less than 0.4 as ‘urban’. This figure is, of course, arbitrary, but it was in fact arrived at in this study by plotting the index for all parishes within the local field (as shown in figure 1b) and finding that, inter alia, the figure of 0.4 did clearly separate out what could be called an ‘urban group’ of parishes from the others.

6. There is some sort of negative relationship between the population size of a parish and the distance of interaction in a relative sense — i.e. the larger the population of a parish, the lower the proportion of marriage contacts over, say, twenty kilometres. In terms of absolute numbers (as opposed to proportions) of marriage distances, however, the larger the parish population, the greater is the contact over, say, twenty kilometres. (This hypothesis is in line with the discussion in the opening paragraphs of this article.)

- Techniques

The graphical and quantitative techniques now introduced allow a progressive analysis of the above hypotheses to be undertaken, and are of four main types:

(i) marriage contact fields;
(ii) chi-square analysis;
(iii) average marriage distances;
(iv) regression analysis.

(1) Marriage contact fields

These can be constructed by placing a grid, centred on the parish under study, over the area of interaction. Examples are shown in figures 1 and 2 where grids have been centred on the parish of Stony Stratford in north Buckinghamshire. This visual method specifically brings out any directional bias in spatial interaction and is probably the easiest method of summarising and analysing the data. Figure 1a shows the case study area divided into 25 grid cells, representing the ‘regional field’ discussed in hypothesis 4, and figure 1b shows the detail of the ‘local field’, divided, in this case, into sixty-four cells. Figures 2a and 2b show the number of marriage contacts falling in each of the grid cells, derived from figures 1a and 1b. In figure 2a note the strong north-west to south-east directional bias and the absence of any definite bias in the local field. This relates back to hypothesis 4, which could be said to be confirmed in this particular example. (A technique which can more objectively test this conclusion will be discussed in (ii) below.) Relating the numbers of marriage con-
tacts back to maps in this way is a useful reminder that the data being manipulated are, in fact, derived from a particular locality. Although the overall aim here has been stated as a search for generalisations, which can then be compared with similar results from other parts of the country, the detail of the analysis, especially when trying to explain specific anomalies, needs to be referred to the real world if any sense is to be made of the results. For example, the regional directional bias identified here may be readily understood by referring to the location of London and the urbanising Midlands, as well as by the orientation of the major lines of communication. Similarly, the general density and distribution of population in an area, particularly in upland Britain where physical features form barriers to movement, as well as channelling it, may have a pervasive influence on the pattern of spatial interaction.

Marriage contact fields, like those in figure 2, can be produced for successive periods in order to examine how this spatial interaction changes over time.

Figure 1a. Sketch map of the 'regional field', centred on Stony Stratford, showing the main settlements, Watling Street and the London and North-Western Railway (constructed in the 1830s).
Figure 1b. Sketch map of the 'local field'.

Figure 2. Observed marriage contact fields for Stony Stratford, 1834-1873. The figures in each grid cell indicate the number of marriage contacts from places in that cell with Stony Stratford. The local field shows the detail of the central grid cell of the regional field. Stony Stratford was identified as an 'urban' parish (see text) and its contacts with other 'urban' parishes are marked with an asterisk in the local field.
(II) Chi-square analysis

Hypothesis 5 proposed that socio-economic factors, as well as distance, would be important in influencing spatial interaction, especially when measured by marriage distances. Ideally, of course, we would wish to examine the socio-economic characteristics of the individual bride and groom, normally by recording their employment or social status, but this information, for both partners, is rarely available in the marriage registers. (Although it may be possible to discover it through the technique of nominal linkage; i.e. by cross-checking individual names from other sources, such as the census, where the required information may be found.)

An alternative method is to classify parishes into socio-economic types (see hypothesis 5 above) and to measure the interaction between both similar and dissimilar types. This is legitimate as long as the analysis undertaken uses aggregated interaction data; it would not be acceptable to analyse specific marriages in this way or to use very small samples, because they may not be representative of the parish as a whole.

Figure 2b uses the grid derived from figure 1b, with its centre positioned over Stony Stratford. It shows that Stony Stratford (which is identified as an 'urban' parish) had twenty-five contacts with other 'urban' parishes and forty-two contacts with 'rural' parishes within the local field. On the face of it, these figures show that Stony Stratford has more contact with 'rural' parishes than with 'urban' ones, but this would be to ignore the proportion of these two types of parish present, both in terms of their number and the population they contain. This could be taken into account by enumerating the number and population size of all 'urban' and 'rural' parishes within the area of the local field. However, an easier method, which does not require this detailed information because it remains constant (or 'given') no matter which parish's contact pattern is being analysed, is to examine interaction data for a number of parishes. I set out below the extra-parochial marriage data for six parishes in north Buckinghamshire, located at a maximum distance of fifteen kilometres from each other (see figure 1b), and with 'urban' and 'rural' types spatially mixed, in order to reduce any impact that distance may have on the data. This is important because we are attempting to isolate the socio-economic factors from the geographic ones.

<table>
<thead>
<tr>
<th>OBSERVED DATA</th>
<th>extra-parochial marriage type</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>urban</td>
</tr>
<tr>
<td>parish type</td>
<td>urban</td>
</tr>
<tr>
<td></td>
<td>rural</td>
</tr>
<tr>
<td>where marriage took place</td>
<td>total</td>
</tr>
</tbody>
</table>

A visual examination of these figures seems to show that, indeed, parishes of the same socio-economic type do interact more with each other than with parishes of other types. It is possible, however, to be much more precise in analysing this data by using a simple technique known as chi-square. Basically, chi-square compares an observed situation (as, for
example, the data given above) with one which we would expect if there were no association between the variables being investigated. The two variables in this case are, firstly (as shown in each row above), the socio-economic character of the parish in which the marriage took place (and where one of the marriage partners resided), and, secondly (as shown in each column) the socio-economic character of the parish from which the extra-parochial marriage partner came. We are interested in whether or not there is any association between these two variables.

We now need to compare this observed distribution with the expected distribution which can be easily calculated. For a given observed figure, its expected counterpart is derived by multiplying the row total by the column total and dividing by the grand total. Thus, the first observed figure of 52 has an equivalent expected figure of $\frac{172 \times 61}{275} = 38.2$.

This calculation produces expected values based on the assumption that there is no association between the two variables. Given the total number of marriages taking place in ‘urban’ parishes (i.e. 172) and the total number of marriages with ‘urban’ extra-parochial marriage partners (i.e. 61), a figure is calculated (i.e. 38.2) purely on the basis of the proportions of the ‘urban’ categories available within each of the two variables. Each of the expected values is calculated in the same manner, thus:

<table>
<thead>
<tr>
<th>EXPECTED DATA</th>
<th>extra-parochial marriage type</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>urban</td>
</tr>
<tr>
<td>parish type</td>
<td>urban</td>
</tr>
<tr>
<td></td>
<td>rural</td>
</tr>
<tr>
<td>where marriage took place</td>
<td>total</td>
</tr>
</tbody>
</table>

It should be noted that an important limitation of chi-square is that not more than one-fifth of the expected values should be five or less. Clearly there is no problem in this case, but if the criterion is not met it would be necessary to combine two or more of the categories within each variable until it is. In this example there are only two categories (i.e. ‘urban’ and ‘rural’) so combination could not, in any case, be undertaken. It would have been quite possible, however, to have devised a socio-economic classification by allocating parishes not just into a broad or urban-rural dichotomy but into three, or more, categories. For example, where sample sizes are large enough, an ‘urban’ type could be defined as having an index of rurality of 0.3 or less (as in this study), an ‘intermediate’ type of between, say, 0.3 and 0.7, and a ‘rural’ type of over 0.7. Finer or different categories could of course be devised. Again it should be stressed that any classification needs to reflect the particular needs of the study. The categories within each of the variables used in chi-square should therefore be large enough not to contravene this one-fifth rule, but also numerous enough to examine characteristics thought important. It will be readily appreciated that the bigger the difference between the observed and expected values, the more likely it is that they are significantly different from each other, and therefore the more likely that our initial
hypothesis will be upheld. The chi-square technique involves calculating this difference then squaring it (i.e. multiplying it by itself) in order to get rid of any negative values, dividing it by the expected value so that the difference calculated is expressed as a proportion of the expected value, and then summing all the values obtained to give a composite chi-square statistic. This procedure is given by the formula:

\[ \text{chi-square} = \sum \frac{(0 - E)^2}{E} \]

where \( \sum \) simply means add up the individual items that follow,

\[ 0 = \text{observed value} \]

\[ E = \text{expected value} \]

The calculation is normally effected by constructing a table:

<table>
<thead>
<tr>
<th>0</th>
<th>E</th>
<th>(0 - E)</th>
<th>(0 - E)^2</th>
<th>(0 - E)^2 + E</th>
</tr>
</thead>
<tbody>
<tr>
<td>52</td>
<td>38.2</td>
<td>13.8</td>
<td>190.4</td>
<td>5.0</td>
</tr>
<tr>
<td>120</td>
<td>133.8</td>
<td>-13.8</td>
<td>190.4</td>
<td>1.4</td>
</tr>
<tr>
<td>9</td>
<td>22.8</td>
<td>-13.8</td>
<td>190.4</td>
<td>8.4</td>
</tr>
<tr>
<td>94</td>
<td>80.2</td>
<td>13.8</td>
<td>190.4</td>
<td>2.4</td>
</tr>
</tbody>
</table>

\[ \sum 17.2 \]

(Note: The apparent symmetry in the above table, giving the same \((0-E)\) and thus \((0-E)^2\) values, is due to the fact that each of the two variables only has two categories: 'urban' and 'rural'. If more categories were to be used this would be unlikely to happen.)

The next step is to work out a figure called the **degrees of freedom**. This is simply a measure of how many chances for variation exists in the observed table of figures we started with. If we accept that the row and column totals are fixed, then the only chance for variation in terms of the columns is one, since once a value is entered in the first column the value in the second column is pre-determined. Similarly, if we examine the two rows, there is again only one chance of variation, since after the first value is entered the second is fixed. If, however, there had been, say, five rows, then the chance for variation would have been four, as the first four figures would be free to vary (assuming they did not exceed the total), leaving the fifth as pre-determined. Thus the degrees of freedom (or the total chances for variation which have called it) equals the number of columns minus one, multiplied by the number of rows minus one. In our case this is \((2 - 1) \times (2 - 1) = 1\). In order to assess the significance of the result, it is necessary to refer to either a chi-square table or graph (found in the standard textbooks). This will show that a figure of 17.2 with 1 degree of freedom falls well below the 0.1% 'critical value' of chi-square. This means that this could only be produced by chance less than 0.1% of the time (or one time in a thousand).

By convention, the observed distribution of data between any two
variables is taken as being 'significant' (i.e. that there is an association between them) if the chance or random element could be responsible 5% or less of the time. You should note, however, that this is only a convention, and that the 'critical value' of any chi-square statistic you calculate should always be quoted. In this case, we note that chance has not been eliminated (i.e. it is possible, though unlikely, that the figures were produced by chance), and that even if the 'critical' value had been, say 10%, this would still mean that chance plays a 'small' role even though, by convention, we say it plays 'too big' a role. The point is that the value of a technique like chi-square is that it enables words like 'small' or 'big' to be discarded in favour of a numerical value; it does not allow us to definitely 'prove' or 'disprove' a hypothesis. In fact, even though by using statistical jargon, we could say our result was 'very significant', all we have really done is to show that within the constraints of the table of data we have constructed, it is unlikely that the figures we have observed are randomly determined. We can only conclude that extra-parochial marriages are not randomly distributed between 'urban' and 'rural' parishes. The notion we have that it is 'urban-urban' and 'rural-rural' links which are over-represented (as compared with 'urban-rural' and 'rural-urban' links) is only added by inspection of the observed table. In other words, it is where the numbers fail that leads us to surmise this, not the value of the chi-square statistic itself. The latter can only indicate the strength, not the form, of a relationship.

Despite these important caveats, however, a technique like chi-square is more powerful than a simple visual inspection of the figures. This is because it allows us to quantify the relationship between two variables, to be precise in allocating the influence of chance, and, thereby, to be able to make direct comparisons with other chi-square results calculated with different data or by different researchers.

The chi-square test is a very valuable and simple technique which has many uses. It can, for example, test the significance of the directional bias observed in marriage contacts at the regional level, discussed in (1) above. If the data presented in figure 2a are taken as a display of observed values, we can proceed to calculate expected values. First of all, however, it will be apparent that, because so many of the cells in fig. 2a have nil values and many others are under 5, it is very likely that the expected values will not meet the one-fifth criterion discussed above. A check would confirm this to be correct. It is necessary, therefore, to combine cells so as to boost the value of each resulting group of cells. Bearing in mind that we wish to test for directional bias in a north-west to south-east direction, the 25 cells can be amalgamated into 4 cell groups, so that there is a north-west, north-east, south-east and south-west group. Ignoring the central cell (which is disregarded because it has no effect on directional bias) it will be seen that each of the 4 cell groups contains 4 whole cells and 4 half cells. This is because the 8 central vertical and horizontal cells each fall into 2 of the 4 cell groups. Thus, the total number of marriage contacts in the north-west cell group is $3 + 2 + 4 + (10 ÷ 2) + (3 ÷ 2) = 15.5$. Producing cell groups in this way also removes the distance variable from the analysis, since each of the 4 groups
is equidistant from the centre, thus the test will focus exclusively on directional bias.

Proceeding on these lines, the following observed values are obtained:

<table>
<thead>
<tr>
<th>Cell Group</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>North-west</td>
<td>15.5</td>
</tr>
<tr>
<td>North-east</td>
<td>6.5</td>
</tr>
<tr>
<td>South-east</td>
<td>15.5</td>
</tr>
<tr>
<td>South-west</td>
<td>3.5</td>
</tr>
<tr>
<td><strong>Total</strong></td>
<td><strong>41</strong></td>
</tr>
</tbody>
</table>

The calculation of expected values in this instance cannot be achieved by using row, column and grand totals as in the previous example, because the data we have is not in this sense cumulative. Simply using common sense, however, we would expect, if there was no directional bias, that each of the 4 cell groups would have an equal share of the total number of marriage contacts available; in other words, each would have 41 divided by 4, or 10.25.

A chi-square calculation table is constructed in precisely the same manner as above and a result of 11.2 is obtained. Similarly, we cannot calculate the degrees of freedom on the basis of number of rows and number of columns. In this case, we simply subtract 1 from the total number of variables. Therefore degrees of freedom equals 3. This makes sense if we consider that, if the total is fixed, in this case at 41, then the first 3 cell-groups are free to vary, but once they are determined the fourth is fixed.

Reference to a chi-square graph or table reveals that the difference between the observed and expected values is between the 0.1% and the 1% critical value of chi-square. In other words, there is by convention a significant difference between the two and the visual impression of directional bias obtained from fig. 2a is supplemented by a more objective piece of evidence. As noted above, however, we should remember that all we have really done is to show that the scatter of Stony Stratford's extra-parochial marriage contacts is unlikely to be the result of random processes. The notion that there is a north-west to south-east orientation to the pattern is one which is added by looking at the data; the chi-square statistic does not tell us this. All we can do is to rule out the idea that the distribution is random, and we must articulate further hypotheses if we are to try to discover the reason for this. For example, reference to fig. 1a will show that Birmingham is situated in the north-west sector and London in the south-east. The directional bias may have been due to these two large towns, and we could test this possibility by doing the analysis once again, excluding them. If the directional bias were thereby removed, we may be strengthened in the notion that it was Birmingham and London 'causing' it, although we should still not rule out the possibility of other 'causes' which we had not thought of.
Using chi-square with geographical data of this kind, where the grid, as the unit of measurement, is arbitrary, does require care. Our expected distribution of values, for example, was produced on the assumption that there was an equal number of potential contacts (or population) in each cell, or cell group. Clearly this is not the case, and we must refine our reasoning, and hence our conclusion, by constant reference to the realities of the geography of the area.

Table 1. Marriage distance data for Stony Stratford, 1754-1913.

<table>
<thead>
<tr>
<th></th>
<th>(1) total marriages</th>
<th>(2) total extra-parochial marriages</th>
<th>(3) extra-parochial marriage %</th>
<th>(4) mean extra-parochial marriage distance (kilometres)</th>
<th>(5) median and quartile extra-parochial marriage distances (kilometres)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1754-1793</td>
<td>370</td>
<td>108</td>
<td>29.2%</td>
<td>29.8</td>
<td>4.1, 11.8, 34.2</td>
</tr>
<tr>
<td>1794-1833</td>
<td>509</td>
<td>164</td>
<td>32.2%</td>
<td>23.9</td>
<td>2.1, 11.3, 31.5</td>
</tr>
<tr>
<td>1834-1873</td>
<td>381</td>
<td>115</td>
<td>30.2%</td>
<td>34.7</td>
<td>2.1, 11.3, 60.3</td>
</tr>
<tr>
<td>1874-1913</td>
<td>361</td>
<td>156</td>
<td>43.2%</td>
<td>49.0</td>
<td>2.2, 21.2, 80.9</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th></th>
<th>(8) mean census population</th>
<th>(9) extra-parochial annual marriage rate (per 1000)</th>
<th>(10) (11) gradient of regression line</th>
<th>(12) correlation coefficient (r)</th>
<th>(13) coefficient of determination (100r²)</th>
<th>(14) % of extra-parochial marriages used in regression</th>
</tr>
</thead>
<tbody>
<tr>
<td>1754-1793</td>
<td>—</td>
<td>—</td>
<td>-2.19</td>
<td>-0.93</td>
<td>86.6%</td>
<td>94.4%</td>
</tr>
<tr>
<td>1794-1833</td>
<td>1564.7</td>
<td>2.6</td>
<td>-2.39</td>
<td>-0.94</td>
<td>88.6%</td>
<td>98.2%</td>
</tr>
<tr>
<td>1834-1873</td>
<td>1873.7</td>
<td>1.5</td>
<td>-2.03</td>
<td>-0.92</td>
<td>84.3%</td>
<td>93.0%</td>
</tr>
<tr>
<td>1874-1913</td>
<td>2070.7</td>
<td>1.9</td>
<td>-1.87</td>
<td>-0.86</td>
<td>74.4%</td>
<td>82.7%</td>
</tr>
</tbody>
</table>

(II) Average marriage distances

In addition to the examination of successive marriage contact fields, the changing nature of spatial interaction can also be analysed by reference to marriage distance data which has been manipulated in certain ways. For example, it is possible to compare the four periods by looking in turn at the following data in Table 1:
Column 3 percentage of extra-parochial marriages.

Column 4 the mean extra-parochial marriage distance. The mean is what we normally term the average. It is simply arrived at by adding each individual marriage distance and dividing the total by the number of distances used.  

Columns 5, 6, 7 the mean and quartile extra-parochial marriage distances. The mean, as shown in column 4, gives no indication of the range of spread of values about it and is badly affected by extreme values. The median, on the other hand, is not so affected, as it is simply the central number, below which, and above which, fifty percent of all the numbers fall. For example, if there are twenty-three distances arrayed in ascending order, the median is the twelfth distance. Similarly, the upper quartile is the median between the actual median and the highest value, and the lower quartile is the median between the actual median and the lowest value. The calculation of median and quartile distances gives some indication of the spread of values, and is, therefore, probably more suitable than the mean for comparing marriage distances.

Columns 8, 9, 10 the marriage rate is the number of marriages in a given year expressed as a rate per thousand of the total population. In the example, only the last three periods have rates calculated for them because of the unreliability of population data before the Census commenced. Marriage rates have been used because they tell us, for a given population, how many marriages there were (according to the registers). This may have an effect, although this is speculation, on the marriage distance patterns produced, both as a result of socio-economic or cultural factors and also as a result of the geographical availability of suitable marriage partners, which may be particularly problematic in a small parish.

When using marriage rates, however, it must not be forgotten that an important factor is the age structure of the population to which it refers. For example, if there were a higher than average proportion of persons in the 20-30 age group we would expect, regardless of any socio-economic, cultural or geographic influences on marriage, that the marriage rate would be boosted. This is not a factor that has been directly taken into account in this case study, although it would have been possible to do so using the published Census, from 1821 onwards, when information about age is first included. It is, however, a factor which should not be overlooked.

(iv) Regression analysis

Regression analysis involves the comparison of two variables in such a way that we can see how a change in one variable results in, or is the result of, a change in the other. The two variables of concern here are distance and the number of marriage contacts. For every extra-parochial marriage we can, of course, calculate its marriage distance, but every
marriage distance is a unique and specific measurement. In order to see the effect of distance on the number of marriages, we need to group the marriages in a certain way. This is done by allocating each marriage to a distance band, by drawing a series of concentric circles, around the parish church under study, in such a way that each successive circle has the same increase in radius as the last. In this case study a series of distance bands were chosen, each with a width of five kilometres. Kilometres were used, rather than miles, simply because the marriage distance was not measured directly on the map, but calculated from the six figure National Grid Reference (based on kilometres) for the two parish churches, using the Pythagoras theorem. Five kilometre distance bands were chosen because any smaller distance did not allow at least one marriage contact to fall within each band, up to a one hundred kilometre limit, beyond which bands containing no contacts start to appear. Column 14 of Table 1 shows the percentage of extra-parochial marriages included as a result of imposing this limit.

Another important decision which had to be made was whether or not to include the number of intra-parochial marriages in the first distance band (i.e. 0 to 5 kilometres). On the one hand it could be argued that because, in 1834 to 1873 for example, 266 marriages took place in Stony Stratford between people who had a 'zero' marriage distance, this is just as crucial in measuring the distance effect on the selection of marriage partners as the fact that eleven marriages took place with people from another parish. On the other hand, one could argue that because we are only concerned to examine interaction between different places, we can ignore the number of intra-parochial marriages. Obviously, whichever way the decision is made will have a dramatic effect on the analysis of data, simply because the number of intra-parochial marriages is so large. In this case study, I chose to use only extra-parochial marriages, mainly because it seems that the majority of earlier studies do the same, and comparison between findings is a very important aim of any work of this nature.

The way in which all these problems are perceived and confronted must, of course, depend upon the preferences, aims, and difficulties encountered by each researcher. It is important to remember, however, that such decisions can have a pervasive effect on the results obtained and they must therefore be only considered with a full realisation of the implications involved.

These problems are a good example of the additional difficulties thrown up by the use of new techniques which are, nevertheless, worth tackling because of the additional understanding we are able to gain.

Some results of these procedures are shown in Table 2. Column 1 indicates the distance band, and column 3 the number of marriages falling in that band. Notice that column 4 is labelled standardised number of marriages. This is because the number of marriages at any given distance must be standardised for area. The reason is that, as distance increases away from the study parish, there are potentially a greater number of parishes supplying marriage partners. As described above, when a series of concentric rings is drawn outward from the study parish, the area of
Table 2. Number of extra-parochial marriage contacts by distance bands, Stony Stratford, 1834-1873.

<table>
<thead>
<tr>
<th>(km.)</th>
<th>(1) Distance band order</th>
<th>(2) ratio of area to first distance band</th>
<th>(3) number of marriage contacts</th>
<th>(4) standardised number of marriage contacts</th>
<th>(5) logarithm of column 4 (y axis)</th>
<th>(6) logarithm of column 1 (x axis)</th>
</tr>
</thead>
<tbody>
<tr>
<td>0-5</td>
<td>1</td>
<td>1</td>
<td>11</td>
<td>11.0000</td>
<td>1.0414</td>
<td>0.0000</td>
</tr>
<tr>
<td>6-10</td>
<td>2</td>
<td>3</td>
<td>34</td>
<td>11.3333</td>
<td>1.0543</td>
<td>0.3010</td>
</tr>
<tr>
<td>11-15</td>
<td>3</td>
<td>5</td>
<td>16</td>
<td>3.2000</td>
<td>0.5952</td>
<td>0.4771</td>
</tr>
<tr>
<td>16-20</td>
<td>4</td>
<td>7</td>
<td>4</td>
<td>0.5714</td>
<td>-0.2431</td>
<td>0.6021</td>
</tr>
<tr>
<td>21-25</td>
<td>5</td>
<td>9</td>
<td>6</td>
<td>0.6667</td>
<td>-0.1761</td>
<td>0.6990</td>
</tr>
<tr>
<td>26-30</td>
<td>6</td>
<td>11</td>
<td>6</td>
<td>0.5454</td>
<td>-0.2633</td>
<td>0.7782</td>
</tr>
<tr>
<td>31-35</td>
<td>7</td>
<td>13</td>
<td>2</td>
<td>0.1538</td>
<td>-0.8130</td>
<td>0.8451</td>
</tr>
<tr>
<td>36-40</td>
<td>8</td>
<td>15</td>
<td>1</td>
<td>0.0667</td>
<td>-1.1759</td>
<td>0.9031</td>
</tr>
<tr>
<td>41-45</td>
<td>9</td>
<td>17</td>
<td>2</td>
<td>0.1176</td>
<td>-0.9296</td>
<td>0.9542</td>
</tr>
<tr>
<td>46-50</td>
<td>10</td>
<td>19</td>
<td>2</td>
<td>0.1053</td>
<td>-0.9776</td>
<td>1.0000</td>
</tr>
<tr>
<td>51-55</td>
<td>11</td>
<td>21</td>
<td>1</td>
<td>0.0476</td>
<td>-1.3224</td>
<td>1.0414</td>
</tr>
<tr>
<td>56-60</td>
<td>12</td>
<td>23</td>
<td>3</td>
<td>0.1304</td>
<td>-0.8847</td>
<td>1.0792</td>
</tr>
<tr>
<td>61-65</td>
<td>13</td>
<td>25</td>
<td>1</td>
<td>0.0400</td>
<td>-1.3979</td>
<td>1.1139</td>
</tr>
<tr>
<td>66-70</td>
<td>14</td>
<td>27</td>
<td>1</td>
<td>0.0741</td>
<td>-1.1302</td>
<td>1.1461</td>
</tr>
<tr>
<td>71-75</td>
<td>15</td>
<td>29</td>
<td>1</td>
<td>0.0370</td>
<td>-1.4318</td>
<td>1.1761</td>
</tr>
<tr>
<td>76-80</td>
<td>16</td>
<td>31</td>
<td>2</td>
<td>0.0968</td>
<td>-1.0141</td>
<td>1.2041</td>
</tr>
<tr>
<td>81-85</td>
<td>17</td>
<td>33</td>
<td>8</td>
<td>0.2424</td>
<td>-0.6155</td>
<td>1.2304</td>
</tr>
<tr>
<td>86-90</td>
<td>18</td>
<td>35</td>
<td>3</td>
<td>0.0857</td>
<td>-1.0670</td>
<td>1.2553</td>
</tr>
<tr>
<td>91-95</td>
<td>19</td>
<td>37</td>
<td>1</td>
<td>0.0270</td>
<td>-1.5686</td>
<td>1.2788</td>
</tr>
<tr>
<td>96-100</td>
<td>20</td>
<td>39</td>
<td>2</td>
<td>0.0513</td>
<td>-1.2899</td>
<td>1.3010</td>
</tr>
</tbody>
</table>

Each successive ring is greater than the inner ring. The greater the area the higher the population it can potentially contain and therefore the more marriage partners it could supply. The number of marriage contacts each ring provides is therefore standardised so that it is expressed as the number of marriages per unit area. This is, in fact, less complicated than it sounds, as the unit of area used is that contained in the innermost ring. If this inner ring is said to have an area of one, then, by using the formula \( A \) \( r^2 \) for the area of a circle, it can be shown that the second ring has an area of three, the third of five, the fourth of seven, the fifth of nine etc. The number of marriage contacts in the second ring are therefore divided by three, the number in the third by five, and so on, to give the standardised number of marriages for each successive distance band away from the study parish. The ratios to be used in each case are shown in column 2.

It is obvious that this standardisation procedure would be more reliable if it were based on the actual population in each ring rather than the area of the ring. The difficulty here is, of course, that we do not have reliable population statistics for parishes falling within each ring until the 1801 Census, and, even after this date, the population figures may hide much inter-decennial variation in actual numbers. The distribution of population
in north Buckinghamshire and adjacent areas, was, throughout the study period, relatively uniform in the sense that there were no areas of sparse population. But in parts of the country where this is likely to be a problem, population figures, rather than land area, should be used if possible. Whether or not this can be done, constant reference to the realities of the local geography of an area (for example, figures 1a and 1b) should guide the interpretation of the results. It can be noted, for instance, that the number of marriages (column 3 of Table 2) in the 81-85 kilometre distance band goes dramatically against the trend of decreasing marriages with increasing distance, probably London and Birmingham are included within this band. If we were able to standardise on the basis of population in this case study it is likely that this band would conform more closely to the trend.

Once the data have been prepared in a suitable form, they can be used to construct a scatter graph. Most graphs, such as those shown in figures 3 and 4, are made up of two axes: a horizontal (or x axis) and a vertical (or y axis). These two axes represent the two variables which are being compared. We want to see how a change in one variable (called the independent or x variable) affects the other variable (called the dependent or y variable). In our example, the independent variable is distance, and the dependent variable is the standardised number of marriage contacts. To produce a scatter graph each point is located with reference to these two axes; for example, if there were fifteen marriages at five kilometres, the point representing this would be level with fifteen on the number of marriages axis and with five on the distance axis. Each point is plotted in the same way on the graph.

![Figure 3. Graph showing the relationship between standardised number of marriages and distance in Stony Stratford, 1834-1873.](image)

![Figure 4. Double-log graph showing the relationship between standardised number of marriages and distance in Stony Stratford, 1834-1873.](image)
Now, it has been hypothesised (see hypothesis 1) that, as distance increases, the number of marriage contacts decreases (i.e. a distance effect). We therefore expect to see a series of points which trend from the top left-hand side of the graph to the bottom right-hand side. Reference to fig. 3 (based on data from columns 1 and 4 of Table 2) will show that this does, in fact, happen, although the trend is not regular but seems to indicate that, with increasing distance, the rate in the reduction of marriage contacts decreases. Figure 3 also shows the best-fit line, (i.e. that which best summarises the trend of the points), as a steep curve.

Regression analysis involves the construction of a best-fit line mathematically, to produce a regression line which can be expressed as a mathematical equation. It is this equation which specifically defines the relationship between the two variables. Although curves are susceptible to mathematical expression, the calculations are complex. One way of getting around this problem is, rather than have two axes with uniform scales, to construct scales based on logarithmic numbers. This produces an axis which increases proportionately rather than in absolute terms, as shown in figure 4 (based on data from columns 5 and 6 of Table 2). When this double-log graph (as it is called) is used with most data which analyses the distance effect, a straight line is produced; i.e. the trend of the points approximates to a linear, or straight-line, relationship. This linear regression line has a relatively simple mathematical equation

\[ y = ax + b \]

where \( y \) (as the dependent variable) is the standardised number of extra-parochial marriage contacts,

\( x \) (as the independent variable) is distance

\( a \) is the slope, or gradient, of the line (i.e. the greater \( a \) is the steeper is the line)

and \( b \) is the value of \( y \) where the line crosses the \( y \) (vertical) axis.

The gradient of the regression line, '\( a \)', in distance effect equations invariably slopes from top left to bottom right, and, by convention, this is termed a 'negative' slope, with '\( a \)' being given as a negative value. If the regression line sloped from bottom left to top right the value of '\( a \)' would be positive. In figure 4, '\( a \)' has a value of -2.03, and reference to Table 1 will show how this value changes for each of the four time periods (column 11). If the distance effect decreases over time, therefore, we would expect the value of '\( a \)' to also decrease.

The regression line equation, therefore, mathematically describes the trend of the data. However, it does not indicate how close the fit is between the line and the points scattered around it on the graph. To measure this, the correlation coefficient is used, which ranges from +1, through 0, to -1. Zero indicates no correlation at all; i.e. the points are randomly distributed right across the graph in such a way that there is absolutely no trend. A value of one indicates perfect correlation; i.e. every point is exactly on the regression line. The coefficient is either positive or negative, depending on whether the gradient of the regression line is positive or negative as described above.
The correlation coefficient of the regression line shown in figure 4 is —0.92, which indicates that all the points are very close to the line. In fact, the regression lines for each of the four periods in this example exhibit very high correlation coefficients, which tends to mean that a great deal of faith can be invested in the ability of these regression analyses to describe the marriage distance data. This is further supported by another necessary statistical technique, which we must carry out, called the standard error of the coefficient. This indicates how ‘significant’ the correlation coefficient is, i.e. how probable is it that the data was not generated by a random process. In general, the larger the sample size (i.e. the greater the number of points) together with a high correlation coefficient, the more ‘significant’ the correlation can be said to be. In the example, all the correlation coefficients produced were highly significant; i.e. all had less than 1 time in 1000 of being chance results. The correlation coefficients do, in fact, seem to vary in relation to the percentage of extra-parochial marriage data used in the regression analysis (column 14 of Table 1). This is, arguably, to be expected, and shows that the results are influenced by the decisions taken in regard to the width of the distance bands and their outer limit.

One further useful technique is termed the coefficient of determination which is calculated in column 13 of Table 1. It is quite simply derived by squaring the correlation coefficient and multiplying the result by 100. Its usefulness is that it tells us how much of the total variation in the dependent variable is associated with, or ‘explained’ by, the variation in the independent variable. In other words, we can say that, using the 1834-1873 data, 84.3% of the variation in the standardised number of extra-parochial marriages is ‘explained’ by the distance between the parish of residence of the groom and that of the bride.

Conclusion

The techniques described above were used to analyse spatial interaction in six parishes in north Buckinghamshire from 1754-1913. All six of the hypotheses stated were supported with one important aberration.

A distance effect was found in the data for all six parishes, and the strength of this effect generally decreased over time. However, in three of the six parishes the second period (1794-1833) was marked by an increase of the distance effect (see column 11 of Table 1 which shows the Stony Stratford figures). The same three parishes also showed relatively high marriage rates (both extra-parochial and total) in this period (see columns 9 and 10). The distance effect in the other three parishes steadily decreased in strength over time, and there was not this variation in the marriage rates. There seems to be, therefore, some sort of relationship between a steadily decreasing distance effect and stable marriage rates on the one hand, and high marriage rates and an increasing distance effect on the other.

Reference to the mean and median distances (as for example in columns
4-7 of Table 1) also seems to indicate a lowering of marriage distances between 1794 and 1833. In addition, the extra-parochial percentages (column 3) show an increase at this time a little above that expected from the overall trend throughout the period, and this seems to confirm the evidence of the marriage rates that, although there were a higher proportion of extra-parochial marriages, the distances which they mark out are less than expected.\textsuperscript{21}

Are there any reasons why spatial interaction was more restricted than expected, and marriage rates abnormally high in this period? (Note, however, that because we do not have reliable population data before this period, we cannot, with certainty, state that the rates had risen at this time; although see Note 22.) The very fact that we begin to ask questions such as these, which arise directly out of the application of the techniques demonstrated, forges the essential link between a perhaps mechanical and pedestrian exercise and the focus of historical and intellectual interest which is the real objective of our endeavours. In the same way that we noticed one figure in particular on Table 2 as being a deviant against the overall trend of decreasing numbers of marriages with increasing distance (i.e. in the 81-85 kilometre band), asked why, and we discovered London and Birmingham fell in that band, we can ask questions about the overall thrust of the results; particularly if they are not what we expected. These questions must include, of course, a realisation that the particular technique employed and a particular procedure followed may themselves, and not the historical reality, be responsible for the pattern of the results. But when we find that a combination of different techniques point to a similar conclusion, we can have a lot more faith that our results are genuinely indicating something of interest.

It is beyond the scope of this article to discuss any possible historical reasons for this apparent retrenchment of marriage horizons, in the last decade of the eighteenth century and first decades of the nineteenth century, beyond suggesting that the reasons may include the level of wage rates and the operation of the Poor Law at the time. I have instead been concerned with means rather than with ends, and have therefore tried to show how a series of techniques of increasing complexity, but not beyond the reach of the amateur local historian, can be used in the analysis of spatial interaction based on marriage distance data. The techniques demonstrated here do not, by any means, constitute an exhaustive list, but they do enable, with diligent and careful use, a greater degree of understanding to be brought to bear upon a particular historical topic.
NOTES

3. The most important historical, geographic, and demographic studies using marriage distances to examine spatial interaction and mobility are:—
5. The six parishes incorporated into the 'parish type where marriage took place' variable are Stony Stratford (urban), Newport Pagnall (urban), Haversham (rural), Little Brickhill (rural), Loughton (rural), Tyningham (rural). Parishes incorporated into the 'extra-parochial parish type' are, of course, all the parishes within the local field.
10. D. Mills, 'Aspects of Marriage: an example of Applied Historical Studies', a Social Science publication, The Open University 1980. Mills contrasts market towns in lowland England with parishes, such as those in Wharfedale, with a large land area but small scattered population (pp. 8-10).
12. The Pythagoras theorem simply states that the square of the hypotenuse (the side of a right-angled triangle which is opposite the right angle) equals the sum of the squares of the other two sides. Thus if we wish to work out the straight line distance between two parish churches using the National Grid Reference, we first have to ascertain the six figure grid reference for each church. This is done either by reading it from an Ordnance Survey map, or by reference to a suitable gazetteer. The first three figures of the grid reference refer to the 'easting' of a church from the National Grid's origin, and the second three figures refer to its 'northing'. Each of these pairs of three figures will locate the church to the nearest 0.1 kilometre. In order to work out the grid reference of the third point of the triangle, the point which makes the right angle, we simply take the lowest 'easting' of the pair of parishes and the lowest 'northing', to make a composite six figure grid reference. For example, if the two parish churches are located at 415370 and 455340 respectively, the grid reference of the right-angled point of the triangle is 415340. It is now a simple matter to calculate the 'easting' distance and the 'northing' distance of the two parish churches from this right-angled point, the former being: 45.5 — 41.5 = 4.0 kilometres
   and the latter: 37.0 — 34.0 = 3.0 kilometres.
   The distance between the parish churches is the hypotenuse of the triangle and is found as described above:
   \[(\text{hypotenuse})^2 = 4^2 + 3^2 = 16 + 9 = 25\]
   \[\text{hypotenuse} = \sqrt{25} = 5 \text{ kilometres.}\]
   This method is particularly useful for the calculation of longer distances, because it is not affected by the inaccuracies which inevitably result from directly measuring the distance on the map using a ruler.

30
13. The reason it is necessary to avoid zero marriage contacts falling in any distance band is because the regression analysis undertaken involves using the logarithm of the number of marriage contacts (as described below) and, as the log of zero does not exist, their inclusion would make a nonsense of the procedure.

14. Few studies discuss these sorts of problems openly, but it is usually apparent from the data presented that the decision has been taken to exclude intra-parochial marriages. For example, compare Table 1 (page 124) with figs 3 (p. 130) and 6 (p. 137) of Perry (1969).

15. See R. Watson, 'Measuring Migration' Local Population Studies, No. 21, 1978, p. 61, for a detailed description of how these ratios are calculated.

16. Bradley, pp. 43-4. If you are unsure how to obtain the logarithm of a number, consult a standard textbook. Nowadays, of course, logs are also available as a function on many pocket calculators. You should particularly note, however, that the convention of giving the logarithms of numbers less than 1.0 as 'bar' logs is misleading when used in regression analysis. For example, the log of 0.5714 is normally given as 1.7569. What this means mathematically, however, is $-1 + 0.7569 = -0.2431$. It is this latter figure which must be used to plot figures on the graph. Similarly, the log of 0.0513 is usually given as 2.7101, whereas mathematically this means $-2 + 0.7101 = -1.2899$.

You should also note that the use of logarithms in this way, to 'transform' the data so that they approximate to a straight line, produces what is called double-log graph. There are some problems involved in using these graphs which mean that you should only use them as indicative of the distance effect and in conjunction with other techniques, as in this paper. However, the graph gives a very succinct description of the distance effect, thus enabling comparisons to be made between sets of data, and it is relatively easy to fit and interpret. See P. J. Taylor, 'Distance Decay in Spatial Interaction', Concepts and Techniques in Modern Geography, No. 2, 1975.

18. Bradley, pp. 36-9. Notes, that the particular measure used in this article is the product-moment correlation coefficient, and, as Bradley states (p. 39) the method is given in the standard textbooks.
21. An examination of the evidence presented in Peel, Constant, and Kuchemann, et al., (see Note 3) also gives some very tentative support to this observation, although the way the data are presented in these papers does not allow any firm conclusions to be drawn.
22. J. T. Krause, 'Some neglected factors in the English Industrial Revolution', Journal of Economic History, Vol. 19, 1959 (reprinted in M. Drake, Population in Industrialisation, London, 1969): '... in England it is relatively certain that the marriage rate rose sharply in the late eighteenth and early nineteenth centuries and then fell in the 1830s. The suggested causes of this development are many: early industrialisation, with its child labour, the Poor Laws, enclosure, and mining,' (p. 106). See also page 109, which mentions 'cultural disorganisation' as a probable cause.