VOLUME 44, No. 2 APRIL 1945

STUDIES ON INFANT MORTALITY

PART I. INFLUENCE OF SOCIAL CONDITIONS IN COUNTY BOROUGHS OF ENGLAND AND WALES

BY BARNET WOOLF, M.A., PH.D. AND JOHN WATERHOUSE, B.A.
Zoology Department, University of Birmingham

(With 8 Figures in the Text)

CONTENTS

1. INTRODUCTION

An infant, in vital statistics, is a baby less than a year old. The Infant Mortality Rate (I.M.) of a community is the proportion of babies born alive who die before reaching their first birthday. The I.M. for any year is usually calculated as follows:

$$\frac{\text{Infant deaths during year}}{\text{Infants born alive during year}} \times 1000.$$

Besides the rate for all babies, separate figures are often given for boys and for girls. Boys are usually less viable. The total figure can be analysed by apportioning the deaths into categories, and calculating the rate per 1000 births for each category. Thus we may divide the deaths according to cause as shown on the death certificate, giving for example the death-rate for single causes such as pneumonia or premature birth, or for groups of causes such as infectious diseases or congenital defects. The criterion may be age at death, i.e. mortality, per 1000 births, during the first day, week, or month of life, or in any subsequent age group. Deaths may be divided up according to the calendar month or quarter in which they occur; and other modes of classification are in use. We may also apply two or more means of analysis simultaneously, as in tables giving mortality rate by cause in each age group, for all babies and for each sex separately.

Continuous records of infant mortality in England and Wales began more than a century ago, and for many years past the Registrar-General has given detailed analyses by age, sex and cause for every local government area. There are also valuable decennial reports relating I.M. to the occupation and social class of the father, as well as a number of special and local investigations. Some other countries have records comparable with those of Great Britain, and most governments issue figures of some sort. There is thus ample raw material for a statistical investigation of infant mortality.

Nor should there be any question of the value of any results we might attain, if they can prove to be of even the slightest assistance to practical workers in preventive medicine. In the 21 years between the two great wars, 1919–39 inclusive, just under a million infants died in England and Wales, including about 400,000 potential mothers. To these we should add well over half a million still-births. If these deaths were unavoidable, one might face them with mournful resignation. But there is a strong *prima facie* case for believing that a large proportion of infant deaths are preventable, and no other vital index approaches I.M. in variability or in sensitivity to social conditions. In 1938 the I.M. for England and Wales as a whole was 53, lower than ever previously recorded. In London boroughs, it varied between 42 for Greenwich and 84, exactly double, for Holborn. In county boroughs, the range was much greater still, from 28 in Bath to 99 in Wigan. In 1930–2 the rate for babies whose fathers were in the Registrar-General's Social Class I (employers and higher grade professional men) was 32·7; for Class V (unskilled workers) it was 77·1. In Europe, in 1937, the extremes were Holland 38, and Roumania 178; while for European islands there was an even wider range, from 33 in Iceland to 243 in Malta. On a world scale, New Zealand came out best, with a figure for the white population of 31, though the Maori population fared much worse. The largest recorded figure was 241 for Chile, but there is good reason to suspect even higher mortalities in countries too backward to issue statistics. These enormous differences are not capricious. Every authority agrees that infant mortality falls steeply with improving standards of living, housing conditions, education, public health and sanitation.

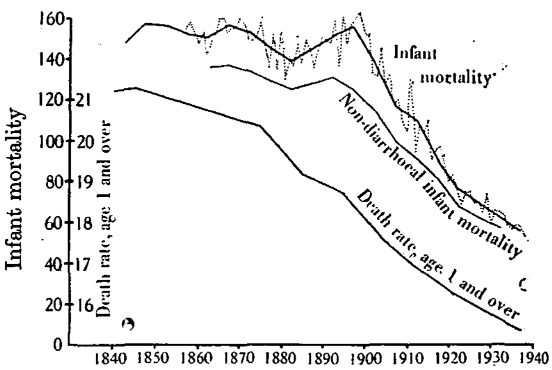

Fig. 1. Infant mortality and death-rate at ages over 1. A hundred years' experience in England and Wales.

Although a falling I.M. is in general closely associated both with falling death-rates in other age groups, and with rising standards of social prosperity, there are important discrepancies for which no explanation is available. It was to elucidate such discrepancies that this inquiry has been undertaken. Why is it, for example, that while general death-rates have been falling and the expectation of life increasing in England and Wales since before 1850, the infant mortality remained roughly constant at around 150 all through the nineteenth century, and then started falling steeply from 1900 onwards? The relevant data are plotted in Fig. 1. The full line at the top shows infant mortality in 5-year averages, with the yearly rate on the dotted line around it. The lowest line gives the 'life-table death-rate' for all persons who survived infancy, being 1000 divided by the expectation of life at age 1 year. The vertical scales have been adjusted so that the two graphs cover roughly the same range. The height above the base-line correctly represents infant mortality, but the general death-rate scale starts at 15. The diagram shows that since 1840 the I.M. has decreased by about two-thirds, while the death-rate above 1 year of age has fallen a little over one-quarter, corresponding with an increase in the expectation of life at age 1 from 47 years to 65 years. But while the general death-rate has been falling all the time, infant mortality remained high until after 1900. Indeed, between 1880 and 1900 it actually increased, and the record high figure of 163 was returned in 1899. There are three widespread misconceptions arising from these facts. It is necessary to clarify them at the outset.

First, it is quite widely believed that the fall in infant mortality can be largely attributed to reduction in the death-rate from diarrhoea. This is a hang-over from past medical anxieties and campaigns. From 1880 to 1900, diarrhoea rose sharply, from a mean yearly death-rate of 14 per 1000 births in 1881–5 to 31 in 1896–1900. A hot, dry summer was a thing to be dreaded, as in 1899, when the diarrhoeal death-rate was 40, one-quarter of the total infant death-rate in a record year. During the present century, until 1921, the peak years shown in Fig. 1 were diarrhoea years. Thus in 1911 it rose to 14. Since then it has always been below 8, and is now about 5, whether the summer is hot and dry, or wet and cold. This is of course a notable achievement. It is true also that a rate of 5, though a great improvement on 30 or 40, is intolerably high when New Zealand, whose diarrhoea mortality was comparable to ours during the nineteenth century, has now brought the figure down to 0·6. But it is not true, either that diarrhoea was mainly responsible for the high infant mortality of the last century, or that its partial conquest is the chief explanation of the improved position of to-day. This is illustrated by the middle curve in Fig. 1, which gives the infant mortality, in 5-year averages, from all causes other than diarrhoea. Although the peak around 1899 is flattened, the non-diarrhoeal does not differ markedly from the total mortality curve. Indeed, it is a gross over-simplification to attribute the changes to any single disease group. Both the steady state of last century, and the improvement during this, mask rises as well as falls in the various certified causes of death. The analysis of these will be far from simple. We hope to deal with this topic in subsequent contributions to this series.

A second misconception is more serious. It arises from the gratuitous and incorrect assumption that all observed differences between people, whether of physical characteristics, health, mentality, morality, or cultural patterns, are solely or mainly determined by 'inheritance', and that changes of social environment can do little or nothing to alter such visible manifestations of the inborn nature of the various 'human stocks'. From this general theory have been derived several social doctrines, varying, according to the aspect most strongly emphasized, from militant eugenics to Nazi racialism. Such doctrines, of course, claim to rest on a scientific basis. But they have been so assiduously spread by propagandist bodies that they are believed in and acted upon by multitudes of people who know nothing of the evidence regarded as convincing by their originators. We do not intend to deal with the general implications of this issue in this communication; but it is necessary to take cognizance of its existence in order to understand a widely held prejudice about infant mortality itself.

The question is, do social measures such as slum clearance, extra milk, health visitors and the like, designed to reduce infant mortality, have a beneficial effect on the *after* health of all babies by giving them a better start in life? Or, on the contrary, do such measures merely postpone and pile up later trouble by interfering with the selective effect of weeding out weaklings? Those who believe that a person's state of health and length of life are almost exclusively fixed by heredity necessarily incline to the second view, which was, in fact, strenuously advocated by Karl Pearson and his pupils.

Although Pearson disclaimed any opposition to social reforms, those chiefly responsible for framing preventive policies obviously felt that his views were obstructing their efforts, and in the first two decades of this century the issue was very hotly contested. The controversy now seems to be dormant; but there was no generally accepted decision, one way or the other. In any new examination of infant mortality, the matter must therefore be reopened, partly because it makes all the difference to one's interpretation of the statistical results, partly because new evidence has accumulated since 1920, and because it is still as true to-day as when Newsholme wrote the words in a Government report of 1910:

> Attempts to reduce infant mortality are regarded by many as an interference with natural selection, which must be inimical to the average health of those surviving. According to this school of thought, efforts to save infant life merely prevent the 'weeding out' of the unfit, and ensure the survival of an excessive proportion of weaklings.

The relevant work of Pearson's school is contained in three papers. The main one (Pearson, 1912) showed that rising infant mortality between 1880 and 1900 in England and Wales as a whole was correlated with falling mortality in the age group 1–5, and in adult life. Pearson concludes 'that a heavy death-rate does mean the elimination of the weaklings'. Snow (1911) came to a similar conclusion from a study of earlier figures, both for England and Prussia. But his results held for rural districts only. Town populations did not fit in. Using an elaborate statistical technique, the 'variate difference' method, which they later (Pearson & Elderton, 1922) admitted was not applicable to the problem, Elderton & Pearson (1914) tried to show that conclusions drawn from a rising I.M. and falling child mortality still held during the new century when both rates were rapidly falling. There is now no point in going into the calculations of Pearson's school in detail, nor in considering the formidable and, in our opinion, decisive objections raised, as for instance by Brownlee (1917), to their work both on statistical and on medical grounds. The data accumulated since 1914 are a much more satisfactory check. With respect to the course of infant mortality, it is clear from Fig. 1 that the period

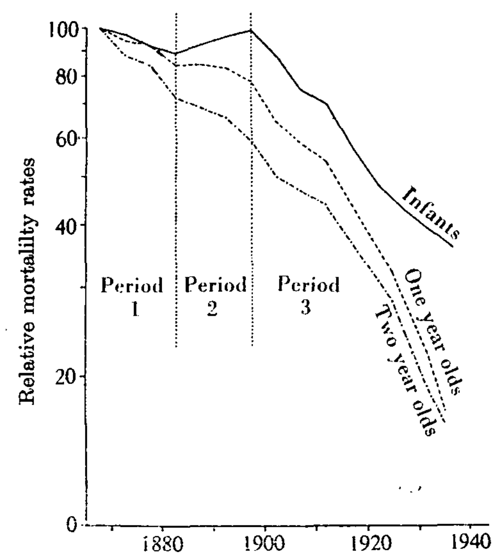

Fig. 2. Changes in infant mortality in England and Wales compared with death-rates at ages 1–2 and 2–3. Logarithmic scale, rates in 1866–70 taken as 100.

1880–1900 was quite exceptional. The relevant facts are set out in Fig. 2, showing infant mortality, death-rate at ages 1–2, and the rate at ages 2–3, averaged over short periods, since 1865. The graphs for ages 3–4 and 4–5 are so close to that for age 2–3 that they would be difficult to disentangle if put into the figure. The vertical scale is logarithmic, so that the slope at any point correctly represents the proportional rate of change. The absolute figures for the three age groups are of course very different, so to facilitate comparison all rates appear as percentages of the value in the period 1866–70.

The data fall naturally into three periods, as shown. In Period I, from 1866 to 1880, infant mortality was slowly falling; in Period II, 1881–1900, it was slowly rising; and in Period III, from 1901 onwards, it has fallen rapidly at a remarkably constant rate. If, as did Pearson (1912), one considers only Period II, completely disregarding the data before and since, one may of course argue that the association of a rising infant mortality with falling death-rates in subsequent age groups proves that high infant mortality has weeded out the weaklings. It is also easy to see that during Period I, though the death-rates were varying together for the country as a whole, they may well have been moving in opposite directions for some sections of the population. If, following Snow (1911), one seeks out those areas that give the desired answer, ignoring those that contradict, one may produce figures seeming to support the Pearsonian thesis. But if one tries to give due weight to the facts as a whole, it seems to us that one is forced to the following interpretation. During the 70 or so years covered, there has been a remarkable fall in infant mortality, accompanied by an even more spectacular fall in mortality rates in early childhood. Since the change of genetic constitution of the people during this period must have been negligible, the whole improvement can be attributed to such things as better housing, feeding, mothercraft and sanitation, reduction of the virulence of pathogens, and so on—all those agencies covered by the word 'environment' when it is used in opposition to 'heredity'.

Improvement of environment has been progressive, and is still going on. The 2-year-olds were the first to give a striking response, then the 1-year-olds, and lastly, after a long time-lag, the effect of better conditions began to show in the infant rate. Rapid decline of infant mortality during Period III has been accompanied, not by a slowing of the fall in the higher age groups as demanded by Pearson's theory, but by an acceleration. At least in part, this can reasonably be attributed to the improved stamina of the entrants to the 1–2 and 2–3 age groups, due to the healthier circumstances of the year of infancy which also enabled a greater proportion of infants to survive. During Period II, while infant mortality was rising, there was a striking check to the downward trend of the 1–2 death-rate, and a smaller though definite check in the 2–3 death-rate. This must have been due to two superimposed sets of circumstances. The special deleterious conditions which were raising infant mortality were also masking the otherwise improving environment of the older babies; and the higher age groups were being recruited from infants enfeebled by a less healthy environment in infancy. Pearson's theory involves the conclusion that, if infant mortality had been falling during Period II instead of rising, mortality at the higher age groups would have risen, or at any rate fallen less rapidly. The explanation here advanced involves the opposite conclusion, that a falling infant mortality would have been accompanied by an even steeper fall of death-rates among the 1- and 2-year-olds. In Periods I and III, we have 'controls' for Period II of a kind rare in vital statistics. Can anyone studying Fig. 2 doubt that the conditions which made infant mortality rise during 1880–1900 also made children of 1 or 2 less healthy, and not more healthy, than they would otherwise have been?

There is more at stake in this matter than laying the ghost of ancient fallacies. As improved social and hygienic conditions bring about a reduction of infant mortality, two opposite effects are possible on higher age groups. One is the Pearson effect—lowered fitness and higher death-rates owing to the survival of congenital weaklings who would be better weeded out. The contrary process was so strongly urged by Newsholme, in opposition to Pearson, that we may call it the *Newsholme effect*. The healthier conditions that express themselves in reduced infant mortality may so improve the health of the survivors that there will be greater fitness and lower death-rates all through the span of life. The data shown in Figs. 1 and 2 indicate strongly that the Newsholme effect is now, and in accordance with present trends will continue to be, overwhelmingly stronger than the Pearson effect, if indeed this effect is real. We express the doubt advisedly. However much we improve environment, it may well be that infancy will remain the most dangerous period of life until old age is reached, so that those who survive cannot be considered 'weaklings' in relation to the environment prevailing at the time. This is at present merely an academic question which only the far future can decide. Meanwhile, it is safe to say that for the present, and for quite a long time ahead, we can diligently strive to reduce infant mortality with full assurance that the benefit in subsequent age groups will far outweigh any possible harm.

Another aspect of the Pearsonian theory has to be considered, especially as we now have ample data by which to judge it. If excess mortality is primarily an expression of inherited weakness, sections of the population whose death-rates are higher must necessarily be 'inferior stocks'. It is, of course, well known that mortality is higher among the poorer than among the better-off. Therefore poverty denotes congenital inferiority, and the poor should, according to taste, be discouraged from breeding, sterilized, or (Fisher, 1930) taxed to provide the children of the rich with family allowances proportional to the parental income.

Now we have reasonably good data showing the change of infant mortality in relation to social class over a period of 20 years. Using the census returns of 1911, 1921 and 1931, the Registrar-General (A.R. 1911; D.S. 1921, 1931) grouped men in England and Wales, according to their occupations, into a limited number of social classes. In 1921 and 1931 there were five. Class I contained the upper and middle classes, Class III skilled labour, and Class V unskilled labour. The other two were intermediate, Class II including roughly the lower middle class and Class IV semi-skilled labour. In 1911 there were three additional classes, textile workers, miners, and agricultural workers being omitted from the general groups and given separately; and there were other differences of tabulation. Nevertheless, the Registrar-General (D.S. 1921) regards Classes I–V of 1911 as reasonably well comparable with those of 1921 and 1931. In each case, the mortality rate for legitimate infants of fathers in the various social classes was computed, giving figures for 1911, for 1921, and the mean for the triennium 1930–2.

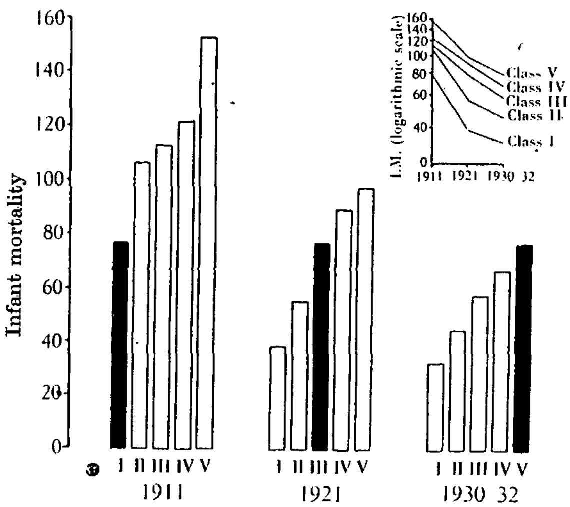

Fig. 3. Infant mortality by social class of father in England and Wales.

The data are shown in Fig. 3. The vertical rectangles give the infant mortality by classes for each census year, and the smaller inset indicates the change over the period in each social class.

Two conclusions stand out. In any given year infant mortality goes up steeply as parental standards of living decrease; and between 1911 and 1931 infant mortality was approximately halved in every social class. We want to draw particular attention to the three rectangles shown in black, namely, those denoting Class I in 1911, Class III in 1921, and Class V in 1930–2. The I.M. values represented by these rectangles are 76·4, 76·9, and 77·1, almost exactly the same. In 10 years, Class III caught up with Class I. In 20 years, children of unskilled labourers caught up with children of employers and professional men. Since genetic changes within the social classes during a period less than one generation must be comparatively trivial, the whole effect must be attributed to improvement in environment. Of course, the environment in the home of an unskilled labourer in 1931 was not exactly like that in the home of a professional man in 1911. To a large extent the poor were still living in overcrowded and insanitary houses, and (Orr, 1935) were unable to buy an adequate diet. These continuing handicaps must have been counterbalanced by such things as more adequate public health measures, more efficient choice of foodstuffs, better mothercraft. In these respects the poor of 1931 must have been considerably better served than the rich of 1911, since, in spite of less money and bad housing, their babies did as well.

Now this is a very surprising conclusion. Though there was obviously a great deal of social improvement between 1911 and 1931, it was not a period of big or dramatic social changes. Yet we must conclude that the environment of a navvy's baby in 1931 (I.M. = 61·5) was slightly better than that of a barrister's baby in 1911 (I.M. = 63). But those who accept the arguments of the Pearsonian school must go further than this. For, in postulating that the relevant environment of Class I in 1911 was equal to that of Class V in 1931, we have made an important tacit assumption. Our conclusion is true only on one condition, i.e. that there is no systematic congenital difference in stamina between babies of Class I and Class V. The Pearson school hold that babies of Class V are more easily killed off. Under conditions where the babies of the rich would have an I.M. of 77, they assert that Class V infants would have a significantly higher mortality—87, or 97, or perhaps 107. Since the environment of the poor in 1931 was able to reduce I.M. to the 1911 level for the rich, and that in spite of the poor child's congenital inferiority, the Pearson school should be even more zealous than their opponents to improve the conditions of the poor. If the relatively slight social changes between 1911 and 1931 could wipe out not only the monetary and domiciliary disabilities of the poor, but the manifest effects of their supposedly hereditary handicaps as well, what might we not expect from really drastic reforms?

So far we have considered only the total infant mortality. The decennial reports also give the mortality of each social class for various causes, as stated on the death certificates. From almost all these causes, environmental conditions—malnutrition, overcrowding, and so on—are known on medical grounds to increase the risk of death; and one duly finds a class gradation of death-rates. There is one cause of death, however, that might be expected to be nearly independent of nurture, and hence to be a measure of the genetic differences between the social classes. This is congenital malformations. Here are the Registrar-General's own comments on the figures. In the 1911 *Annual Report* he writes:

Mortality from congenital malformations is, as might probably be expected, equal for all classes as well as being much the same for illegitimate as legitimate infants.... It should be noted, however, that the mortality of the children of textile workers from this cause is exceptionally high, for it may be that this is related to the fact that the wives of such workers frequently work in the mills themselves throughout the greater part of their pregnancies.

In D.S. 1921 he states:

Mortality from this cause is, as might be expected, very much the same in all classes, though... there is some indication that if the mother is engaged in manual labour during pregnancy the risk of malformation fatal to the life of the infant is increased.

The Registrar-General then goes on to discuss the occupations with high infant death-rates from this cause. These include various grades of textile workers, costermongers, and clergymen, all of them men whose wives are liable to excessive work during pregnancy. He concludes:

If the incidence of mortality corresponds with that of malformation, the latter is practically the same in all classes, but if greater care in Class I prevents death, at least during infancy, from some malformations which would be fatal in Class V, it is possible that these congenital defects may be somewhat more frequent in the upper than in the lower ranks.

In D.S. 1931 the picture is similar:

The incidence of congenital malformations was, as in 1921–3, scarcely influenced at all by social class.

Here therefore there is no evidence of congenital inferiority among the poor. If anything, the suggestion points in the opposite direction. So far as we know, eugenists have not put forward proposals for coping with the possible dysgenic effect of wealth in keeping alive the weaklings who obstinately happen to turn up even in the best families.

It would, in fact, be easy to make a case for the inferiority of the rich more plausible than for that of the poor. From the occupational figures for infant mortality, one could cite not only deaths from congenital malformations, but also those from injury at birth. These are highest in Class I, and decrease with lowering of social class. Unless one is willing to postulate greater risk to a baby delivered by a well-paid medical man rather than by a midwife or medical student, the data point to greater unfitness for child-bearing among well-to-do mothers. In connexion with an assertion made without any supporting evidence by Fisher (1930), that the rich excel in moral qualities, one could make play with the fact that deaths from cirrhosis of the liver and from suicide among the rich far exceed those among the poor, and that the disparity is increasing, as indicative that the rich are becoming more and more disposed to alcoholic excess and more suicidal. One could continue the catalogue. Needless to say, eugenists would explain away such claims by appeal to environmental agencies; and we should not be disposed to dispute their denials. But it is difficult to see why we should discount the disabilities of the rich on environmental grounds, while ascribing the handicaps of the poor to innate inferiority.

In the course of our investigations, we shall not shut our minds to the possibility of congenital differences between classes. On the contrary, we shall be on the alert for any signs. Our object is to attain as complete a quantitative account as possible of all the factors that influence infant mortality. But we refuse to postulate imaginary biological inferiority when known environmental differences are ample to account for the facts. With respect to viability this much is clear. There is no genuine evidence of a differential distribution of genotypes at different economic or social levels. For any quality, the range of individual variation within a class is certainly far greater than the difference between class averages; and wherever relevant observations are available, environmental factors are known to have produced differences greater than those which now distinguish one class from another. If mortality of Class V fell from 152·5 in 1911 to 77·1 in 1930–2 and has probably gone on decreasing at the same rate ever since, we may well ask what there is to prevent Class V from reaching Class I level except lack of Class I amenities and conditions of life? Later in this paper we give estimates of the number of infant deaths associated with a given degree of overcrowding, or male unemployment, or of industrial employment of women. The intransigent eugenist may say that men are unemployed because their genes are bad, or women work in cotton mills because their genes are bad, or families are overcrowded because their genes are bad; and that even if the fathers were given a living wage, the mothers better factory conditions and adequate leave before childbirth, and the families healthy homes, babies would still die at the same rate. We think that any reasonable person will reject this view after examining the evidence which we present.

A third misconception about the fall of infant mortality can be dealt with more briefly. In a word, it is *complacency*. There is no need to single out individuals by specific quotation. Ministerial and political utterances and official reports abound in self-congratulations, linked with hints that no further big improvements are possible. This attitude is as inhibitory to sanitary improvement as the eugenic argument which we have examined. For the proper application of the results of statistical research, the relevant issues are not how far we have improved on a regrettable past. They are:

https://doi.org/10.1017/S0022172400035907 Published online by Cambridge University Press

how many present deaths are preventable, and what are the measures needed to avert them? When the problem is put in this way, there is little scope for relaxing effort. Fig. 3 shows that though infant mortality is falling in all classes, the ratio of the worst class rate to the best is still as bad as ever—over twice as great. The same is true of differences between county boroughs, as shown below. The ratio between worst and best is not decreasing. In 1935, when the infant mortality for county boroughs was 66, the Registrar-General (Text, 1935), after dealing only with the most obvious directions in which improvement is possible, concluded that 'it ought to be possible for every northern town to achieve a rate below 50 and for every other town to achieve a rate below 40. The realization of such rates would mean an annual saving of more than 4000 infant lives in the county boroughs alone.' Since the rate for Class I in 1930–2 was 32·7, the Registrar-General's statement can hardly be called reckless. In this series of papers it is our object to obtain an estimate, as precise as available data allow, of what reduction of infant mortality is possible by specific improvements in social conditions.

2. THE DISTRIBUTION OF INFANT MORTALITY IN COUNTY BOROUGHS OF ENGLAND AND WALES, 1921–38

In the present paper, we consider only total mortality from all causes during the first year of life, deferring to later communications the discussion of deaths from separate causes or at different age levels. Using the figures for England and Wales, our aim is to discover to what extent the variation of mortality between one place and another is associated with differences of unemployment, poverty, housing and other social agencies. Our calculations must therefore be restricted to those places and years for which the necessary social indices are available. The limiting factor proved to be the statistics of local unemployment, which compelled us to confine attention to one class of locality, the county boroughs, for the 11-year period 1928–38.

About one-third of the population live in county boroughs, which during the relevant period numbered 83, being 79 in England and 4 in Wales. London has its own peculiar municipal arrangements, and except for Croydon, East Ham, and West Ham, with a total population (1931 Census) of about 670,000, the Greater London conurbation of over eight million people was not represented among the county boroughs. But outside London, these areas include all the big towns, ranging from prosperous seaside and residential resorts like Bath and Bournemouth to the crowded industrial towns of the north. With three exceptions, all the county boroughs had over 50,000 inhabitants; one, Birmingham, has over a million citizens; and the average population is over 160,000. The range of recorded infant mortalities in individual boroughs during the period studied was very great; the lowest was 28, the highest 129. County boroughs therefore offer very suitable material for a statistical study of infant mortality. Their populations are big enough to reduce the sampling error in their infant mortality figures to a small fraction of the total variance, and

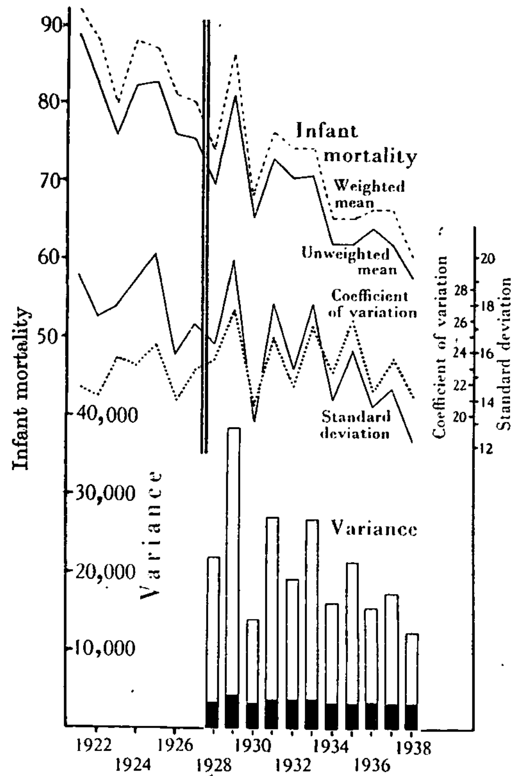

Fig. 4. Infant mortality in county boroughs of England and Wales. Weighted and unweighted means and yearly variance.

they exhibit a wide range of economic, industrial, social and climatic conditions, with a correspondingly big spread of I.M.

Some essential data about the variances to be explained are illustrated in Fig. 4. All the vertical scales have their zeros on the base-line, so that heights about this line correctly represent relative magnitudes. The period 1928–38 is shown to the right of the double vertical dividing line, but the data are taken back to 1921 to show the trend. The top broken line gives the average infant mortality in county boroughs, weighted by the number of births in each area. This graph is roughly parallel to that for infant mortality in the country as a whole, but is on the average about $8\frac{1}{2}$ units higher.

Infant mortality in the county-borough group is consistently above the national rate, approximately the same in other urban areas, and lower in Greater London and in the rural areas. For reasons explained below, we have in this study given equal weights to the county boroughs, whatever their size. The unweighted average is shown as the top full line in Fig. 4. It will be seen to lie below the graph of weighted averages, indicating that on the whole the more populous county boroughs tend to have the higher mortality. The mean distance between the two graphs is under 4 units.

The standard deviation of the infant mortality rate is shown in the lower full line. It fluctuates about the same general downward trend as in the top graphs of infant mortality. The dotted line shows the coefficient of variability, which is

$$\frac{\text{Standard deviation}}{\text{Average infant mortality}} \times 100.$$

The trend of this is very accurately horizontal. On the left of our dividing line the mean I.M. is 80·5, while in the period 1928–38 it is 66·7; the corresponding standard deviation figures are 18·2 and 15·6; but the mean coefficient of variability remained steady, the values being 22·6 and 23·4%. This indicates that as infant mortality falls, the relative disparity between the best and the worst county boroughs remains unchanged. Fig. 3 leads to the same conclusion about the disparity between the social classes. There is a remarkable regularity about the short-term fluctuations of variability. During 1928–38 there was a regular 2-year periodicity, the odd-numbered years showing high variability between county boroughs and the even-numbered years low variability. This comes out clearly in both graphs. During 1921–7 the regularity is not so well marked. Each graph shows one exception. But whereas the exception among standard deviations was in 1923, that among coefficients in variability was in 1927. We may therefore presume that the circumstances producing the 2-year cycle were acting in this period also. This rhythm looks at first as if it may be connected with the 2-yearly measles cycle. But this cannot be so, because the bad measles years are even-numbered, coinciding with our years of low variability. The years of high variability for infant mortality are also peak influenza years. These two facts may be connected. The Registrar-General (Text, 1934) gave evidence of a positive correlation between the death-rates for influenza and for premature birth. But we propose to defer attempts to account for the alternation until the completion of our studies on the individual causes of death. The rectangles at the bottom of Fig. 4 show the variance, that is, the sum of the squared deviations of infant mortality from the yearly mean. The highest variance, in 1929, is more than three times as great as the low value of 1938. The solid black portion in each rectangle represents the sampling error variance of the I.M. For each county borough in each year the sampling variance was calculated from the formula

$$\frac{\text{I.M.} \times (1000 - \text{I.M.})}{\text{Total live births}}.$$

The results were summed in years. If the yearly totals were equal to the variances actually observed, the differences between places could be entirely ascribed to chance fluctuations. Since the observed variances are much greater, the differences between places cannot be attributed to sampling alone. That being so, the solid areas in the rectangles represent variance explained by sampling error, leaving the clear portions to be accounted for in other ways. The proportion of the total variance which cannot be attributed to chance varies from about 73% in 1938 to 91% in 1929.

We have now performed the operation that is quaintly called 'extracting all the relevant information out of the data', and orthodox statistical etiquette demands that we should henceforth forget about the raw facts and direct our attention solely to idealized normal curves. Nevertheless, although we know the means and standard deviations, we still feel that the data may have something to teach us.

Accordingly, county boroughs were ranked for each year in order of increasing infant mortality. Fig. 5 is a distribution contour diagram, rank being plotted against infant mortality. The broken line at the bottom connects the lowest recorded infant mortalities, year by year. Similarly, the top broken line shows the highest figures. The other lines, reading upward, show the 3rd, 16th, 29th, 42nd, 55th, 68th and 81st ranks. These numbers form an arithmetical series with common difference 13. The whole distribution in any year is thus divided into eight ranges. The outermost ranges cover the lowest two values and the highest two. The other ranges divide the rest of the distribution, omitting the four extreme values, into six equal portions. The 42nd contour line is of course the median.

Two things are clear at a glance. One is the general slow downward trend at all ranks in the distribution. The other is the regularity of the lower outline, as compared with the wild fluctuations of the upper limit. It is possible to lay a straight-edge between the 3rd and the 16th contour without cutting either boundary. The same thing cannot be done at any higher range. We would like to draw attention to the specially bad year 1929. The fortunate places in the lowest range were almost unaffected. As one ascends to the higher ranges, the deviation from trend gets more and more prominent. The same is true of the other bad years.

Places with a relatively low infant mortality are insulated against adverse conditions. Whether the year be favourable or not, they are almost completely undeflected from their steady trend of slow improvement. In the middle ranges of the distribution, the effect of bad years begins to tell. The full brunt falls on the places in the highest ranges of mortality.

Fig. 6 shows an alternative way of plotting the data. The vertical axis represents ranking, with lowest mortality at the top and highest at the bottom. The contour lines are drawn at mortality

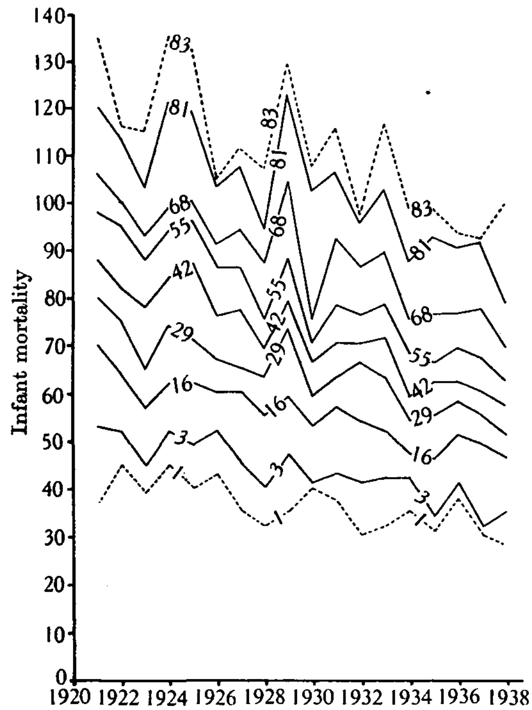

Fig. 5. Infant mortality in county boroughs of England and Wales. Distribution contour diagram showing mortality in each year in the borough 1st, 3rd, 16th, etc., in rank when the boroughs are arranged in order of increasing mortality.

values of 30, 40, 50 and so on, the highest line being at 120. These lines may be called *isothans*. The zone between two successive isothans, say those for I.M. = 60 and I.M. = 70, shows, year by year, how many places had mortalities from 60 to 69, and where they stood in rank compared with other county boroughs.

This figure, like Fig. 5, brings out the relative stability of I.M. in the fortunate places and the fluctuations in towns with the highest mortalities. It gives a clearer picture of the steady trend to improvement. Taking for example the zone between 60 and 70, one sees it travelling downwards across the diagram. In 1921 an I.M. in the sixties range was relatively very good. By 1938 it had become relatively bad, well below the median of the distribution. The lines for 110 and 120 now intrude only in the bad years. During the nineteenth century, they were high up among the best rankings. By 1938 the zone 90–99 is almost eliminated, while a new zone of mortalities under 30 appears in the top right-hand corner.

Many statistical studies have been and continue to be made in which the investigators use mathematical devices such as correlation coefficients, which are valid only if the variates are normally distributed, without any evidence being adduced that this assumption is justified. Under the influence of modern American developments in statistical control in industry (Shewhart, 1939; Simon, 1941) this is coming to be regarded

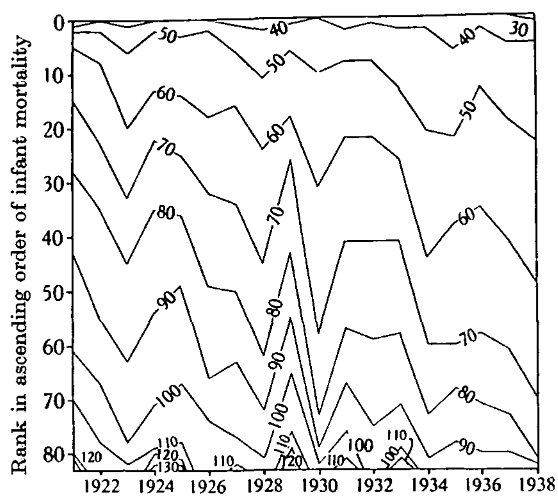

Fig. 6. Infant mortality in county boroughs of England and Wales. Isothan diagram, lines of equal infant mortality showing rank year by year when the boroughs are arranged in order of increasing mortality.

as unsound procedure. At the same time, evidence is accumulating that many statistical methods which in strict theory require the variates to be normally distributed are in fact sufficiently accurate for all practical purposes, even when there are quite marked departures from normality. Modern standard text-books, such as Yule & Kendall (1940, chapter 23), therefore recommend a preliminary inquiry to establish that the distribution is a 'single-humped form not very far removed from the normal'. The position is of course thoroughly unsatisfactory, and further mathematical clarification is urgently needed. It is not only useless, but positively misleading, to undertake elaborate tests of normality such as described in Fisher (1938) when the investigator intends to disregard moderate degrees of skewness and kurtosis, without having any clear guidance from statistical theorists as to the allowable limits of departure from normality.

In our own computations described below, we use the method of multiple regression. As is well known, a regression function does not depend on the frequency distributions of the independent variates. With regard to the distribution of the dependent variable, the standard texts do not make the position at all plain. We think it necessary to distinguish clearly between the two separate aspects of a statistical investigation—obtaining the most plausible values of the parameters, and assessing their reliability or significance. So far as the first process is concerned, there is good reason to believe that normality is irrelevant. So long as certain symmetry conditions are satisfied, the method of least squares seems to lead to correct regression equations even when the distribution of the dependent variate is strongly skewed, rectangular or even U-shaped. This statement is based on the general experience of users of the regression technique, supplemented by some mathematical investigations of our own and numerous experimental cases constructed with the aid of the random drawings given by Shewhart (1931) from normal, triangular and rectangular universes. The whole subject obviously needs further mathematical study.

For the application of the recognized tests of significance, the only necessary condition seems to be that the residual deviations shall be normally distributed round the regression line. If the deviations are also homoscedastic, and the independent variate happens to be normally distributed, then this condition requires the initial distribution of the dependent variate to be normal also. In all other circumstances, normal distribution of the deviations round the regression line will in general be incompatible with normal distribution of the initial values of the dependent variable. Let us consider the most elementary possible case, a simple regression where the values of the independent variate are equally spaced and the residual deviations about the line are normal and homoscedastic. The initial distribution of the dependent variable must then be the summation of a series of similar normal curves laterally displaced, i.e. markedly platykurtic. Since independent variates may have any form of regular or irregular distribution, regression lines may be curved, and residual deviations may not be homoscedastic, it is obvious that the final conditions for a valid test of significance are capable of emerging from any form of distribution of the initial values of the dependent variate. Fortunately, in our computations on infant mortality described below, these highly aberrant possibilities do not arise. Our independent variates, though their distribution is by no means accurately normal, give single-humped curves. If the distribution of our dependent variate is approximately symmetrical, we can apply the regression technique with reasonable confidence that the calculated values of the parameters will not be seriously misleading.

For the valid application of orthodox tests of significance, it is commonly assumed that it is sufficient to establish normality of distribution by some so-called precise method such as that described in chapter 3 of Fisher (1938). In our opinion, this procedure is neither necessary nor sufficient. The test of significance essentially consists in a comparison of the area of the distribution curve outside two predetermined ordinates—the 'tails' of the curve—with the area inside those ordinates. For the purposes of the test, the precise shape of the curve is irrelevant. What matters is that the ratio of tail to body shall be the same as it would be in an ideal normal distribution. Now the usual so-called exact test of normality slurs over precisely this point. Because a distribution has been shown to be sensibly normal in the middle ranges, it by no means follows that this finding can be extrapolated all the way to plus and minus infinity. Indeed, in many biological phenomena, even where the middle of a distribution is very nearly normal, the tails are sparse and truncated. For example, Simpson & Roe (1939) base their application of statistical methods to palaeontology on the general rule that in the distribution of bone dimensions within a species, extreme values *never* exceed 3 S.D. from the mean. In a normal curve, about 4·6 % of the area lies more than 2 S.D. from the mean. If in fact the distribution has more than 4·6 % of its area in the tails, the usual tests of significance are too lenient. If, as so frequently happens with biological variates, the tails are smaller than in a normal distribution, then the tests are too severe. The so-called precise method for testing normality allows excesses or deficiencies in the all-important tails to be swamped by close fitting in the relatively unimportant middle regions of the distribution. It is far more satisfactory to use a simple test of whether the number of observations having the attribute of being beyond the chosen level of significance differ significantly in number from expectation on the assumption of normality.

Our preliminary examination of the distribution of infant mortalities among county boroughs is therefore directed to answer three specific questions. Is the distribution approximately symmetrical? If so, regression methods will probably not be seriously misleading. Is it approximately normal? Does the area of the tails differ significantly from expectation? On the answer to these questions depends the amount of trust we can place in orthodox tests of significance.

The relevant data are shown in Tables 1 and 2. For Table 1, we have calculated for each year the unweighted mean I.M. and the S.D., using the Bessel formula $\sqrt{[\Sigma D^2 (n-1)]}$. The distribution was then divided into groups at intervals of 1 S.D. The last three columns show the number of values below and above the mean, and of those which differ from the mean by more than 2 S.D. The table shows that there is a tendency to skewness in the distribution of I.M. From 1928 to 1938 there are 473 values below the mean, or 51·8 % of the whole, and 8 values in the lower tail (more than 2 S.D. from mean) against 23 values in the upper tail. From 1921 to 1927, however, the asymmetry was in the opposite sense. Moreover, the 473 values below the mean do not differ significantly from the 50 % value of 456·5. The difference between observed and theoretical values is 16·5, against a S.D. of 15. The mean is sufficiently near to the median for the distribution to be regarded as approximately symmetrical. Coming now to the symmetry of the central hump, we note that between 1921 and 1938 there were 488 readings within 1 S.D. below the mean and 473 above. The difference between these figures is 15, and the S.D. of the difference is 25·5. The distribution is sufficiently symmetrical to warrant regression methods. It also happens that the bulk of the distribution does not depart seriously from normality. In a normal distribution, one expects about 68 % of the values to lie within 1 S.D. of the mean. During 1928–38, 584 readings out of 913 were in this zone, giving a proportion of 64 %. During 1921–8 there were 377 out of 581, a proportion of 65 %. This is close enough to normality for our purposes.

Our tests of the tails of the distribution are shown in Table 2. From 1921 to 1927, $15 \pm 3\cdot82$ values were found in the tails, against the expected number of 26·44. The difference between observation and theory is 11·4, giving a t value of 2·99. The departure is highly significant. Similarly for 1928–38, and still more for the whole period 1921–38, the tails are significantly sparser than would be expected in a normal distribution. During the period 1928–38 the area of the tails is only 75 % of expectation. We can therefore make the recognized tests of significance knowing that our results will probably be considerably more reliable than the tests indicate.

We want to make it clear that this preliminary investigation is not to be regarded as establishing the reliability of our computations. The real test of validity comes later, in the internal consistency of the results as a whole and the separate tests of the various parameters. This examination of the distribution is merely an extra precaution, in accordance with the best modern practice, to show that the methods we propose to use are not likely to be seriously misleading. It will we think also have become plain that we have grave doubts about many currently accepted statistical methods and procedures. We naturally, however, want our work to be accepted by statisticians of all schools. We have therefore adopted the most conservative possible policy. In estimating our parameters, we have used no method that is not generally accepted, but we have refrained from using any method, however highly recommended, which we believe may possibly be fallacious, or misleading. We therefore hope that our methods will stand up to the more searching criticism of statistical techniques that is now beginning. On the other hand, in tests of significance we have used the most severe of those currently accepted, even when we ourselves have reason to believe that a more lenient test would be

Table 1. *Distribution of* I.M. *in county boroughs*, 1921–38

			Below mean			Above mean						
Year	Mean	S.D.	−3 S.D. to −2 S.D.	−2 S.D. to −1 S.D.	−1 S.D. to mean	Mean to 1 S.D.	1 S.D. to 2 S.D.	2 S.D. to 3 S.D.	3 S.D. to 4 S.D.	Below mean	Above mean	> 2 S.D. from mean
1921	88·6	19·3	2	12	28	28	11	2	—	42	41	4
1922	82·6	17·5	1	15	27	25	15	—	—	43	40	1
1923	75·8	17·9	1	15	23	29	14	1	—	39	44	2
1924	82·2	19·1	—	19	20	34	7	3	—	39	44	3
1925	82·7	20·2	1	15	23	31	12	1	—	39	44	2
1926	75·9	15·9	1	14	24	31	13	—	—	39	44	1
1927	75·4	17·2	1	12	28	26	15	1	—	41	42	2
Total 1921–7			7	102	173	204	87	8	—	282	299	15
1928	69·4	16·3	1	13	31	20	16	2	—	45	38	3
1929	80·7	21·6	1	16	27	22	15	2	—	44	39	3
1930	65·0	13·0	—	12	28	32	7	3	1	40	43	4
1931	72·6	18·1	—	13	32	22	14	2	—	45	38	2
1932	70·1	15·3	1	16	25	24	17	—	—	42	41	1
1933	70·4	18·0	1	15	25	25	15	2	—	41	42	3
1934	61·6	13·9	—	16	29	23	13	2	—	45	38	2
1935	61·7	16·1	—	14	26	28	13	2	—	40	43	2
1936	63·3	13·2	—	14	30	26	10	3	—	44	39	3
1937	61·6	14·5	3	8	33	21	15	3	—	44	39	6
1938	57·4	12·2	1	13	29	26	13	—	1	43	40	2
Total 1928–38			8	150	315	269	148	21	2	473	440	31
Grand total 1921–38			15	252	488	473	235	29	2	755	739	46

Table 2. *Tests for the tails of the distribution of* I.M.

	1921–7	1928–38	1921–38
No. of values > 2 S.D. from mean	15	31	46
S.D. of above	3·82	5·47	6·68
Values > 2 S.D. from mean expected if distribution is normal	26·44	41·54	67·98
Expected − observed values	11·44	10·54	21·98
t	2·99	1·93	3·29
Ratio of observed to expected values	0·57	0·75	0·68

https://doi.org/10.1017/S0022172400035907 Published online by Cambridge University Press

sufficient. Thus we have used the Fisher (1938) test for the significance of regression coefficients, though its author (Fisher, 1942) has since thrown doubt on its meaning and relevance—doubts which we very fully share. We were able to adopt this conservative policy because our results are not just of borderline significance. We decided at the outset that it was useless to undertake an inquiry of the kind to be described in this and subsequent papers unless it was done on such a large scale that the results were overwhelmingly significant according to any test that could be applied.

3. THE METHOD OF MULTIPLE REGRESSION

In order to estimate the influence of social conditions on the distribution of infant mortality, we have used the product-moment method of multiple regression. Separate calculations were made for each of the 11 years. If M is the infant mortality of a county borough, and $A, B, C, \ldots$ the numerical values for that borough of the social indices being considered, then we obtain for each year an equation of the form

$$M = K + aA + bB + cC + \ldots \pm d, \quad (1)$$

where $a, b, c, \ldots$ are constants, the regression coefficients, and d is the difference between the observed value of M and that calculated from the regression equation. The mathematical procedure ensures that the constant K and $a, b, c, \ldots$ are so fixed that Σd^2 is a minimum. If D is the difference between any individual value of M and the mean value, then ΣD^2 is of course the total variance of M, while Σd^2 is the residual variance after factors $A, B, C, \ldots$ have been taken into account, and $\Sigma D^2 - \Sigma d^2$ is the variance in M associated with the regression factors, the *covariance*. In the course of calculating the regression equation one obtains the value of R, the multiple correlation coefficient between M and the independent variables. The ratio of the covariance, V_c, to the total variance, V, is identical with R^2. It is convenient to have a name for this ratio. In American usage it is called the 'coefficient of determination' (cf. Ezekiel, 1930), but we prefer a word that does not suggest that covariation implies causation. We shall call it the *explanation*. Of course, some part of the total variance may be accounted for in other ways than by the regression. For example, the black rectangles in Fig. 4 show the sampling variance, which we shall call V_s. If the variance accounted for other than by the regression or sampling is called V_t, and the residual or unexplained variance is called V_u, then it is clear from the additive nature of variance that

$$V = V_c + V_s + V_t + V_u, \quad (2)$$

where $V_s + V_t + V_u$ is equal to Σd^2 in equation (1). In equation (2) the *total explanation* is the ratio of $V - V_u$ to V. We propose to denote this ratio by the symbol E. The appropriate suffix will denote the various portions into which the explanation can be partitioned. Thus

$$E_s = V_s/V;\ E_t = V_t/V;\ \text{and}\ E_c = V_c/V = R^2.$$

It is, of course, understood that no part of the variance is regarded as being explained unless the appropriate tests of significance have been satisfied.

The mathematical consequences of imperfections or errors of the figures are very different according to whether they affect the dependent or the independent variables. In the simplest case, where the dependent variable, M, is being related to only one independent variable, A, any errors in measuring M, provided they are random and uncorrelated with A, reduce the explanation but do not affect the regression coefficient. Their effect is to increase V_s or V_u and to add a corresponding amount to the total variance V, without altering V_c. Errors in measuring A are much more serious. They do not affect the value of V, but decrease the explanation by reducing V_c, diminishing the regression coefficient and thus distorting the relationship between M and A. Although these facts are pointed out in the standard texts, the practical consequences are far too frequently disregarded. There are frequently many alternative ways of measuring or expressing any given social agency. For example, as described below, there are a large number of available standards of overcrowding. The standard most closely relevant to infant mortality is that which, in conjunction with the other social indices, gives the highest explanation, which can conveniently be judged by the value of R^2 associated with the regression equation. The only way to discover the most relevant standard is to try them all out and select the one which gives the maximum E_c. Not only is this laborious but necessary preliminary investigation usually omitted, but completely inappropriate indices are frequently used, with the result that a relevant agency is grossly underestimated or even reported as not having any significant influence. An example of this is the quite common use of the proportion of domestic servants in the population as an index of the poverty or wealth of an area, as in Stevenson (1921). No very profound inquiry is needed to establish the fact that high proportions of domestic servants are found not only in well-to-do areas, but also in those poverty-stricken places like Newcastle-on-Tyne and Merthyr Tydfil where the staple industries employ mainly men, and there is consequently no alternative outlet for female labour. An index that is high in both the richest and the poorest places naturally leads to the most erroneous conclusions when used as a measure of poverty.

In a multiple regression the effects of inaccurate or inappropriate figures are more complicated. Random errors of the dependent variable still decrease explanation by loading V with extra unexplained variance, without changing the multiple regression equation. In equation (1), their only effect is to increase the term $\pm d$. Errors in any or all of the independent variables reduce the explanation by a decrease in V_e, and the constants in the regression equation are altered. It will happen only extremely rarely that the errors will be so distributed that each of the regression coefficients suffers the same proportional reduction. In equation (1), suppose that for the values of A there are substituted alternative figures A' which are less accurate or relevant, while the values for B and C remain as before. The explanation will be lowered by a reduction in V_e. But this will appear chiefly in a decrease in the regression coefficient a' as compared with the original coefficient a. The values of b and c will be much less affected, and that very often in a positive direction. In other words, if one is using a regression equation to estimate the relative contributions of the various social agencies to the total covariance, then the influence of any agency which is wrongly or inappropriately measured may be very much underestimated, and that of more accurately measured agencies may be grossly exaggerated.

We therefore considered it necessary to make extensive preliminary inquiries and calculations in order to obtain the most relevant available measure of each independent variable. But even this is not sufficient. The reliability and appositeness of the figures will still vary greatly among the different social agencies. In interpreting the results, it is necessary to take this into account. For example, we give below our reasons for believing that our measures of unemployment and of badly paid occupations are less reliable than our index of overcrowding. If this is so, our regression equations will tend to underestimate the number of infant deaths that could be averted by the abolition of unemployment and low wages. The need for such a critical appraisement of conclusions is not lessened, but rather made more imperative, by the spurious atmosphere of precision conferred on the data by elaborate methods of statistical treatment. Nor can this realistic assessment come from mathematical knowledge or facility. It requires social information and analysis, an examination of how the indices were secured, and what social processes connect the phenomena whose interaction is the subject of the inquiry. Fortunately, we can be confident that the covariance will not be overestimated. Since it is impossible to obtain perfectly accurate and relevant measures of all the agencies, there must be some loss of explanation. It is possible in a regression covering only a single year that chance deviations may combine to give a fictitious increase in explanation. That this could happen in all 11 years is so unlikely as to be negligible. The relative contributions of the different social variables to infant mortality may not be correctly assessed. But we can be sure that our estimate of their total influence when acting jointly is a minimum figure, possibly considerably less and certainly no more than the result that would be obtained if all the indices reached the ideal limit of perfect accuracy and relevance.

4. INDICES OF SOCIAL CONDITIONS

Since it is obviously out of the question to try out all the versions of the social indices in every possible combination for each of the 11 years, we made a large number of exploratory calculations of correlation coefficients and regression equations. About 30 different indices were calculated and tried out in many combinations. We are fairly certain that the five indices finally selected give as high E_c values as possible. There is no point in going over this work in detail. We have extracted and arranged in logical order those results that seem of most social interest.

For reasons given below, we had to omit one county borough, Bootle, from our calculations. All correlation coefficients and regression equations are based on the figures for 82 county boroughs, and were done by the methods described in a later section.

(a) *Overcrowding*

There is an embarrassingly large choice of possible indices. The housing volume of the 1931 Census gives for each place the proportion of individuals and of families living at densities of more than 1, $1\frac{1}{2}$, 2 and 3 persons per room, as well as the average number of persons per room in the whole community. The figures based on family units seem to be the most relevant in assessing the proportion of infants subjected to overcrowding. A standard of overcrowding often adopted in social investigations is a density of more than 2 persons per room, denoted by A in Table 3. Correlation coefficients between this index and I.M. for five sample years are shown in column A of Table 4. Scatter diagrams showed that the regression is not linear, but falls off at high values for overcrowding. Accordingly, the logarithm of the index was tried, and it gave the appreciably higher correlations in the column 'log A' of the table. The index seems unlikely a priori to be the most relevant. As shown in Table 3, less than 5 % of familes in county boroughs are overcrowded to this extent, and it seems most improbable that the housing conditions of over 95 % of families are ideal for infant welfare. About 13 % of families in county boroughs live more than $1\frac{1}{2}$ persons per room. The correlations of I.M. with this index, shown in column B of Table 4, are appreciably higher than those for index A. The scatter diagrams still show departure from linearity, and improved correlations

https://doi.org/10.1017/S0022172400035907 Published online by Cambridge University Press

were obtained by the use of log B instead of B. We next tried index C, families with more than 1 person per room, which covers about 28 % of families. On the whole, this gave higher correlations than any yet obtained. Scatter-diagrams showed no sign of departure from linearity. Correlations between I.M. and persons per room are shown in the last column of Table 4. They are remarkably close to the results obtained with index C. The two indices run so nearly parallel that we measured the correlation between them. It is 0·9837. The two indices, calculated in different ways and expressed on different scales, are to all intents and purposes identical. We chose to use C rather than persons per room because the taken under the Ministry of Health in 1936. This investigation had the practical object of discovering the number of new houses required in each area to abolish 'overcrowding' as defined in the Housing Act of 1935. The Survey, in our opinion, was defective both in principle and in method of execution. In its planning it shared the two major defects which enormously reduced the statistical value of such elaborate investigations as the New Survey of London Life and Labour and the Merseyside Survey. In the first place, instead of covering the whole population, it was confined to 'working-class houses', with an arbitrary and ambiguous set of criteria of which houses should be included and which omitted.

Table 3. *Housing indices—county boroughs*

	Weighted mean	Lowest	Highest
1931 census:			
Families with more than:			
A 2 persons per room	4·7 %	0·8 %	19·2 %
B 1½ ,,	12·9 %	4·3 %	35·0 %
C 1 ,,	28·2 %	14·1 %	52·2 %
1936 Survey:			
D Families below standard	4·2 %	0·3 %	20·6 %
E Families at or below standard	7·6 %	1·5 %	29·9 %
Persons per room, 1931	0·86	0·64	1·23

Table 4. *Correlation of housing indices with infant mortality*

	A	log A	B	log B	C	D	log D	E	Persons per room
1928	—	—	—	—	0·5875	—	—	—	—
1929	0·5592	0·6683	0·6172	0·6670	0·6823	—	—	—	0·6931
1930	—	—	—	—	0·5845	—	—	—	—
1931	0·6005	0·6720	0·6429	0·6681	0·6936	—	—	—	0·6909
1932	—	—	—	—	0·6244	—	—	—	—
1933	—	—	—	—	0·6556	—	—	—	—
1934	0·5864	0·6652	0·6319	0·6718	0·6742	—	—	—	0·6818
1935	—	—	—	—	0·6808	—	—	—	—
1936	0·5105	0·6103	0·5467	0·6025	0·5980	0·4725	0·5741	0·5053	0·6253
1937	—	—	—	—	0·6272	—	—	—	—
1938	0·3719	0·4695	0·4251	0·4848	0·4707	—	—	—	0·4734

figures are easier to interpret. An increase of 1 % in overcrowded families has a more obvious social meaning than an increase of 0·1 in the average number of persons per room.

All the indices so far tried were obtained from the 1931 Census, and might be expected to become less and less applicable at years far removed from the census day. In fact, as column C of Table 4 shows, there is no indication from correlation coefficients that there was any decrease in relevance except in 1938. This may indicate that the efforts at rehousing during this period were doing no more than keep housing conditions from getting worse, without effecting any appreciable improvement. But we thought it desirable to try out the indices obtainable from a special Housing Survey under-

Secondly, instead of stating the number of families at various levels of occupation density, one single overcrowding standard was used. This standard, like the poverty line used in the social surveys, was invented a priori and is obviously far below the biological optimum. It was most elaborate, involving counting children as fractions of a person and small rooms as fractions of a room. The collection of information was left to local authorities and was inadequately financed, resulting in wide differences in method and reliability between one place and another. The results represent the percentage of working-class families overcrowded by the Survey Standard. After all this elaboration the figures, indicated by the symbol D in Table 3, correspond quite closely in range with those obtained at the

2-persons-per-room level in 1931. If one includes those families which just reach the standard, E in Table 3, the percentage of families covered rises to 7·6%. The correlations between I.M. and D, log D and E for 1936 are shown in Table 4. They are distinctly lower than those for A, log A and B in the same year. As the 1936 Housing Survey indices gave lower relevance even for 1936 than the Census figures obtained in 1931, we thought it useless to test them any further.

It is possible that the indices more highly correlated with I.M. may also be so much more highly correlated with the other independent variables that they give less explanation in a multiple regression than apparently less promising indices. We made enough trial regression calculations to convince ourselves that our index C really retains its superiority when combined with other social indices.

Because of its high correlation with I.M., the linearity of the regression, and the ease of interpretation, we chose index C as the best available single measure of overcrowding in its relevance to infant mortality. Henceforth we will use the letter H to mean the percentage of families in any county borough reported in the 1931 Census to be living at a density of more than one person per room.

(*b*) *Unemployment*

The only available source of data is the *Local Unemployment Index* issued monthly by the Ministry of Labour to private subscribers from January 1927 until the outbreak of war in 1939. We have to thank the Statistical Officer of the Ministry for kindly lending us his file copy. The *Index* did not profess to give anything more than very rough figures. The data refer to Labour Exchange areas, which do not necessarily coincide with municipal divisions. For each area, figures are given month by month of the percentage of insured workers unemployed. Until the end of 1936, men, women and juvenile percentages were given separately, but in 1937 the classification was changed to males and females. The number of workers upon which the percentage was based was always the estimated total of persons in the area insured against unemployment in July of the previous year. The percentages were thus subject to a change of base every January, so that one year's figures are not strictly comparable with those of another year. The figure used for unemployed workers was the number signing at Labour Exchanges within the area on a specified date in the middle of each month. As a great many workers signed at an Exchange in the area in which they worked rather than where they lived, the figures only very roughly corresponded with the number of resident workers unemployed. There were other sources of inaccuracy. In some cases, unemployed workers were drafted to training centres away from their usual places of residence. They were counted as unemployed in the town containing the centre, which thus acquired a fictitiously high percentage. The basis of calculation for districts in the Greater London area was radically changed half-way through the record, and there were changes also for other places.

We had hoped to include in our calculations all the major municipal areas, not only county boroughs but also municipal boroughs and administrative counties. But examination of the *Index* quickly showed that Labour Exchange areas did not correspond even approximately to municipalities except in county boroughs. In most cases the county borough was named in the *Index*. In a few instances a different name was used, e.g. North Shields could be approximately identified with the county borough of Tynemouth, and Canning Town and Stratford with the borough of West Ham. We carefully went through the list with the aid of Phillip's *Administrative Atlas of England and Wales*, and succeeded in finding an explicit or disguised entry for every county borough with the exception of Bootle, which was presumably swallowed up for Labour Exchange purposes in its larger neighbour, Liverpool. It seemed to us that the unemployment percentage most likely to be relevant to infant mortality was that for men. We therefore copied from the *Index* the figures month by month for percentage of men unemployed, except that for 1937 and 1938 we had perforce to use data for males, which included men and boys.

For one day during our period there is a set of figures available which is theoretically free from the major defects of the *Index* data. The 1931 Census gives for each area the number and percentage of occupied males and females who stated they were out of work on the census date, 28 April 1931. These figures should give a true picture of unemployment by locality, covering all residents rather than only that section of the population subject to state insurance, classified by place of domicile and not of registration. Comparison of the Census and *Index* figures showed general correspondence, but frequent discrepancies. For instance, the *Index* for April 1931 gave Sunderland and Gateshead the similar figures of 47·4 and 45·1%, while the Census values for male unemployment were 36·6 and 27·2%. There are good indications that the Census data are not as reliable as one might have hoped. Although they should include all the unemployed, and not only those coming under state insurance, the Census unemployment total for England and Wales is less than that given by the *Ministry of Labour Gazette*. Presumably many persons who were out of work omitted to state the fact in their census returns. There is no means of finding out whether this source of error affects all places equally, in which case

https://doi.org/10.1017/S0022172400035907 Published online by Cambridge University Press

it is not very important, or whether it is highly localized.

It seemed worth while trying to improve the *Index* figures by combining them with the Census information. For this purpose we calculated a factor for each place by the formula

$$\text{Factor} = \frac{\text{Census unemployment, April 1931}}{[\text{Mean of } \textit{Index} \text{ unemployment, April and May 1931}]}.$$

For example, in Chester the Census male unemployment was 12·4%. The *Index* unemployment in mid-April and mid-May was 16·3 and 16·5, giving a mean of 16·4. The factor is 12·4/16·4 = 0·756. If we multiply the *Index* unemployment in Chester at any other time by 0·756, we will obtain the unemployment on the Census basis, provided we can assume an unchanged ratio of insured workers to total occupied males, and of unemployed workers signing at Chester Exchanges to unemployed Chester residents. Figures adjusted in this way were computed for each of the 82 county boroughs for the whole period covered by the *Index*. When the basis of calculation of the *Index* unemployment changed for any borough, an appropriate adjustment was made in the conversion factor. Correlations were worked out for each year between the mean *Index* unemployment and the adjusted Census values. The correlation coefficients varied in different years from 0·84 to 0·93, with a mean of 0·88. A more informative measure of concordance is the covariance ratio, which is the square of the correlation coefficient. This varied from 0·70 to 0·86, with a mean of 0·78. This is not a high covariance ratio for two sets of figures which purport to measure the same social phenomenon. Quite divergent results might easily be obtained in a multiple regression according to whether *Index* or Census unemployment figures are used. Such divergences were actually observed.

There is another important decision to be made in relating unemployment to infant mortality. What is the most relevant period of unemployment to take into account? Infant mortality is given for calendar years. Take, for instance, the year 1930. Those babies who die during January 1930 cannot be affected by the degree of unemployment in the following December. On the other hand, the reduced standard of living in the family consequent upon unemployment of the father may affect the baby not only from birth onward, but also during the period of gestation. For a baby just under 1 year old dying in January 1930, one might have to take into account unemployment during the whole of 1929 and the greater part of 1928. We first tried averaging the unemployment over 3-monthly periods, and working out a multiple regression equation between infant mortality and the 12 quarterly figures of unemployment in the same year, the previous year, and the year before—a procedure analogous to the 'discriminant function' technique. But this gave wildly erratic results. A multiple regression with 12 independent variables based on 82 points will frequently give unstable values, especially when the independent variables are as highly intercorrelated as the unemployment percentages in the same borough at different times within 3 years. We next averaged unemployment over calendar years, and sought the best single measure of this kind. We had a choice of six possibilities—the same year, the previous year and the year before that, using either *Index* unemployment or the adjusted Census figures. After a large number of trial regressions, it was clear that on the whole the *Index* figures gave more explanation than Census-adjusted percentages. The similarity between same year and previous year figures was much closer, but on balance the unemployment of the previous year gave the highest explanation when used in combination with the other social indices. We therefore chose as our unemployment figures the percentage of men or of males unemployed according to the Local Unemployment Index, averaged over the year previous to that in which the infant mortality was incurred. We denote this figure by U. For the sake of comparability we used this measure in every equation. But in 3 years out of 10 the adjusted Census figures gave higher explanation, as did same year's unemployment in 5 years out of 11. In view of these facts, as well as the defects in the method of compiling the figures, we must regard our unemployment data as very inaccurate or highly randomized, and so as likely to underestimate seriously the contribution of unemployment to the total amount of infant mortality.

Since the record of unemployment does not go back beyond 1927, and we were correlating the figures with the infant mortality of the following year, the earliest year we could deal with was 1928. The record of infant mortality ceased in 1938, the war-time figures not being made public. That is why our calculations are confined to the 11 years 1928–38 inclusive.

(c) *Social class*

We wished to include some measure of family income, so we naturally investigated the applicability of the Registrar-General's division of the population in the 1931 Census into the five social classes already referred to above. The number per mille of occupied males in each class in all county boroughs is conveniently given by the Registrar-General (Text, 1934). We have compiled two indices, the percentage in Classes I and II, denoted by W

https://doi.org/10.1017/S0022172400035907 Published online by Cambridge University Press

(for well-to-do) and the percentage in Classes IV and V, denoted by P (for poor). The correlation between the two indices is -0.7839. We tried these indices out singly and in combination, in conjunction with the other social measures. If only one were to be used, there was very little to choose between them. When both were used in a multiple regression equation, the decrease in explanation on omitting either was very small, as might be expected from the high correlation between the two indices. In these equations, the regression coefficient of P was as a rule higher than that of W. With some reluctance, we decided to simplify our equations by including only one measure of social class. Accordingly our final regressions contain P, the percentage of occupied males in the 1931 Census considered to be in semi-skilled or unskilled occupations.

This is not necessarily a true measure of the number of low-paid workers. For instance, the most poverty-stricken borough, by this index, is not one of the towns in the depressed areas, but the relatively well-off fishing port of Grimsby, which has a P value of 52 %. This arises because fishermen, who form 13 % of Grimsby's occupied males, are regarded as semi-skilled and placed in Class IV. On the other hand, the hewers and getters, who form the majority of the notoriously low-paid coal miners, are included among the skilled workers in Class III. Thus Merthyr Tydfil, perhaps the hardest hit of all boroughs during the depression, has a P figure of 42·5 %, to which another 24 % would be added if the hewers and getters were put among the poor. Since a baby's chance of survival is much more likely to be affected by the wages its father brings home, rather than by the skill he displays in earning them, we cannot consider P as other than an inaccurate measure of poverty in its relevance to infant mortality. This is, of course, no reflexion on the Registrar-General's classification, which does not profess to be based on relative incomes. We would rather direct attention to that notorious gap in our social statistics, the lack of any reliable information about wages and earnings. We spent a lot of time trying to devise a measure of relative wage-levels in county boroughs, and considered various possible methods of adjusting the occupations included in the various social classes to make them correspond more closely to income gradations. But the available data are so fragmentary and unreliable that we thought it best to keep our index P as it stood, while noting that it is not a very faithful measure of relative poverty.

(d) *Women employed in industry*

The inclusion in our equations of a measure of the industrial employment of women was not part of our original plan. The unexpectedly large influence of this factor emerged from our computations. We were of course aware that factory work by mothers, especially in the textile and pottery areas, was regarded during the first two decades of this century as having an important influence on infant mortality. This is especially stressed by Newsholme (1910, 1913, 1914) in his special reports on the subject. More recently, the Registrar-General, in his *Decennial Supplement for* 1931, attributed high mortality among the babies of textile workers to 'the frequent employment of the wives of textile workers in the same occupations as their husbands'. But we felt that a phenomenon that affected relatively few towns and families would be unlikely to show up in our regressions against the more universal social facts of bad housing, unemployment and poverty. After completing a set of equations relating infant mortality to the indices H, U, P and L, we computed a representative sample of residual deviations, showing in what direction and how much the observed mortality in individual towns diverted from the calculated figures. It then became apparent that factory work by women was probably much more important than we had anticipated. If one took the most poverty-stricken towns, with the highest infant mortalities, some, such as the places in Durham and Northumberland, tended to have lower mortalities than indicated by the equation. In the textile towns of Lancashire and Yorkshire there were very large divergences in the opposite direction. The only obvious difference between the two kinds of borough was the low prevalence of female industrial work in the one group, and the high rate in the other. We accordingly sought a measure of this factor to include in our equations. In the 1931 Census there are no data for employment of married women by both locality and industry. We therefore had to be content with figures for employment of all women, which may be expected to be reasonably highly correlated with the degree to which married women go out to work. Two indices were tried. The first, the percentage of females aged 14 and over who were recorded as occupied, gave insignificant correlations with infant mortality. For three sample years the figures were -0.0058, 0.0461 and 0.0877. On consideration, it seemed plausible that two opposite effects were involved. The index included the number of domestic servants, which on the whole would be negatively correlated with mortality, as well as the number of women in industry, which we expected would give a positive correlation. We therefore constructed a second index by calculating for each place

$$\frac{\text{Females employed in manufacture} \times 100}{\text{Total females aged 14 and over}}.$$

The numerator was obtained from Table 4 of the Industry Tables volume of the 1931 Census. This gives details for each sex of the number engaged in each industry, distinguishing employers and managers, operatives, those working on their own account, and those out of work. We assumed that the unemployed would be mainly wage-workers. We therefore added together the operatives and the unemployed in industrial groups III–XIV inclusive, comprising all the manufacturing and similar trades, but excluding agriculture, distribution, administration, the professions and personal service. The denominator was obtained from the Occupation volume of the Census. We denote this index by F (for females in factories). It gave correlations with mortality ranging from 0·2145 to 0·4111, as shown in Table 5. But its effect on the regression was much higher than these rather low correlations might suggest, because it was also very slightly correlated with the other social indices. Indeed, as shown in Table 6, it was negatively correlated with H, and with U in 9 years out of 11. When added to a regression involving the other four indices, it increased the covariance by an average amount of 14·5%, a substantial contribution to the total explanation.

Table 5. *First-order correlations between infant mortality and independent variables*

	H	U	P	F	L
1928	0·5875	0·4992	0·5827	0·3543	0·5451
1929	0·6823	0·5407	0·6160	0·3565	0·6404
1930	0·5845	0·5995	0·5620	0·2145	0·5950
1931	0·6936	0·6940	0·6050	0·2213	0·5873
1932	0·6244	0·5369	0·5852	0·4111	0·5967
1933	0·6556	0·5959	0·6183	0·2732	0·5158
1934	0·6742	0·5421	0·5239	0·2599	0·5009
1935	0·6808	0·5494	0·5990	0·2335	0·6124
1936	0·5980	0·4511	0·5936	0·2927	0·6211
1937	0·6272	0·5088	0·5267	0·2450	0·6023
1938	0·4707	0·4389	0·4417	0·2583	0·4395
Pool	0·5784	0·4298	0·5253	0·2624	0·5239
Means	0·7347	0·6587	0·6653	0·3368	0·6557

Table 6. *First-order correlations among independent variable*

(1) *Correlations with U* (=unemployment in previous year)

	H	P	F	L
1928	0·6187	0·5227	−0·3501	0·3580
1929	0·5498	0·4892	−0·1371	0·2884
1930	0·5918	0·5041	−0·0676	0·3569
1931	0·5702	0·5380	0·1428	0·4790
1932	0·6874	0·6457	0·0626	0·5224
1933	0·6850	0·5617	−0·1145	0·4923
1934	0·6412	0·5288	−0·1768	0·5181
1935	0·6187	0·5076	−0·1469	0·5046
1936	0·6122	0·4910	−0·1573	0·5060
1937	0·5562	0·4605	−0·1637	0·4460
1938	0·4969	0·3880	−0·1896	0·4005
Pool	0·5236	0·4461	−0·0849	0·3950
Means	0·6378	0·5353	−0·0986	0·4767

(2) *Correlations among other variables*

	P	F	L
H	0·6870	−0·0217	0·5288
P	—	0·0768	0·5007
F	—	—	0·2867

(e) *Latitude*

We included this as an independent variable because the Registrar-General has repeatedly drawn attention to the association of increasing latitude with increasing mortality. Latitude does, indeed, give quite big regression coefficients with infant mortality. But we give below our reasons for regarding the apparent effect of latitude as primarily associated with poverty rather than with climate.

https://doi.org/10.1017/S0022172400035907 Published online by Cambridge University Press

We use L to denote the number of degrees of latitude north of 50° 30′.

These five variables H, U, P, F and L were used in our regressions with M, the infant mortality. A number of other social indices were also tried and discarded for various reasons. These are discussed below.

5. COMPUTATIONAL METHODS

At the outset, we had to decide whether to treat each county borough as a unit of equal value, or whether to assign weights according to the presumed reliability of the figures. It was clear that any method of weighting would so increase the labour of computation as to make it impossible for us with our resources to complete the investigation in a reasonable time. However, that was not the main reason which induced us to give each borough an equal weight. In the first place, there is the difficulty about the appropriate weight to assign. The object of weighting is to minimize the unexplained variance. The accepted method is to weight each reading inversely as the variance of the dependent variable, in this case the infant mortality. But this does not completely achieve the desired result. Places with high initial variance might quite well give mortalities very close to those calculated from the regression, and so have low residual variances; and vice versa. A method for minimizing residual variances could no doubt be devised, but the mathematical procedure would be impossibly complicated. Secondly, it is not only error in the dependent variable that goes to make up the unexplained variance. The errors in the independent variables make a substantial and probably the major contribution, and there is no reason to believe that these errors are parallel to those in the infant mortality. In the unemployment figures, for example, the smaller isolated county boroughs probably give truer returns than the larger places which are part of big conurbations or are surrounded by suburbs, so that the number of insured workers signing on does not correspond at all closely with the number of resident workers.

It is relevant to note that in correlating general mortality with social indices, the Registrar-General (Text, 1934) found that weighting made no substantial difference to the results. If anything, there were indications of a decrease in explanation when weighting by population. The correlation coefficients were:

Thirdly, and most decisively, we are not so much interested in the variances as in the regression equations. Our multiple correlation coefficients are all so high that there is no doubt of their significance. There is therefore no object in weighting, unless we can anticipate that any resultant changes in the regression equation will more truly represent the relations between the variables. But there is good reason to believe that increased weight to the bigger boroughs would give a worse rather than a better regression equation. About 19·5% of the population in county boroughs is contained in the three largest cities, Birmingham, Liverpool and Manchester, and 30% in these plus Leeds, Sheffield and Bristol. These six places would therefore be assigned about 30% of the total weights. But it is precisely in these places that one would expect the least clear-cut results from a correlation of mortality with social indices. They contain within their borders a most heterogeneous collection of districts and strata of the population. The infant mortality in the separate wards of a city like Birmingham shows discrepancies as great as those between the best and the worst of the smaller county boroughs. By relating the average mortality of such a place to the average overcrowding, unemployment and so on, one is liable to obtain a blurred and inaccurate result, as compared with that to be expected from the much more homogeneous smaller towns which are either mainly industrial or, like Bath and Bournemouth, mainly residential. By giving each place an equal weight, we did indeed introduce a slight difficulty in interpreting the regressions in terms of actual infant deaths. But we believe that the social picture is truer than that to be expected from a weighted regression.

For purposes of calculation, each of the variables was coded in a series of whole numbers. Thus an infant mortality of 29 or 30 was coded as 1, 31 or 32 as 2, and so on, giving a total range of 52 steps from 0 to 51. Latitude was coded by steps of 15′, giving 19 steps. All the other variables had well over 30 grades. The grouping was fine enough to avoid the need for corrections for grouping error. The regression matrix was made up from the first-order correlation coefficients, and usually evaluated by the Doolittle method. Our procedure is well described by Love (1936). In addition, we obtained the inverse matrix by adding five columns to the Doolittle solution, substituting for the observed correlations with M the values 1, 0, 0, 0, 0; 0, 1, 0, 0, 0;

		Unweighted (r_u)	Weighted (r_w)	Increase in explanation ($r_w^2-r_u^2$)
Mortality with:	Latitude	0·65	0·72	+0·096
	Persons per room	0·77	0·70	−0·103
	Class IV and V	0·685	0·67	−0·020

and so on, according to the procedure recommended by Fisher (1938). This not only enabled us to compute the standard deviations and significances of the regression coefficients, according to the formula given by Fisher, but also to obtain a check on our solution by calculating the regression coefficients from the inverse matrix. In addition to the eleven equations for the separate years, two summary regressions were computed. In the pool equation, all the 902 mortality figures (82 × 11) were combined, each value of M being associated with the appropriate figure for H, U, P, F and L. The U figure was, of course, the unemployment in the previous year, just as in the separate yearly equations. For the means equation, we calculated the mean mortality of each place, $\overline{M}$, and $\overline{U}$ the mean unemployment for 1927–37 inclusive. The other variables do not of course vary with time. The first-order correlations used in these computations are shown in Tables 5 and 6.

6. REGRESSION EQUATIONS AND PARTITION OF VARIANCE

The thirteen regression equations are shown in Table 7. The regression coefficients in the separate years vary among themselves, but their averages agree fairly closely with the corresponding figures in the pool and the means equations. The values of the coefficients are of course proportional to the units in which the indices are expressed. Because the coefficient of L is four times that of H, it does not follow that latitude is four times as important in explaining infant mortality. If latitude were measured in minutes instead of degrees, its average coefficient would be 0·034 instead of 2·01. Appropriate comparisons of the contributions of the various agencies are given below. The table also shows the residual differences between the actual values of I.M. and those calculated from the regressions. The column headed 'Original' gives the S.D. of the I.M. values about the appropriate mean. In the 'Residual' column are set out the S.D. of the observed mortalities from the values calculated from the regression equations, corresponding to the term $\pm d$ in equation (1) above. The 'Minimum' column gives the sampling deviations, representing the irreducible minimum of error due to the finite number of deaths in each county borough. It will be seen that the regressions have brought the error of estimate quite near to the theoretical minimum.

There is no doubt about the statistical significance of the multiple regression coefficients, which are shown in Table 10. The exceptionally low figure of

Table 7. *Regression equations*

1928	$M=24·2$	+	$0·27H$	+	$1·00U$	+	$0·34P$	+	$0·65F$	+	$0·93L$
1929	$M=17·4$	+	$0·79H$	+	$0·67U$	+	$0·33P$	+	$0·60F$	+	$4·3\ L$
1930	$M=31·1$	+	$0·17H$	+	$0·65U$	+	$0·26P$	+	$0·16F$	+	$3·1\ L$
1931	$M=23·8$	+	$0·67H$	+	$0·66U$	+	$0·22P$	+	$0·17F$	+	$2·3\ L$
1932	$M=24·7$	+	$0·62H$	+	$0·05U$	+	$0·39P$	+	$0·44F$	+	$2·3\ L$
1933	$M=13·3$	+	$0·57H$	+	$0·45U$	+	$0·56P$	+	$0·43F$	+	$0·35L$
1934	$M=25·5$	+	$0·72H$	+	$0·33U$	+	$0·03P$	+	$0·37F$	−	$0·02L$
1935	$M=16·3$	+	$0·63H$	+	$0·26U$	+	$0·31P$	+	$0·26F$	+	$2·6\ L$
1936	$M=24·0$	+	$0·35H$	+	$0·10U$	+	$0·44P$	+	$0·24F$	+	$3·0\ L$
1937	$M=24·8$	+	$0·54H$	+	$0·34U$	+	$0·11P$	+	$0·27F$	+	$2·8\ L$
1938	$M=29·1$	+	$0·24H$	+	$0·49U$	+	$0·22P$	+	$0·29F$	+	$0·53L$
Average	23·1		0·51		0·46		0·29		0·35		2·01
Pool	$M=22·6$	+	$0·58H$	+	$0·24U$	+	$0·34P$	+	$0·32F$	+	$2·2\ L$
Means	$\overline{M}=22·9$	+	$0·51H$	+	$0·44\overline{U}$	+	$0·30P$	+	$0·35F$	+	$1·7\ L$

Standard error of estimate

	Original	Residual	Minimum
1928	15·80	9·92	6·25
1929	22·14	12·29	6·72
1930	12·98	8·83	6·17
1931	17·96	10·88	6·59
1932	15·14	9·61	6·57
1933	17·94	11·85	6·73
1934	13·90	9·44	6·25
1935	15·78	10·31	6·26
1936	13·30	9·22	6·33
1937	14·76	10·26	6·16
1938	11·76	9·51	6·08
Average	15·59	10·19	6·37
Pool	17·05	12·62	6·38
Means	13·38	6·59	1·92

$R^2 = 0{\cdot}3941$ was obtained in 1938. The regression uses up 5 degrees of freedom, leaving 76 for the unexplained variance ratio, 0·6059. The value of Snedecor's F is therefore $0{\cdot}3941 \times 76/0{\cdot}6059 \times 5$, which works out at 9·9. The 1% point for these degrees of freedom is 3·3, so the regression is highly significant. The next lowest R^2, 0·5526 in 1937, gives an F value of 18·8, and the highest, 0·7141 in 1929, an F of 38. The means regression shows an F of 52·5, and the pool, which is based on 902 readings, has an F of 158.

The individual regression coefficients, taken as isolated results, do not all reach the conventional level of significance. But there can be no doubt of their significance when they are considered in their entirety. Two principles of testing are in use. In one, the null hypothesis or external test, one calculates the chances that the results could have come from an uncorrelated universe, making the usual assumptions about normality of distribution. In the other, one tests for internal consistency, letting the observations supply their own estimate of error. The data for the external test are shown in Table 8. For each coefficient the t value is shown, calculated from the values of regression coefficients and their standard deviations which are set out in Table 9.

Table 8. *Significance of regression coefficients. Null hypothesis test*

	H		*U*		*P*		*F*		*L*	
	t	Sig.	t	Sig.	t	Sig.	t	Sig.	t	Sig.
1928	1·50	—	4·49	Yes	1·45	—	5·82	Yes	0·80	—
1929	3·71	Yes	3·64	Yes	1·02	—	4·85	Yes	2·86	Yes
1930	1·08	—	3·82	Yes	1·27	—	1·85	(+)	2·98	Yes
1931	3·45	Yes	4·12	Yes	0·87	—	1·63	(+)	1·76	(+)
1932	3·45	Yes	0·38	—	1·67	(+)	4·62	Yes	1·98	(+)
1933	2·59	Yes	2·72	Yes	2·04	Yes	3·60	Yes	0·24	—
1934	4·24	Yes	2·50	Yes	0·14	—	3·72	Yes	0·02*	—
1935	3·37	Yes	1·76	(+)	1·31	—	2·47	Yes	2·02	Yes
1936	2·13	Yes	0·77	—	2·09	Yes	2·52	Yes	2·58	Yes
1937	3·01	Yes	2·26	Yes	0·47	—	2·56	Yes	2·22	Yes
1938	1·46	—	2·96	Yes	1·02	—	2·98	Yes	0·45	—
Pool	8·69	Yes	5·21	Yes	3·96	Yes	8·58	Yes	4·95	Yes
Means	4·26	Yes	4·12	Yes	1·98	Yes	5·41	Yes	2·10	Yes

Table 9. *Significance of regression coefficients. Internal consistency*

	H		*U*		*P*		*F*		*L*	
	R.C.	S.D.	R.C.	S.D.	R.C.	S.D.	R.C.	S.D.	R.C.	S.D.
1928	0·2692	0·1801	0·9974	0·2220	0·3366	0·2315	0·6460	0·1110	0·9250	1·152
1929	0·7906	0·2129	0·6732	0·1851	0·3294	0·3229	0·6004	0·1237	4·296	1·500
1930	0·1732	0·1601	0·6546	0·1712	0·2586	0·2041	0·1602	0·0865	3·089	1·038
1931	0·6650	0·1930	0·6593	0·1600	0·2212	0·2537	0·1744	0·1070	2·290	1·298
1932	0·6152	0·1784	0·0530	0·1382	0·3856	0·2316	0·4406	0·0953	2·290	1·159
1933	0·5740	0·2218	0·4541	0·1669	0·5578	0·2729	0·4304	0·1197	0·3496	1·439
1934	0·7236	0·1709	0·3342	0·1338	0·0296	0·2176	0·3666	0·0984	−0·0216	1·137
1935	0·6252	0·1857	0·2568	0·1460	0·3106	0·2367	0·2620	0·1060	2·592	1·283
1936	0·3534	0·1658	0·1003	0·1304	0·4424	0·2118	0·2384	0·0947	2·970	1·153
1937	0·5440	0·1807	0·3419	0·1511	0·1092	0·2348	0·2694	0·1051	2·784	1·257
1938	0·2438	0·1666	0·4884	0·1648	0·2226	0·2176	0·2910	0·0975	0·5256	1·160
Average	0·5070		0·4557		0·2912		0·3527		2·008	
S.D.	0·2106		0·2764		0·1497		0·1632		1·373	
S.D. of Av.	0·0635		0·0833		0·0451		0·0492		0·414	
t	7·98		5·47		6·46		7·17		4·85	
Pool	0·5804	0·0667	0·2353	0·0452	0·3414	0·0863	0·3210	0·0374	2·212	0·447
Means	0·5068	0·1189	0·4440	0·1077	0·3000	0·1517	0·3548	0·0656	1·664	0·792

For 76 degrees of freedom the critical value of t may be taken as 2. In Table 8 the word 'Yes' is put against all coefficients reaching this level. The chances are less than 1 in 20 that coefficients of this magnitude, either positive or negative, could have arisen by chance. If one tests the possibility that a *positive* value as great as observed could have arisen by chance, the appropriate t figure is about 1·65. Those coefficients which pass this test are marked with the sign (+). It will be seen that on this basis all the F coefficients and most of the H and U figures are significant, even when taken in isolation. For the rest, it is enough to note that, except for L, all the coefficients are positive. If the universe correlation were zero, such a thing would happen by chance only once in 2^{11} trials, giving a probability of 1 in 2048. One coefficient of L is negative. One would obtain a result at least as favourable as this by chance only 12 times out of 2048 trials, giving a probability of 1 in 170·7. By means of a combined χ^2 test, taking account of the magnitudes of the deviations from zero, one could increase these probabilities many-fold. But this is not necessary. Taken as a whole, there is no doubt that the regression coefficients are overwhelmingly significant.

For the internal test, we computed the average of the coefficients for each social index over the 11 years, the S.D. of the distribution and of the average, and finally the t value for the average. For 10 degrees of freedom, the 1% point for t is 3·17. The results are shown in Table 9. The lowest value is 4·85 for the average of L. The average values of the coefficients are all significant beyond any doubt.

It will be noticed that the S.D. of each coefficient round its mean is closely similar to the individual S.D. for the separate years. It seemed worth while testing the hypothesis that the differences between years might have arisen by chance from an underlying constant value equal to the mean. For this purpose we calculated a set of t values for each coefficient, the deviation from the mean divided by the S.D. For instance, for H in 1928,

$$t = (0{\cdot}5070 - 0{\cdot}2692)/0{\cdot}1801 = 1{\cdot}32.$$

For each t, the area in both tails was calculated for 76 degrees of freedom, using Student's extended tables as given by Peters & Van Voorhis (1940). The natural logarithms of the eleven probabilities for each index were summed, and multiplied by -2, giving the χ^2 value for 22 degrees of freedom as described by Fisher (1938). The χ^2 figures and the corresponding probabilities are as follows:

	H	U	P	F	L
χ^2	27·6	44·7	10·9	38·2	24·2
P	*ca.* 0·20	$<0{\cdot}01$	*ca.* 0·98	$<0{\cdot}02$	*ca.* 0·30

If we regard each of the eleven regressions as an independent test of the constancy of the regression coefficients, then H and L are consistent with a uniform yearly effect, and P is even more close than would be expected by chance. The coefficients of U and F vary more than could be reasonably expected if their relation to infant mortality was really the same in different years.

We must now inquire how much of the variance can be accounted for. The relevant data are shown in Table 10. We will first consider the figures for the separate years. The 'Variance' column shows the total sum of squares of deviations of infant mortality about the yearly mean. R^2 is obtained from the multiple regression, and is equivalent to E_c, the proportion of variance accounted for as covariance with the five social indices. V_c is the product of the previous two columns. It is the covariance. V_s is the sampling variance, calculated for each year as described above, and E_s is the ratio of V_s to the total variance. By adding E_c and E_s, we obtain the total explanation, E. We would first direct attention to this last column. Between 1928 and 1932 approximately 80% of the variance is accounted for in each year, between 1933 and 1937 more than 75%, and only in 1938 does the explanation drop to 66%. These figures are obviously very high for a social investigation of this kind, and we will give reasons below for regarding the explanation as being as complete as could reasonably be expected. The fall in explanation in the later years may be partly due to the lessening relevance of the

Table 10. *Explanation of variance*

	Variance	$R^2 = E_c$	V_c	V_s	E_s	E
1928	20,465	0·6346	12,987	3,202	0·1565	0·7911
1929	40,166	0·7141	28,683	3,700	0·0921	0·8062
1930	13,817	0·5714	7,895	3,126	0·2262	0·7976
1931	26,456	0·6600	17,465	3,557	0·1344	0·7944
1932	18,804	0·6192	11,643	3,544	0·1885	0·8077
1933	26,366	0·5949	15,685	3,711	0·1407	0·7356
1934	15,858	0·5725	9,079	3,199	0·2017	0·7742
1935	20,393	0·6036	12,309	3,210	0·1574	0·7610
1936	14,502	0·5544	8,040	3,285	0·2265	0·7809
1937	17,872	0·5526	9,876	3,108	0·1739	0·7265
1938	11,334	0·3941	4,467	3,029	0·2672	0·6613
Totals	226,033	(0·6111)	138,129	36,671	(0·1622)	(0·7733)
Pool	262,320	0·4687	122,950	36,671	0·1398	0·6085
Means	161,553	0·7755	125,284	3,334	0·0206	0·7961

https://doi.org/10.1017/S0022172400035907 Published online by Cambridge University Press

1931 Census data used to calculate the social indices. But we feel that the discrepancies are less impressive than the regularities. The reliability of such consistent results must be regarded as very high.

The R^2 for the means equation is 0·7755, higher than in any individual year. But this does not correspond with increased covariance. Actually, the V_c of the means equation is appreciably less than the total of the separate years. We did not think it worth while making the laborious tabulations needed to calculate the sampling variances of each of the 82 means, because the sum cannot be very far off 36,671/11, which is 3334. Taking this value for V_s, we get 0·7961 for the total explanation in the means equation. This corresponds closely with 0·7733, the ratio of explained to total variance summed over the eleven years. The R^2 in the means equation exceeds the figure for any separate year because the smaller V_c forms a bigger fraction of a still smaller total variance. On the other hand, the low R^2 in the pool equation, though associated with a smaller covariance, greatly exaggerates the difference between this covariance and that obtained from the sum of years. These discrepancies indicate how misleading correlation coefficients can be unless they are considered in relation to the associated regressions. The variance in the pool can of course be analysed as follows:

	Sum of squares	D.F.	Mean square
Between years	36,287	10	3629
Between places	100,767	81	1244
Within years and places	125,266	810	155
	262,320	901	—

This clinches what must already be sufficiently obvious, that the differences between years and between places are statistically highly significant.

It is interesting to see which of the independent variables are affected by the lowered covariance in the pool and the means equations. The regression computations give the data required for estimating the contribution of each social variable to the total covariance. In the usual notation (Ezekiel, 1930) the explanation attributable to any independent variable is its β multiplied by its first-order correlation with the dependent variable. American writers call this the coefficient of separate determination, but in our terminology it would be the *separate explanation*. The separate explanation multiplied by the variance gives the separate covariance. The separate covariances for each social index were summed over the 11 years. We give the results in Table 11, which also shows the separate covariances in the pool and the means equations. It will be seen that all three methods agree reasonably well about the amount of covariance associated with H, P and F. For L, the means give about 25 % less covariance than the sum of years and the pool, which agree remarkably closely. The only serious divergence is for U, where the pool covariance is 52 % of that given by the sum of years. It is not difficult to see why the pool equation should underestimate the influence of unemployment on mortality. Unemployment is the only index used which varies year by year. The figures for different years are not strictly comparable, for reasons already explained. By pooling this series of figures, and measuring deviations from the common mean, one introduces additional error and randomization, which is apparently sufficient to reduce the separate covariance by 48 %. In all other respects the pool agrees remarkably closely with the sum of years, both in the partition of covariance and the regression equations. On comparing the sum of years with the means equation, one finds a substantial discrepancy in the separate covariance of L. We cannot account for this explicitly, but we give reasons later for believing that in relation to infant mortality L is a conglomerate index representing poverty factors rather than temperature. For the other indices, the two methods of computation give concordant results. The covariances in the means equation are consistently rather smaller than those in the sum of years. These differences seem to be strictly parallel to the reduction in the total variance induced by the process of averaging, for they do not appear in the regression equations. Indeed, as a study of Tables 7 and 9 will show, the averages over the eleven years of the regression coefficients for H, U, P and L are remarkably close to those in the means equation, only L being about 17 % lower in the means. The concordance is even better when due account is taken of the different numbers of deaths, in each year, as in the Balance Sheet of Infant Deaths given in the next section. In the pool the coefficients of H, P and L are higher than in the years average. But this is an example of the behaviour of regression coefficients discussed above. When one agency is inaccurately measured, the coefficients of other indices highly correlated with it tend to increase. As will be seen below, for purposes of calculating infant deaths the pool equation gives

Table 11. *Covariance associated with each social index*

	Sum of separate years		Pool	Means
H	46,453	33·7 %	49,325	42,941
U	34,001	24·6 %	17,609	32,553
P	16,900	12·2 %	18,863	16,478
F	19,219	13·9 %	15,644	17,415
L	21,556	15·6 %	21,509	15,897
Total	138,129	100 %	122,950	125,284

results much closer than the discordance in regression coefficients might suggest to the figures obtained from the yearly and the means equations. Table 11 shows that H is the most important contributor to the covariance, accounting for about one-third. U is responsible for about one-quarter, and the other three indices together are associated with the remaining 42 %.

7. THE BALANCE SHEET OF INFANT DEATHS

So far we have been discoursing in terms of correlations, sums of squares, and other abstractions of the statistician. Before our results can afford any guide to practice, we must translate our variances measured in 'square deaths' into actual numbers of families who were poor or overcrowded and of babies who died. To do this, we must assume that the regression equations in Table 7 give a reliable basis for calculating the infant death-rate at any level of social environment that can be specified by means of our five indices. We must first discuss how far and with what reservations this assumption is justifiable. If the regressions plus sampling error had covered all the variance, there would be no problem. We have to show that our regression equations are unlikely to be invalidated by the unexplained variance.

Our explanation is of course remarkably high. In the means equation the non-explanation is $1-0{\cdot}7961$, which is $0{\cdot}2039$. If the whole of this were due to covariation of infant mortality with some independent variable uncorrelated with our five social indices, its maximum first-order correlation coefficient with I.M. would be $\sqrt{0{\cdot}2039} = 0{\cdot}4516$, lower than for any of our five indices except F. But we shall show that the residual variance can be plausibly accounted for on other grounds. It can in fact be apportioned under four headings.

(a) *Errors in measurement of infant mortality*

We have already discounted the sampling variance of the mortality. But this is calculated on the assumption that births are accurately registered, and that no mistakes are made in the recognition or attribution of deaths. Nor have we taken account of the possible errors arising from the yearly fluctuations in the numbers of births. Many babies who die in one calendar year are born in the previous year. By relating their deaths to births in the current year one may introduce an error, which is discussed by Kuczynski (1935). A rough calculation indicates that the variance from this source is unlikely to exceed 10 % of the sampling variance. But even this would be an appreciable fraction of the non-explanation. Then there are actual errors in both the births and the deaths. It is hard for the midwife or doctor to draw with accuracy the line between stillbirths and deaths in the first few minutes of life. Babies may migrate with their parents out of the town of their birth, and die in a new place of residence, causing error in the vital statistics of both places. Births or deaths may be concealed, or dates and ages may be wrongly reported. All such inaccuracies would add to the variance of the mortality figures. One can only guess at the magnitude of this error variance. But if one makes the plausible assumption that it is more or less randomly distributed among places, it will not affect the validity of the regression equations.

(b) *Errors in the independent variables*

Loss of explanation may result from two kinds of flaws in the indices expressing the independent variables. First there is imperfect relevance. Our E_c values would have been smaller if instead of our chosen housing index we had used the percentage of families living more than 2 per room. None of our indices is the most relevant possible, but only at best the most relevant of the measures available in a fragmentary body of social statistics. Then there are actual inaccuracies in the measurements on which the index is based. We have taken great pains to avoid these sources of error. But we cannot claim to have eliminated them. On the contrary, we have given reasons for believing that at least three of our indices are very faulty. All such imperfections will increase the residual variance at the expense of the regression. The loss of explanation is to a large extent minimized by the use of several intercorrelated indices. Deficiencies in one are counterbalanced by the others. Nevertheless, we believe that flaws in the indices are important sources of error. To the extent that they occur, our computation of the number of deaths attributable to the social variables will be an understatement of the true situation.

(c) *Omission of relevant social agencies*

We have already described some indices whose inclusion would probably have increased the total explanation. There would have been an increase in E_c if we had included index W as well as P to give a multiple measure of social class, or housing index A of Table 3 to give extra weight to gross overcrowding. The same considerations apply to the addition of unemployment figures for the current year as well as for the year previous to the mortality. It is easy to specify other indices whose relevance may reasonably be anticipated, for instance unemployment among women. There are a number of social agencies whose relevance to infant mortality is scarcely open to doubt, but which had to be excluded because

https://doi.org/10.1017/S0022172400035907 Published online by Cambridge University Press

they could not be measured without a prohibitive amount of inaccuracy. These include:

Size of built-up area. The consistent difference between the unweighted and the weighted means which is shown in Fig. 4 is sufficient indication that large towns tend to have bigger mortalities than smaller places. This is borne out by the higher level of mortality in county boroughs as compared with the less populous municipal boroughs and urban districts, and the still smaller mortality in rural areas. But when it comes to expressing this effect in figures, almost insuperable difficulties arise. An index based on size of population does not meet the case. Birmingham, for example, would have a high index and Smethwick a low one, although both are parts of the same built-up area and any effect of high population such as greater prevalence of infection would presumably affect babies equally on both sides of the borough boundaries. The same considerations apply to such places as Manchester and its smaller neighbour Salford, to the Tyneside towns, and especially to the three county boroughs in the Greater London area. The only valid measure of the effect would be the population of the total conurbation. This would be difficult to determine without gross ambiguity. We did try out an index proportional to the population of each county borough. As we feared, the correlations with infant mortality were trivial.

Persons per acre. Successive reports from the Registrar-General's office have been pointing to the positive relation between mortality and density of population ever since the time of William Farr. But the construction of an accurate index is a formidable task. The usual criterion, persons per acre, cannot be applied to big areas like county boroughs. Some cities, Leeds and Birmingham for example, have extended their boundaries to include undeveloped land to be used for future housing schemes. Their figures for persons per acre would be fictitiously low. In other cases the value of the index will depend on exactly where the borough boundary runs in relation to open spaces or factory areas on the outskirts. We tried out a persons-per-acre index. It gave an obviously distorted picture of comparative densities of population, and its correlation with infant mortality was negligible.

Infant welfare services. No practical worker in preventive medicine has the slightest doubt that the progressive decline in infant mortality during this century is to a large extent due to better public care for infant life, coupled with improved mothercraft. It seems likely that towns with good infant welfare activities will save more infant lives than those with indifferent services. But to prove or disprove this conjecture from the available figures is quite impossible. The Local Government Financial Statistics of the Ministry of Health give annually for each county borough the amount spent on Maternal and Child Welfare. But these figures depend a great deal on the book-keeping system of the local authority. Some councils debit the general Public Health Fund with the cost of activities that others would put to the account of the Maternal and Child Welfare Committee. The Ministry's Returns do not specify how much is spent on infants, as distinct from toddlers and mothers. Nor do they take account of voluntary or charitable effort, of whether the authority has to rent premises or possesses its own, or of many other facts relevant to the assessment of the adequacy of a service for infant welfare. We tried correlating infant mortality with two indices—the amount spent per head of population, and per birth. The second index gave an appreciable negative regression on infant mortality in London boroughs, where the Public Health organization and the methods of accounting are presumably more comparable. But in county boroughs both indices gave regression coefficients that did not differ sensibly from zero.

Size of family. As a rough index of the relative number of children per family we used G, the gross reproduction rate for county boroughs in 1931, as calculated by Charles (1936). Unlike the indices just considered, this social variable in conjunction with H, U, P, F and L did give appreciable extra covariance, ranging in different years from practically zero to one-fifth the residual variance. Nevertheless, we decided to omit this index, for three reasons. First, we have evidence that the relation between G and I.M. is not simple and straightforward. With some causes of death, G gives positive regression coefficients, and negative ones with others. We hope to deal with the G effect in a subsequent communication. Second, the inclusion of G would raise the extraneous issue of how much infant mortality would change if families were larger or smaller. We want to confine our attention for the present to the effect of social conditions on those infants actually born, rather than on those who would be born if there were a change in reproduction rates. Third, the inclusion of G as a sixth independent variable does not markedly alter the regression coefficients of the other five indices. Its omission therefore does not affect the validity of the conclusions to be drawn from the regressions.

It is easy to think of other social agencies which may be relevant to infant mortality. In so far as they are highly correlated with the social variables which appear in the regression, the omission of these extra indices is to a large extent compensated. For instance, density of population in persons per acre must be closely correlated with overcrowding as measured by persons per room, so our H index is a measure of both effects. Nevertheless, the inclusion of more variables would undoubtedly greatly increase the explanation. If we put into one equation all those indices we tried out, we believe the non-explanation would be decreased by at least a quarter, and quite probably by more than half. The omission of relevant variables has the same effect on the regression as imperfections in the indices actually used. Potential covariance is lost, and the effect of social conditions on the number of deaths is underestimated.

We think it will be conceded that omitted variables and errors of measurement are probably ample to account for the small amount of unexplained variance, and that their effect on the regression would be to cause it to understate the number of deaths attributable to social conditions. But there is a fourth possible source of loss of explanation which, if it existed, would be more serious. This is:

(*d*) *Departure from linearity of regression*

As we propose to use the regression equations for calculating the number of deaths to be expected at various levels of the social indices, which in a sense involves extrapolating, any marked departure from linearity would bias our estimates. There are no really satisfactory mathematical tests for curvature in multiple regressions. But we can apply a graphical test which we think is reasonably decisive. If an equation of the first degree does not correctly represent the regression, then the residual deviations will not lie evenly about the line. There might be an excess of positive deviations at high and low values of the dependent variable, and an excess of negative deviations at medium values; or vice versa. In such a case, it would be inaccurate to use the regression formula except in the middle range, and impermissible to extrapolate beyond the limits of recorded values. But if the positive and negative deviations are evenly balanced all along the line, one may legitimately use it for reasonable extrapolation if there are other grounds for supposing that the relation between the variables is linear. In our balance sheet below, we stress only the total deaths over the 11 years. The totals are substantially the same whether derived from the sum of years or the means equations. To establish linearity, it is therefore sufficient to show that deviations from the means equation lie evenly at all values of $\bar{M}$. We therefore calculated for each of the 82 county boroughs the theoretical mortality assuming the truth of the regression equation, and the deviation of the observed $\bar{M}$ from this value. The results are shown in Fig. 7. The horizontal scale is calculated mortality, and the vertical is deviations. It will be seen that there is no tendency to curvature; and this is borne out by the usual tests involving the computation of deviations of means of arrays. The figure also shows that the points cluster more closely about the line at the lower values of $\bar{M}$. This agrees with the data displayed in Figs. 5 and 6. Mortality in the lower ranges is less variable than among the more poverty-stricken towns which are more at the mercy of changes in adverse conditions. Furthermore, any additional social variables, such as density per acre, have their main effect on places with high mortality. The loss of explanation due to the omission of these factors tends to be concentrated along the higher part of the line, increasing the error of estimate in that region as compared with the lower ranges. It follows that the regression is even more reliable for low mortalities than the amount of non-explanation might lead one to believe.

Before we proceed to our balance sheet of infant deaths, there is one last matter that must be dealt with. We have included latitude as a social variable because its correlation with mortality is so frequently stressed by the Registrar-General. It is true that latitude gives appreciable separate covariance with infant mortality. But we do not accept the conclusion, unhesitatingly drawn by the Registrar-General, that deaths attributable to latitude are caused by climatic conditions, with the inference that nothing can be done about them. In England and Wales there are many things beside the climate that get more severe as one travels north. Take housing, for example. West Ham in the south is rather more unfortunate than West Hartlepool in the north in percentage of families living more than 1 person per room—44·5 % as opposed to 40·8 %. But nobody who knows the facts will agree that housing conditions in West Ham are slightly worse than in West Hartlepool. The index can only measure rooms as equal units. It cannot take account of the smaller rooms, the more defective sanitation, the back-to-back houses and all the other features which tend to make an overcrowded dwelling in the north more lethal to babies than one in the south with a similar ratio of persons to rooms. To express housing conditions fully, one should supplement the overcrowding index in the regression by an index of housing quality. In default

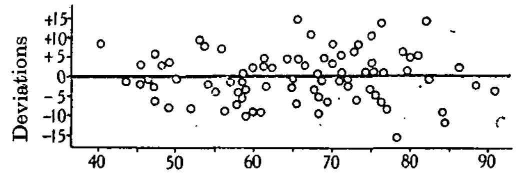

Fig. 7. Deviations of mean infant mortality in county boroughs from theoretical value calculated from regression equation.

of anything better, latitude would be a good approximation to such an index. In the same way, our poverty index, P, does not truly express relative income levels. It is well known that wage rates get lower as one goes north, and this was brought prominently to our notice when we were collecting data in our attempt to improve the P index. If one wants to measure the effect of low standard of living, P plus L would probably give a truer picture than P alone. Then there are the social agencies such as persons per acre, which we could not include because no reasonable measure was available. In most cases, latitude could be used as an imperfect substitute. In our trial computations, we were impressed by the fact that every time a fresh relevant variable was introduced into a regression, a substantial part of the covariance previously associated with latitude was transferred to the new index. We therefore regard our index L as being primarily an omnibus measure of miscellaneous omitted social agencies. We do not, of course, rule out a direct climatic effect. The differences between years shown in Fig. 4 form a *prima facie* case for suspecting that weather changes may be important. We hope to deal with this point in a future communication. Our evidence so far inclines us to the belief that the effect, if any, of adverse weather conditions, does not occur unless there is concomitant overcrowding and low standard of living. If this is the case, even the climatic aspect of the correlation of mortality with latitude would be a poverty index.

Fortunately, it is possible to test whether the latitude effect exists in the absence of poverty. In his Decennial Supplement for 1931, the Registrar-General gives figures for infant mortality by social class of father in each of his twelve regions. Taking each region in turn, we started at the south with the aid of Phillip's Administrative Atlas, and added the populations of all the local government areas until half the total population of the region was reached. The latitude at which this happened was taken as the latitude of the region. There was no ambiguity about this figure to within a quarter of a degree. The data for the regions arranged in ascending order of latitude are shown in the top part of Table 12. It is easy to see that for Class I mortality is much the same at all latitudes, while for Classes III, IV and V it increases as one goes from south to north. To clinch the matter, we calculated for each Class the regression of latitude on infant mortality. The equations are shown at the foot of Table 12, and are plotted in Fig. 8. We do not, of course, claim much precision for the regression coefficients. Nor does the slope represent the effect of latitude alone, but rather of latitude and all the social agencies, such as overcrowding, low wages and unemployment, which are positively correlated with it. But we regard the five equations, taken together, as entirely convincing. If a baby is as well looked after as those in Class I, it does not matter whether it lives at Land's End or Berwick-on-Tweed. Once again we have the phenomenon brought out in Figs. 5 and 6. The infants of the well-to-do are insulated against the risks of babyhood. The very poor are fully exposed to them.

We now come to our final accounting. The rele-

Table 12. *Infant mortality by latitude and social class: regions* 1930–2

		Infant mortality in social classes				
Region	Latitude	I	II	III	IV	V
South-west	$50\frac{3}{4}°$	35·0	43·7	45·4	54·4	58·6
South-east	$51\frac{1}{2}°$	33·1	38·4	41·6	46·8	54·0
Greater London	$51\frac{1}{2}°$	29·6	39·8	49·7	62·3	71·2
Wales I	$51\frac{3}{4}°$	43·1	52·2	70·0	73·0	77·4
Midland I	$52\frac{1}{2}°$	32·7	44·5	58·9	63·8	77·7
East	$52\frac{3}{4}°$	29·9	41·7	46·7	55·9	61·5
Wales II	53°	38·6	54·5	61·7	69·6	70·3
Midland II	53°	35·6	45·0	61·2	64·5	76·0
North IV	$53\frac{1}{2}°$	31·1	51·2	66·8	78·9	93·3
North III	$53\frac{3}{4}°$	34·4	46·9	67·8	74·7	84·6
North II	$54\frac{1}{4}°$	39·2	52·6	61·7	73·1	82·3
North I	55°	37·8	50·3	71·9	86·8	100·6

Regression equations

Social class	Regression equation
I	$M = 34{\cdot}9 + 0{\cdot}07L$
II	$M = 41{\cdot}2 + 2{\cdot}4\ L$
III	$M = 45{\cdot}6 + 5{\cdot}7\ L$
IV	$M = 50{\cdot}9 + 7{\cdot}1\ L$
V	$M = 55{\cdot}6 + 8{\cdot}8\ L$

Where L = degrees of latitude north of $50\frac{1}{2}°$.

https://doi.org/10.1017/S0022172400035907 Published online by Cambridge University Press

vant data are shown in Tables 13 and 14. From the regression equations one can calculate the theoretical infant mortality for any desired values of the five social indices. In the column of Table 13 headed 'All indices at best observed' we have done this for a hypothetical borough with a value for each index at or near the lowest actually observed. U and F are taken as zero, H as 15 %, P as 20 % and L as 50° 30′. The yearly values are not quite concordant, but when used to calculate an average, they give the figure 36·5 with a S.D. of only ± 0·8. The means equation gives exactly the same figure, 36·5. The pool equation, with its over-emphasis on P and H, gives the somewhat higher figure of 38·1. We do not give S.D. for these figures. There is no satisfactory procedure for estimating the precision of the K term of a regression equation. Such methods as are available assume that everything is normally distributed, and that the deviations are smallest in the middle range, whereas our regressions are most reliable for low values of M. It will be seen from Figs. 5 and 6 that 36·5 is a reasonable estimate of the observed I.M. in the most fortunate group of boroughs. We may say with considerable confidence that if social conditions in all boroughs could have been levelled up to the best existing, the mean I.M. over our 11-year period would have been about 36·5 instead of 70·3, a reduction of very nearly 50 %.

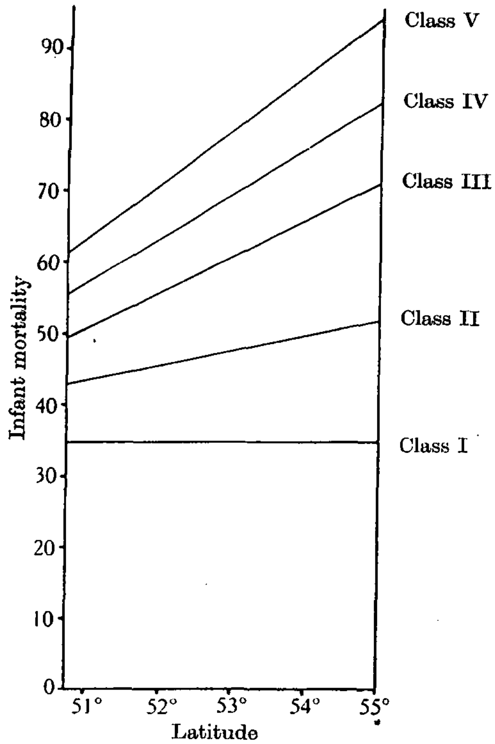

Fig. 8. Infant mortality in regions of England and Wales, 1930–2. Regression lines showing variation of infant mortality with latitude in each social class.

But the 'best observed' still includes a considerable proportion of families poor and overcrowded. Is there any way of estimating the number of deaths to be expected if our five indicators of poverty all pointed to zero? This can obviously be done by calculating M from the regression equations when H, U, P and F are each taken at 0, and L at 50° 30′, just north of the southernmost county borough. In a sense this is an extrapolation beyond the range of the data. But from another standpoint this is not the case.

For instance, the I.M. of Birkenhead in 1933 was 100; in the contiguous county borough of Wallasey it was 54. This does not mean that the mortality risk of every baby in Birkenhead was 100. Wallasey contained 22·3 % of men in Classes I and II. The mortality of their babies must have been below the borough's average of 54, just as the mortality of

Table 13. *Calculated infant mortality at various levels of the social indices*

Year	All indices at best observed	All indices at best possible	Overcrowded poor	Unemployed overcrowded poor
1928	35·0	24·2	87·1	186·9
1929	35·9	17·4	139·7	207·1
1930	38·8	31·1	81·7	147·1
1931	38·2	23·8	117·9	183·8
1932	41·6	24·7	130·3	135·6
1933	33·0	13·3	127·3	172·7
1934	36·9	25·5	100·7	134·2
1935	31·9	16·3	116·1	141·7
1936	38·1	24·0	110·7	120·7
1937	35·2	24·8	96·8	131·0
1938	37·3	29·1	77·0	125·9
Average	36·5 ± 0·8	23·1 ± 1·6	107·8 ± 6·3	153·3 ± 8·8
Pool	38·1	22·6	120·1	143·6
Means	36·5	22·9	107·6	152·0

https://doi.org/10.1017/S0022172400035907 Published online by Cambridge University Press

Wallasey's 24·9 % of Classes IV and V must have been well above 54. But Birkenhead also has its Classes I and II population—not as many as Wallasey, but the substantial figure of 12·0 %. The mortality of the babies of these men could not have been widely different from those of similar people across the borough boundary, below 54. Therefore, to reach Birkenhead's average I.M. of 100, there must have been a number of babies experiencing a mortality risk of well over 100—probably over 150. If one had all the data, one could isolate from both Birkenhead and Wallasey social groups with identical average I.M. risks. There might be a group with an I.M. of 30. These would be more numerous in Wallasey, but not entirely absent from Birkenhead. At the other end of the scale would come a group at 150, or possibly higher, relatively numerous in Birkenhead, with its 29·7 % in Group V and its 32·7 % of overcrowding, but not completely unrepresented among the 16·9 % in Group V and the 17·6 % overcrowded in relatively prosperous Wallasey. The place with the lowest mortality experience, Bath, had 22 % overcrowding, a P index of 26 % and an average unemployment of nearly 13 %. In the boroughs with highest mortality, close on half the population was neither overcrowded, unemployed nor poor. The contrasts between the ends of the social scale must be many times bigger than that between the best and the worst county boroughs.

But our regression equations perform precisely this function of dividing the population into strata with different characteristic I.M. rates. To take a simple numerical example with only one independent variable, suppose we had the regression equation

$$M = 30 + 0{\cdot}7H.$$

Provided the regression was linear, this would tell us that we could calculate M for any place by dividing the population into two portions—those living less than 1 per room, with a mean I.M. of 30, and those living more than 1 per room, with a mean I.M. of $30 + 70 = 100$. If the explanation were substantially complete, and the range of H values large, we could say with considerable confidence that the I.M. of the uncrowded, averaged over all places, was *actually* 30, and the I.M. of the crowded, taking one place with another, was *actually* 100. All our social indices, with the exception of L, are in the form of the percentage of the population possessing a given attribute, such as being unemployed, or being in Group IV or V. Our regressions cover a wide range of values of the indices, the deviations lie evenly plus and minus all along the range, and substantially all the variance is plausibly accounted for. We can therefore use them with considerable confidence to isolate certain broad strata of the population common to all county boroughs. To test the method, we calculated the regression of the average I.M. of the 3 years 1930–2 on the two independent variables P and W. The equation was

$$\bar{M} = 56{\cdot}0 + 0{\cdot}79P - 1{\cdot}14W.$$

By equating P (Classes IV and V) and W (Classes I and II) to zero we can 'recover' the I.M. for Class III. The value obtained is 56·0. By working out separate equations for each of the 3 years, and averaging, we obtained 54·1. The Registrar-General's figure for the I.M. of Class III during 1930–2 is 57·6.

Applying this method to our main equations, we calculated the mortality in the following strata of the population:

(*a*) 'Best Possible'. H, U, P, and F all equated to zero, L to $50\frac{1}{2}°$.

(*b*) 'Overcrowded Poor'. H and P at 100 %, U and F at zero, L at the mean latitude of 52·9°.

(*c*) 'Unemployed Overcrowded Poor'. As in (*b*) except that U also is equated to 100 %.

The values obtained are shown in Table 13. For the most fortunate stratum of the population, the eleven yearly equations give an average I.M. of $23{\cdot}1 \pm 1{\cdot}6$, as compared with the pool figure of 22·6 and the means estimate, 22·9. The agreement is extraordinarily close. We regard this as a well-established estimate. If anything, it is rather high, because the errors in the five indices and the omission of relevant variables tend to make the regressions display less than the true covariance. We can therefore state this important conclusion. *In the section of the population of county boroughs not subject to our adverse social agencies, infant mortality averaged* 23·1. *Where any* I.M. *figure was significantly above* 23·1, *the excess deaths could have been prevented by the abolition of overcrowding, unemployment, low-paid occupation, and industrial work of women.* These excess deaths in county boroughs amounted to about two-thirds of the total. For the 'overcrowded poor' and the 'unemployed overcrowded poor' the yearly values are much less concordant. They reflect the two-year cycle whose main brunt fell on the poorer strata of the population. Once again the means agrees remarkably closely with the average from the eleven yearly equations. The pool equation, with its inefficient measure of unemployment, gives higher values for the poor and lower for the unemployed poor. Over the whole period the mean I.M. for the stratum of the population living in overcrowded houses on the wages of a semi-skilled or unskilled man may be taken as about 108. The I.M. for families subject to continuous unemployment, and living at the level of state benefit or local relief, may be taken as over 150. These figures were subject to quite violent 2-yearly fluctuations. The long-term trend of I.M. among the unemployed was downward. In the four years 1928–31, the average was 181·2, and in 1935–8 it was 129·8, a decrease of over 50. For the employed poor, on the other hand, the trend was practically horizontal, the average being 106·6 in 1928–31 and 100·1 in 1935–8. Some of the improvement in the I.M. in unemployed families may possibly be attributed to increases in scales of benefit and relaxation of means tests. It is worth noting that amid the general improvement this section of the population still had an I.M. at the high nineteenth-century level.

The standard of living of the unemployed poor was approximately at the level of the Poverty Line (Jones, 1934) which was devised by Bowley (1915, 1925) and extensively used in social surveys as a criterion of poverty. When these surveys reported that only 5–15 % of the population of various towns were in poverty, they were using a 150 I.M. standard. The 'overcrowded poor' would be living at approximately the Rowntree (1937) 'Human Needs' standard, which is much the same as the 'Subsistence Standard' put forward by Sir William Beveridge (1942) on the advice of Bowley and others. On the basis of past experience, families living continuously on the Beveridge scale may expect an I.M. of 100 or more.

In all the above calculations, we have equated *F* to zero. Table 7 shows the regression coefficient of *F* to be about 0·35. When *F* is 100, it increases *M* by 35. In other words, the mortality risk of babies whose mothers are at work is increased by 35 per 1000. Actually, this is a minimum figure, because our *F* represents the percentage of *all* women at work. In a town where say 50 % of the total women work at industry, this will include well over 50 % of the single women, and much less than 50 % of the married women. The true extra mortality risk must therefore be much above 35—quite possibly 60 or 70. We hope to investigate this more closely in a future communication. It is not a valid deduction that all manual work by married women, whatever the conditions, is necessarily harmful to their infants. The regressions can only tell us what happens under existing circumstances as to hours, factory conditions, provision for time off for pregnancy, and so on. We intend to go into the question as thoroughly as the data allow. It is clear that the excess risk has been decreasing rapidly during this century, presumably keeping step with shorter hours, better factory conditions and maternity benefit under National Health Insurance.

In Table 14 our equations are translated into actual infant deaths. Take, for instance, the year 1928. From the regression equation, the number of births, and the mean value of each independent variable, we can calculate the theoretical number of deaths. This is found to be a little less than the observed number, because we are using unweighted means of the variables. The deficiency is the difference between the broken and the full line at the top of Fig. 4, and is relatively small. We could calculate means weighted by births for all five variables for 11 years, but the heavy computation involved would not be justified, since the balance sheet cannot claim to be anything more than roughly approximate. We therefore multiplied the deaths in each category by the factor $\frac{\text{Weighted mean I.M.}}{\text{Unweighted mean I.M}}$. The effect of this is to exaggerate slightly the unexplained deaths, and underestimate the number explained by the regression. We are therefore slightly understating the case. The calculations were made by taking the infant deaths actually recorded in the

Table 14. *Balance sheet of infant deaths*

Year	Deaths	Unexplained	*H*	*U*	*P*	*F*	*L*	Total explained
1928	16,994	6,076	1,855	3,039	3,014	2,437	573	10,918
1929	19,185	4,200	5,235	2,258	2,835	2,177	2,480	14,985
1930	15,266	7,426	1,138	2,147	2,209	576	1,770	7,840
1931	16,635	5,548	4,266	3,083	1,844	612	1,282	11,087
1932	15,575	5,580	3,821	322	3,113	1,498	1,241	9,995
1933	14,680	2,813	3,349	2,735	4,229	1,375	179	11,867
1934	13,369	5,626	4,393	1,909	233	1,219	−11	7,743
1935	13,325	3,588	3,792	1,257	2,448	870	1,370	9,737
1936	13,501	5,197	2,106	451	3,426	778	1,543	8,304
1937	13,483	5,550	3,341	1,324	871	906	1,491	7,933
1938	12,401	6,442	1,481	1,473	1,758	968	279	5,959
Totals	164,414	58,046	34,777	19,998	25,980	13,416	12,197	106,368
	100·0 %	35·3 %	21·1 %	12·2 %	15·8 %	8·2 %	7·4 %	64·7 %
			32·7 %	18·8 %	24·4 %	12·6 %	11·5 %	100·0 %
Pool	164,414	56,747	40,063	11,539	30,621	12,130	13,314	107,667
	100·0 %	34·5 %	24·4 %	7·0 %	18·6 %	7·4 %	8·1 %	65·5 %
			37·2 %	10·7 %	28·4 %	11·3 %	12·4 %	100·0 %
Means	164,414	57,375	34,932	21,848	26,871	13,389	9,999	107,039
	100·0 %	34·9 %	21·3 %	13·3 %	16·3 %	8·1 %	6·1 %	65·1 %
			32·6 %	20·4 %	25·1 %	12·5 %	9·4 %	100·0 %

82 county boroughs in each year, as shown in the table, and dividing them up on the assumption that they belonged to a place with the unweighted mean values for each of the six variables. The pool and means equations were, of course, based on total deaths over the 11 years.

The column headed 'Unexplained' shows the number of deaths that would theoretically have occurred, if H, U, P and F were all equated to zero, and L to 50° 30′, corresponding with the 'Best possible' column of Table 13.

Once again the sum of years and the means calculations agree very closely, while the pool underestimates deaths attributable to U and partially compensates by overestimating the influence of H, P and L. Of total deaths, about a third are not associated with our social variables, 21 % are attributable to overcrowding, 16 % to low-paid occupation, 12 % to unemployment, 8 % to industrial work by women, and 7 % to 'latitude', which can be interpreted as poverty agencies not otherwise specified. Of preventable deaths, one-third are associated with overcrowding, one-quarter with low wages, one-fifth with unemployment, and one-eighth with industrial work by women. It is, of course, somewhat artificial to isolate the social agencies in this way. It is in practice impossible to abate overcrowding alone or unemployment alone without altering the incidence of other symptoms of poverty. This is shown, for example, by M'Gonigle & Kirby (1936). It is more realistic to group the figures as follows:

Unexplained	58,000	
Poverty syndrome:		
Overcrowding	35,000	
Low wages	26,000	
'Latitude'	12,000	
Total		73,000
Unemployment	20,000	
Poverty syndrome and unemployment		93,000
Industrial work by women	13,000	
Total	164,000	

If instead of taking our origin at 'Best possible' we take it at 'Best observed', deaths attributable to overcrowding would be 16,000 instead of 35,000, and those ascribed to low wages would be 12,000 instead of 26,000. It is also instructive to give the data in terms of infant mortality rates:

	Best possible	Best observed
'Unexplained'	23·1	36·4
H	13·8	6·3
U	8·0	8·0
P	10·3	4·5
F	5·3	5·3
L	4·9	4·9
Total explained	42·3	29·0
Total	65·4	65·4

This indicates that if all county boroughs were levelled up to the best existing, I.M. would fall by about 29. If all poverty agencies were eliminated, the rate would be decreased by over 42. If unemployment were reduced to zero, one would expect a reduction of about 8, with a further fall if there were an increase in the earnings of the lowest paid workers. The decline in I.M. since the outbreak of war is of this order of magnitude and may fairly be attributed to the virtual disappearance of unemployment. But even a rate of 50 is far above what could be achieved by the elimination of poverty altogether.

Our calculations relate only to county boroughs, containing about one-third of the population of England and Wales. Since the infant mortality in other urban and in rural areas is consistently lower, it is fair to take 23·1 as an attainable figure for the country as a whole. We can then make the following calculation:

England and Wales, 1928–38

			I.M.
Live births	6,812,668		
Infant deaths	421,001	100 %	61·8
Calculated deaths at I.M. = 23·1	157,373	37·4 %	23·1
Preventable deaths	263,628	62·6 %	38·7

Social agencies over the 11 years can be held to account for over a quarter of a million infant deaths, comprising over 60 % of the total. If they were prevented, there would be a fall of over 38 in the rate of infant mortality.

In conclusion, we would like to refer again to the theme of our introductory section. We have shown that infant deaths reflect adverse social conditions with remarkable fidelity. If genetic differences are important in determining the relative level of infant mortality, we can only marvel at the precision with which genetic inferiority is punished by unemployment and overcrowding, and still more at the rapid mass mutations that debase the genetic constitution of a city immediately its staple industry experiences a sudden upsurge of unemployment.

8. SUMMARY

1. The Introduction (pp. 67–73) describes the course of infant mortality in England and Wales over the past century, and critically reviews arguments advanced to prove that variations in the infant mortality rate (I.M.) are caused by genetic differences with respect to viability.

2. The I.M. in county boroughs shows a 2-yearly cycle of variability, affecting places with high mortalities.

3. We have devised and tested various social indices and have selected five which gave the highest

https://doi.org/10.1017/S0022172400035907 Published online by Cambridge University Press

joint covariance with infant mortality in county boroughs during the 11 years 1928–38. These indices are:

H, percentage of families living more than 1 person per room.

U, percentage of men unemployed.

P, percentage of occupied males in the Registrar-General's Social Classes IV and V.

F, percentage of women employed on manufacturing processes.

L, latitude.

4. We have computed multiple regression equations involving I.M. and the five indices for each of the 11 years, and two summarizing equations. The regressions plus sampling variance account for about 80% of the total variance in infant mortality. The regression is linear.

5. Latitude does not affect infant mortality in Class I. For this and other reasons we regard the latitude effect as expressing miscellaneous poverty indices omitted from our equations.

6. The regression equations enable us to divide the population into various strata with characteristic average infant mortality rates. These include:

'Better off' (all poverty indices = 0)	I.M. = 23·1
Overcrowded poor	I.M. = 108
Unemployed overcrowded poor	I.M. = 153

Babies whose mothers work in industry suffer an additional mortality risk of at least 35 per 1000, and possibly more. The figure 23·1 is the I.M. rate that would prevail if our five poverty symptoms could be eliminated.

7. In county boroughs two-thirds of infant deaths would be avoided by the abolition of conditions defined by our indices. Of the preventable deaths, one-third are associated with overcrowding, one-quarter with low-paid occupations, one-fifth with unemployment, and one-eighth with industrial employment of women. In England and Wales, over 250,000 deaths in 11 years, about 63% of the total, can be attributed to adverse social conditions.

We have to thank the Rockefeller Foundation for a personal grant to one of us (J. W.) out of an allocation for research work in Prof. Lancelot Hogben's Department, and the Halley Stewart Trust for a grant for mechanical computing equipment. Our thanks are also due to the various members of the Zoology Department in the University of Birmingham who assisted in computing at various times, and especially to Prof. Hogben for his unfailing interest, advice and support.

REFERENCES

BEVERIDGE (1942). *Report on Social Services*. London.

BOWLEY (1915). *Livelihood and Poverty. London.*

BOWLEY (1925). *Has Poverty Diminished?* London.

BROWNLEE (1917). *J. Roy. Statist. Soc.* **80**, 222.

CHARLES (1936). Article in *Political Arithmetic*, edited Hogben (1936).

ELDERTON & PEARSON (1915). *Biometrika*, **10**, 488.

EZEKIEL (1930). *Methods of Correlation Analysis*. New York.

FISHER (1930). *Genetical Theory of Natural Selection*. London.

FISHER (1938). *Statistical Methods for Research Workers*, 7th ed. Edinburgh.

FISHER (1942). *Ann. Eugen., Lond.*, **10**.

JONES (1934). *Social Survey of Merseyside*, **1**. London.

KUCZYNSKI (1935). *Measurement of Population Growth*. London.

LOVE (1936). *Statistical Methods Applied to Agriculture*. Shanghai.

M'GONIGLE & KIRBY (1936). *Poverty and Public Health*. London.

MINISTRY OF HEALTH (1936). *Report of Overcrowding Survey, England and Wales.*

MINISTRY OF HEALTH (annually). *Local Government Financial Statistics.*

MINISTRY OF LABOUR (1927 to 1939). *Local Unemployment Index.*

NEWSHOLME (1910, 1913, 1914). *Special Reports on Infant and Child Mortality to Local Government Board.*

ORR (1935). *Food, Health and Income*. London.

PEARSON (1912). *Proc. Roy. Soc.* B, **85**, 469.

PEARSON & ELDERTON (1922). *Biometrika*, **14**, 281.

PETERS & VAN VOORHIS (1940). *Statistical Procedures and their Mathematical Basis*. New York.

REGISTRAR-GENERAL. All references give date to which report refers, not date of publication. Abbreviations: A.R. Annual Review; D.S. Decennial Supplement, Part 2 (Occupational Mortality); Text, Text Volume of Annual Statistical Review.

ROWNTREE (1937). *Human Needs of Labour*. London.

SHEWHART (1931). *Economic Control of Quality of Manufactured Product*. New York.

SHEWHART (1939). *Statistical Method from the Viewpoint of Quality Control*. Washington.

SIMON (1941). *Engineers' Manual of Statistical Methods*. New York.

SIMPSON & ROE (1939). *Quantitative Zoology*. New York.

SNOW (1911). *Intensity of Natural Selection in Man*. London.

STEVENSON (1921). *J. Roy. Statist. Soc.* **84**, 90.

YULE & KENDALL (1940). *Introduction to Theory of Statistics*. London.

(*MS. received for publication* 15. I. 44.—Ed.)