Single measures of deprivation

Keith Folwell

Abstract

Study objective **– A number of measures have been developed which attempt to combine a range of variables into a single, more easily understood dimension of "deprivation". These extend from fairly simple additive measures through to those based on more sophisticated statistical techniques. All attempts to simplify a number of variables into a single, summary measure have limitations. This paper compares a number of more commonly used techniques and discusses their relative strengths and weaknesses.**

Design **– Data from the 1991 census is used to show the relative capabilities in discriminating between areas of (a) the Department of Environment's Z score index, a simple additive measure; (b) the Jarman index, not strictly a measure of deprivation but, apart from its importance to health workers, of interest as a weighted index to contrast to simple additive indices; (c) a multivariate technique, namely factor analysis, drawing on the London Research Centre's experience of its use; (d) the index of local conditions, commissioned by Department of the Environment from the University of Manchester.**

Conclusions **– Contrasting these different methodologies highlights relevant considerations in choosing a measure of deprivation, including ways in which the method of construction can dictate how a measure may be used. In particular, simple additive indices should be avoided as they hide too much information and if badly constructed can be meaningless, while weighted indices demand critical use since they tend to lack generality.**

(*J Epidemiol Comm Health* 1995;**49**(Suppl 2):S51–S56)

Deprivation is generally recognised as a composite concept, in that there is no single variable that can be said to measure it but rather a number of variables must be combined in some way. Thus, for example, poverty as measured by household income is usually recognised as an important component of deprivation, but other variables affecting quality of life may also need to be taken into account. Very often, data on important components of deprivation is lacking and proxy variables are used instead.

The methods used to combine component variables differ greatly and have surprisingly unremarked effects on the resultant measures. Researchers choosing a measure need to be aware of these differences in order to make a sensible choice. This paper examines representative examples of four different methodologies and discusses their relative merits and uses. It is aimed primarily at the researcher looking for a ready made deprivation measure. While many of the criticisms may be well known to those working in the field, they remain ignored for the most part and the non-specialist should be aware of the possible dangers of misinterpretation or misuse of a measure.

London Research Centre, 81 Black Prince Road, London SE1 7SZ
K Folwell.

Correspondence to:
K Folwell

Simple additive indices

The easiest way to combine a range of variables into a single measure is to add them up. One of the most commonly used indices during the 1980s was devised by the Department of the Environment.[1] It is sometimes referred to loosely as the Z score, as it is based on the addition of standardised scores for component variables. However, since the term Z score has a more precise meaning to statisticians, I have preferred to use the term Z score index.

The basic methodology for additive indices is summarised by the equation

$$\text{Index} = \sum_{i=1}^{n} z_i \qquad (1)$$

where n is the number of variables being combined into the index and z_i is the score on variable i standardised with respect to England and Wales, that is

$$z_i = \frac{x_i - \bar{x}_i}{s_i} \qquad (2)$$

x_i is the score for an area (ward, enumeration district), $\bar{x}_i$ is the England and Wales mean calculated over the same area level, and s_i is the SD for the same areas in England and Wales. Very often each x_i is transformed by a simple function such as a log or square root; the effect is to reduce skew and/or kurtosis in the probability distribution of a variable so that it approximates more closely to a normal, or Gaussian, curve.

The Department of the Environment used eight variables in constructing their original index.

An indication of how additive indices work, and how they can hide information, is illustrated in figures 1 and 2. In figure 1, a simplified index is calculated for three wards using only three variables. The three wards have very close scores on the combined index, but a glance at figure 2 shows that their profiles on the three component (standardised) variables are radically different. In particular, the shape of the line for Regent's Park ward in Westminster is an inverted version of that for

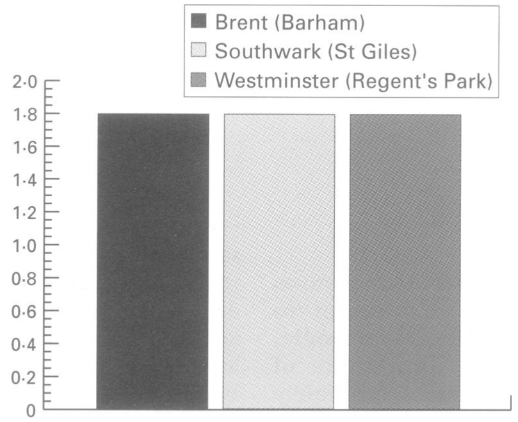

Figure 1 Scores on a simple additive index for three electoral wards in London.

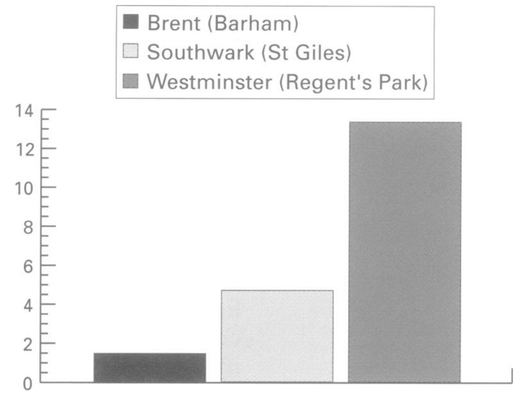

Figure 3 Weighted index – scores for three London wards.

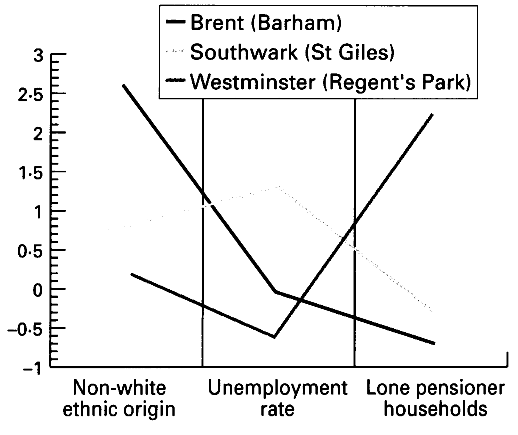

Figure 2 Scores on component variables for three electoral wards in London.

St Giles ward in Southwark; the former has slightly below average unemployment, the latter somewhat higher than average.

All methods of combining indices hide information in this way. The additive index assumes that the proportion of lone pensioner households is of equal importance to the concept of deprivation as the unemployment rate or the proportion of persons in households with a head born in the New Commonwealth or Pakistan (the proxy used for non-white ethnic origin). There is no prima facie case for believing this assumption and as we shall see below there are reasons to treat it with suspicion.

Weighted indices

The methodology for calculating a weighted index is broadly similar to simple additive indices, except that equation (1) changes to

$$\text{Index} = \sum_{i=1}^{n} w_i z_i \qquad (3)$$

where w_i is a weight by which variable x_i is multiplied after standardisation. This method accords each component variable a relative importance – that is, if the weight w_p is greater than w_q then variable x_p contributes more to the concept of deprivation under consideration than x_q. Changing the weights simply gives a different concept of deprivation, where different variables have different relative importances.

Jarman's underprivileged area score[2] is a good example of a weighted index, since it is well known and the weights were derived methodologically. The weights were derived by a survey of general practitioners, asking them to rate a number of measurable items as to the degree to which they were likely to increase a GP's workload. The index is not, therefore, a general measure of deprivation, but was specifically designed to help decide the allocation of resources in the health service.

Figure 3 shows how the addition of weights to an index changes the relative position of the three wards for which we calculated the Z score index previously. The same three variables were used, but this time with their Jarman weights (variables were chosen which are used in both the Z score index and the underprivileged area score). One area now stands out from the others – Regent's Park ward has a score about three times that of its nearest rival, St Giles ward in Southwark. This is because high weights are afforded to age related variables, since the very old and the very young add disproportionately to their numbers to the workload of a general practice.

Figure 3 makes it obvious why the underprivileged area score is not a very good general measure of deprivation, and conversely why it is a successful measure of issues affecting the health service.

Other deprivation measures such as the Townsend and Carstairs indices[3–5] use weights to adjust the relative contributions of component variables, but these weights are usually integers and are not derived by the application of any methodology. The Z score index is sometimes described as a weighted index, since the unemployment rate is often given a weight of 2. However, this is merely because at the time of its first calculation, 10% census data was not available and the preferred item – population in low socioeconomic groups – could not be known. Unemployment rate was effectively added in twice to preserve a balance between economic, housing, and so called "social" variables.

Multivariate techniques

It is not always possible to discern a methodology for deriving weights independently. Multivariate statistical techniques such as factor analysis and the related technique of principal components analysis offer a way of calculating composite scores using weights derived from the data itself. Highly correlated variables are grouped together on a single factor and factors may be designed to be orthogonal to each other, or at least correlating only very slightly. This is important since it makes explicit which component variables may reasonably be added together; figures 1 and 2 illustrate the problems which can arise if variables are combined inappropriately. The factor analysis computes a transformation matrix, which acts essentially as a set of weights. Factor scores are computed as

$$\mathbf{f} = M\mathbf{x} \qquad (4)$$

where **f** is a vector of factor scores, M is the transformation matrix and **x** is a vector of scores on the component variables. A score on the jth factor F_j is thus calculated as

$$F_j = \sum_{i=1}^{n} w_{ij} x_i \qquad (5)$$

where each w_{ij} is the element in the ith column and the jth row of the transformation matrix M.

A study conducted by the London Research Centre for the London Planning Advisory Committee (LPAC) identified eight factors from census of population data.[6] One was an economic factor, grouping such variables as unemployment rate and low socioeconomic group. Interestingly, age related variables (proportion of pensioner households, children under 5) were associated with a completely different factor. In this study factors were held completely uncorrelated with each other, and so the addition of economic and age variables as in the Z score index is invalidated. Economic and age variables will vary largely independently of one another, so that cases such as that illustrated in figures 1 and 2 will arise frequently. Economic and age related deprivation are conceptually different, and adding them together without an effective weighting system leads to an index which is meaningless since similar scores can mask vastly different profiles.

As $x_i \in \{x_1, x_2, \ldots, x_n\}$ where n is the number of variables and $F_j \in \{F_1, F_2, \ldots, F_m\}$ where m is the number of factors then M provides a mapping from an n-dimensional to an m-dimensional coordinate system. It thus provides not a single summary measure but a range of simpler measures which can be used to profile an area. Scores on an economic or age factor can be used on different occasions and for different purposes, separately or in conjunction, to describe an area, but should not be added together.

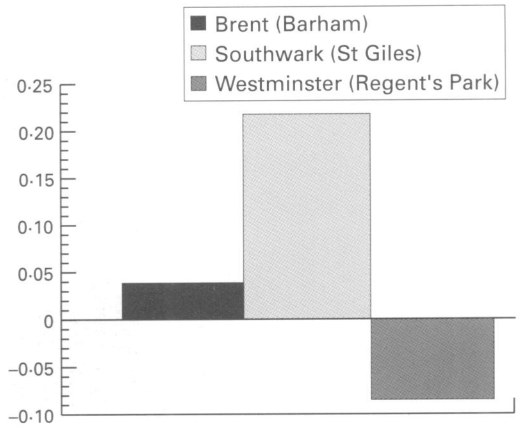

Figure 4 Scores based on an economic "deprivation" factor for three London electoral wards.

It is possible to extract as many factors as there were original variables, but since the purpose of factor analysis is to simplify data and because each additional factor tends to account for decreasing amounts of variance in the original variables, this is not normally done. However, variables which do not have high loadings on any of the more important factors can be considered not to have contributed much to the investigation; factor analysis can in this way be used as a tool for deciding which variables to drop from a study, as well as identifying weights for composite measures.

Figure 4 shows the same three London wards as before, plotting their scores on the "economic deprivation" factor from the LPAC study. Economic variables obviously have the highest weights when computing a score on this factor, and so it is no surprise to find the wards ranked in the same order as their unemployment rates from figure 2.

Critiques of multivariate analysis would most likely point out that the factors derived are highly dependent on the variables chosen originally. The addition or elimination of a few variables can give rise to entirely different factors, or affect the structure so that different variables are associated with different factors.

Signed χ^2

One further measure gaining interest is the Department of the Environment's index of local conditions, recently developed on their behalf by the Centre for Urban Policy Studies at the University of Manchester.[7 8] It is an additive index but differs in the method used to standardise scores. Since greater reliability can be credited to values on a variable measured against a larger base population, the index gives greater weight to variables in areas with larger populations. The standardisation method used is the signed χ^2 statistic, calculated as

$$\chi^2 = \mathrm{sign}\left(\log\left(1 + \frac{(O_{1i} - E_{1i})^2}{E_{1i}} + \frac{(O_{2i} - E_{2i})^2}{E_{2i}}\right)\right) \qquad (6)$$

Here, O_{1i} represents the observed value for variable i, E_{1i} the expected value (calculated by applying the rate for England to the base

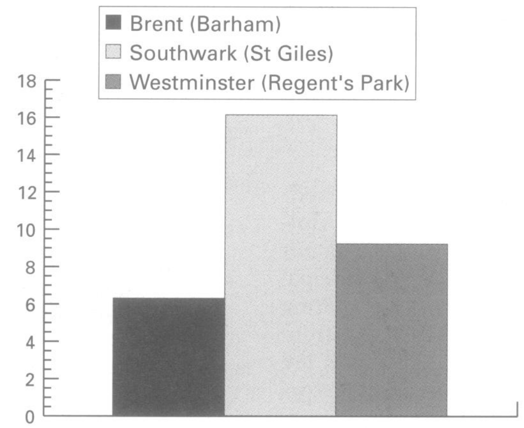

Figure 5 Index of local conditions scores for three London electoral wards.

population in the area), O_{2i} is the number in the relevant population without a characteristic and E_{2i} the expected value without a characteristic. Absolute numbers, rather than proportions, are used; for example, if O_{1i} is the number of the economically active who are unemployed, then O_{2i} will be the number of economically active who are not unemployed. The sign function is defined as

$$\text{sign}(x_i) = \left[\begin{array}{l} x_i, \text{ if } (O_{1i} - E_{1i}) \geq 0 \\ -x_i, \text{ if } (O_{1i} - E_{1i}) < 0 \end{array}\right. \qquad (7)$$

A composite score is then calculated as for an additive index, by summing the transformed values for the component variables

$$\text{Index} = \sum_{i=1}^{n} \chi_i^2 \qquad (8)$$

The rationale for a χ^2 transformation depends on the fact that a measure based on an area with a small population is less reliable than one based on an area with a large population. We have more confidence in a 30% unemployment rate if it is calculated from 100 economically active persons than 10. The problem is made worse with census data where counts can be randomly adjusted by ± 1 in small areas to preserve anonymity. The χ^2 calculation gives greater weight to figures based on large samples. However, this leads to difficulties in interpretation, most obviously since two areas with the same scores on component measures can end up with different χ^2 scores simply because they have different populations.

Figure 5 illustrates how our familiar three London wards perform on the index of local conditions. Since all the wards are broadly similar in size, the differences are more probably attributable to the particular mix of component variables used.

Discussion

This paper has described a variety of methods purporting to measure deprivation. It is apparent that there is no agreed definition of what deprivation is or how it should be measured. Researchers are left to choose one of a number of constructs fit for the purpose in hand. In some cases, the choice is predetermined; health authorities competing for funding will choose the Jarman index while local authorities will choose the Z score index or the index of local conditions, since these measures are accepted by the relevant funding authorities. For other purposes, the choice can be daunting.

Many researchers choose additive indices on the grounds that since they are constructed simply, they must be easy to understand and interpret. Nothing could be further from the truth; as we have seen, they are often constructed quite crudely so that what they are measuring can be at best questionable, at worst meaningless. Adding certain variables together is rather akin to adding apples and bananas together – it can be done, but the resulting figure of so many "fruit" hides information rather than illuminates. Weighted indices address this problem in effect by deciding that one banana is equal to a certain number of apples, which is effective for particular purposes when such equivalences can be shown to hold. Weights could be related to the number of calories or grammes of protein in an "average" apple or banana, for example. It is easy to see with this analogy that a set of weights based on the amount of protein would be appropriate for one purpose, but inappropriate for other purposes where calories would be a better choice. Factor analysis could be said to identify different ways of classifying and scoring the characteristics of fruit – for example, citrus fruit would have a high score on an "acidity factor" – but this is probably carrying the analogy too far.

A different set of considerations comes into play when comparisons over time are needed. It would normally be desirable to update indices using better information where available. For example, many indices constructed on 1981 census data (including Jarman and the Z score index) used the proportion of the population in households with a head born in the New Commonwealth or Pakistan as a proxy for ethnic origin. Now that ethnic origin is asked in the 1991 census it would make sense to use that in new indices. But if a comparison with 1981 is needed, this cannot be done without destroying the comparability of the two sets of data. Factor analysis is particularly vulnerable to lack of comparability. Since the factors are constructed from the interrelationships between the data variables, then when these interrelationships change so can the number and characteristics of the factors. Constructing scores from 1991 data on factors based on relationships between variables that held in 1981 does not necessarily tell you how the picture has changed. Here, the problem is similar to that experienced by economists in constructing a price index; as the prices of goods and services change, people's patterns of consumption change and the "basket" of goods and services that is used to calculate the index has to be altered to reflect this.

In making comparisons over time there is therefore a case for avoiding the combinatorial approach altogether and choosing a single representative variable that correlates highly with other variables one might wish to consider. Unemployment rate is a good proxy for most other economic variables and also correlates highly with a number of other measures such as standardised mortality ratios. Campbell, Radford, and Burton[9] advocate this approach in studies where the use of the Jarman index would be inappropriate. Even with single measures, however, there is often a problem of finding a consistent time series.

Mention might be made of multiple regression studies, where a number of variables are used to "predict" a score on an unknown measure. Input variables can be related by a regression equation to an output the value of which is known at one period in time, or at one geographical level, but is required at another time or for another area where only the input variables are known. The incidence of a particular disease, for example, may be known at a gross level but some indication may be needed at ward level: a regression equation can be constructed using predictor variables such as social class, housing amenities, and so on which are available at ward level from the census. The equation shows a similarity to the foregoing indices of deprivation

$$y = \sum_{i=1}^{n} a_i x_i + b \qquad (9)$$

where the regression coefficients a_i act as weights applied to the inputs x_i and b is an error term. Such measures are usually constructed to predict a specific outcome, such as morbidity or admission rates, and similar restrictions on their general applicability obtain as do to weighted indices such as the Jarman index. In addition, they do not always generalise from one area to another; Noble *et al*[10] for example regressed take up of housing benefit on a number of variables, and discovered different regression equations for Oxford and Oldham.

While occasions will arise where a measure is chosen for purely pragmatic reasons, we can conclude with some guidelines which have arisen in the foregoing:

- Simple additive indices should be avoided whenever possible. If a factor analysis shows that the component variables of an additive index are uncorrelated, then there is a high possibility that a meaningless measure will result. If the factor analysis shows a sufficient amount of covariance, then the use of the factor scores will balance the contribution of each component variable to the summary measure much better than the standardised score.
- Weighted indices should not be used uncritically, since the weights are usually only valid for a specific purpose. The researcher should be satisfied that the weights are appropriate for the purpose in hand and that they have been derived by a sound methodology.
- Comparisons over time present particular problems. In the absence of an agreed definition of deprivation and a means to measure it which is constant over time, decisions which are sometimes necessarily arbitrary can be avoided by choosing a single variable which correlates highly with other components of a summary measure.
- The use of standardisation techniques such as the signed χ^2 statistic should be considered very carefully. They are appropriate to small areas, especially where deliberate data corruption to preserve anonymity is known to have occurred. Caution should be exercised in interpreting scores, particularly where areas of unequal base population sizes are being compared. An additive index based on signed χ^2 is subject to the same criticisms as simple additive indices using other standardisation techniques – i.e., the addition of inappropriate variables can result in a meaningless measure.

1 Department of the Environment. *Urban deprivation.* Inner Cities Directorate. Information note no 2. London: Department of the Environment, 1982.
2 Jarman B. Identification of underprivileged areas. *BMJ* 1983;**286**:1705–9.
3 Townsend P, Phillimore P, Beattie A. *Inequalities in health in the Northern Region: an interim report.* Newcastle: Northern Regional Health Authority and Bristol University, 1986.
4 Carstairs V, Morris R. *Deprivation and health in Scotland.* Aberdeen: Aberdeen University Press, 1991.
5 Morris R, Carstairs V. Which deprivation? a comparison of selected deprivation indices. *J Public Health Med* 1991;**13**: 318–26.
6 London Planning Advisory Committee. *Regeneration areas: technical background report for discussion. Review of advice and guidance working paper.* London: London Planning Advisory Committee Economic Issues Working Party, 1993.
7 Department of the Environment. *Index of local conditions.* London: Department of the Environment, 1994.
8 Bradford M, Robson B, Tye R. An urban deprivation index 1991. In: Simpson S, ed. *Census indicators of local poverty and deprivation: methodological issues.* London: Local Authorities Research and Intelligence Association, 1993.
9 Campbell DA, Radford JMC, Burton P. Unemployment rates: an alternative to the Jarman index? *BMJ* 1991;**303**: 750–5.
10 Noble M, Cheung SY, Smith G, Smith T. Using housing benefit records in mapping poverty. *Policy and Politics* 1995;**23** (in press).

Open discussion

DRAPER – Mr Folwell, I did not understand your closing remark. You say the purpose is for multivariate analysis. Are you talking about component analysis/factor analysis, where you are looking at the internal structure of the data, or are you talking about regression analysis – in which case you need some single "objective" measure, as it were, as your dependent variable in the analysis?

FOLWELL – I was advocating the component analysis/ factor analysis because it gives flexibility as well. If there are a number of factors, one of which is age related and another economics related, the age related factor can be used if age is important to health problems, for one application, and the economic factor for another.

DIGGLE – From the statistician's perspective, if you have a number of explanatory variables then what you do with them depends on the question you are trying to answer. Multiple regression methodology is designed to give you the best linear predictor from all your explanatory variables, so if you want to predict a response, then the appropriate combination of factors depends on that response. There is no mystery in that. Presumably the indices have been constructed for other purposes because if you construct an index as a linear combination of factors then it cannot be better for prediction than multiple regression, but it might be better for other purposes.

Therefore the question you are trying to answer has to be defined before you can even discuss what a good index is.

FOLWELL – I would not suggest using factors which are uncorrelated as components in an index: I would suggest using each factor as a separate index. You would not expect necessarily to find linear relationships. Because they are uncorrelated, relationships between a score on one factor and a score on another would not be expected. What I am really saying is that you cannot have a single component measure, you need to look at different dimensions and to plot areas in multidimensional space.

GORDON – I disagree with not using factors that are uncorrelated. If deprivation is multivariate then surely factors that are uncorrelated are wanted because they are measuring different aspects of deprivation. Variables that are highly correlated are actually measuring the same thing. An index composed of highly correlated variables is therefore very unreliable because it is measuring the same thing over and over again.

FOLWELL – I do not think you are disagreeing: I am saying these factors are uncorrelated and that is what is wanted. You take many correlated variables and you end up with a few uncorrelated ones, which gives you different dimensions of this thing called "deprivation" you are trying to measure. But they are different flavours and probably completely different concepts, like age related deprivation and economic deprivation. I would suggest that these are totally different concepts, and therefore it is false to add them together.