DOCUMENT RESUME

ED 464 195 UD 034 990

AUTHOR	Gorard, Stephen; Taylor, Chris
TITLE	A Comparison of Segregation Indices Used for Assessing the Socio-Economic Composition of Schools. Measuring Markets: The Case of the ERA 1988. Working Paper.
INSTITUTION	Cardiff Univ. (Wales). School of Social Sciences.
SPONS AGENCY	Economic and Social Research Council, Lancaster (England).
REPORT NO	WP-37
ISBN	ISBN-1-872330-347
PUB DATE	2000-00-00
NOTE	39p.
CONTRACT	R000238031
AVAILABLE FROM	For full text: http://www.cardiff.ac.uk.socsi/markets.
PUB TYPE	Reports - Evaluative (142)
EDRS PRICE	MF01/PC02 Plus Postage.
DESCRIPTORS	Disadvantaged Youth; Educational Discrimination; Educational Research; Elementary Secondary Education; Foreign Countries; *Indexes; *Poverty; Research Methodology; *Socioeconomic Influences; *Socioeconomic Status
IDENTIFIERS	England; Wales

ABSTRACT

This paper examines the changing socioeconomic composition of English and Welsh schools over 11 years following the Education Reform Act of 1988, using many indicators of socioeconomic disadvantage and methods of measuring the changing spread of disadvantage among schools (segregation). It explains how to assess segregation (percentage differences and indexes of segregation), then details the isolation index, the dissimilarity index, the Gini coefficient, and the segregation index, an alternative to the others. These indexes form a kind of hierarchical family. Using data on all English and Welsh schools over 11 years, results find high correlations for patterns of poverty among the various indexes. The segregation index and the dissimilarity index are the most highly correlated, while the isolation index has the strongest relationship with the population composition. The choice between the segregation index and the dissimilarity index and the others may make little practical difference in real-life situations, though differences are important and worthy of further investigation. The paper notes that the segregation index is the only index that can separate the overall relative growth of patterns of poverty from changes in the distribution of patterns of poverty among schools. (Contains 65 references, 17 tables, and 6 figures.) (SM)

ED 464 195

CARDIFF UNIVERSITY
PRIFYSGOL CAERDYDD

MEASURING MARKETS: THE CASE OF THE ERA 1988

www.cardiff.ac.uk/socsi/markets

A Cardiff University School of Social Sciences ESRC-funded Project

WORKING PAPER 37

A comparison of segregation indices used for assessing the socio-economic composition of schools

Stephen Gorard and Chris Taylor

2000

Cardiff University School of Social Sciences
Glamorgan Building
King Edward VII Avenue
Cardiff
CF10 3WT

Email: gorard@cardiff.ac.uk

ISBN 1 872330 347

A comparison of segregation indices used for assessing the socio-economic composition of schools

Stephen Gorard and Chris Taylor

Cardiff University School of Social Sciences

Glamorgan Building

King Edward VII Avenue

CF10 3WT, UK

email: gorard@cardiff.ac.uk

Introduction

This paper stems from research funded by the UK Economic and Social Research Council (Grant No. R000238031), so far represented by Gorard and Fitz (1998a, 1998b, 2000a, 2000b), Gorard (1999a, 1999b, 2000a, 2000b, 2000c, 2000d), Taylor et al. (2000a, 2000b) and White et al. (1999a, 1999b). Using official statistics, we examine the changing socio-economic composition of all schools in England and Wales over an 11 year period following the Education Reform Act of 1988. In the study we have used many indicators of socio-economic disadvantage, and many methods of measuring the changing spread of disadvantage between schools (segregation). We have found one approach most useful, and this paper attempts to explain why.

The role and validity of various indices of segregation have been a focus of considerable debate and speculation over the last fifty years in social science research. Similar debates to ours, concerning the relative changes over time in the socio-economic composition of the student intake to schools, have occurred in many fields including the analysis of: residential patterns by ethnicity; gendered patterns of occupation, and polarised income patterns in family economics. The debate starts from three important assumptions. First: in many respects an analysis of segregation between schools is very similar to an analysis of segregation between any other institutions or organisational categories, and to more general analyses of societal inequality. Second: there are many different interpretations of the term 'segregation'. Third: there are consequently many alternative indices, or methods of calculating the level of 'segregation' in an institution such as a school. Each of these issues is dealt with briefly in this paper. However, our main purpose here is to describe a typology of the common segregation indices, and from this to consider their potential benefits and drawbacks for our primary empirical concern with the distribution of poverty and other indicators of disadvantage in British schools. We therefore show here, for the first time, several major advantages in terms of compositional invariance in using what we describe as the segregation index (S) over the more common dissimilarity index (D). Specifically we show that D is influenced by population composition under certain circumstances in which S is not. S, on the other hand, is not symmetrical having two components (S and S*) which when added together produce D. We also show that in various guises, S is already a commonly used measure of inequality and unevenness, probably as popular as D. The choice between these two thus becomes one between symmetry and strong compositional invariance.

What is segregation?

We start by defining what we mean here by 'segregation' and therefore what this paper is, and is not, about. This is necessary because of confusion within the literature, perhaps caused by 'the complexity and ambiguity of the concept of segregation..[and] a reluctance on the part of those with substantive interests to articulate their concepts and justify their selection of measures' (James and Taeuber 1985, p. 24). It is closely related to concepts such as polarisation, stratification and inequality, and may even be considered identical to these in certain conditions. Massey and Denton (1988) and Massey et al. (1996) identify five dimensions of segregation, of which only two concern us here - namely evenness and exposure. The other three - concentration, centralisation and clustering - while different in conception are relatively unclear theoretically, methodologically, and empirically. It is surely significant that when Massey and Denton examined 20 indices of segregation for their underlying latent variables via factor analysis, these three were not robust factors and not clearly delineated from each other or from the first two.

In residential segregation, clustering is a measure of the extent to which a section of the population lives in enclaves rather than being spread out over a region. Centralisation is a measure of the extent to which a section of the population occupies a central core in an urban region. Concentration is a measure of the relative space per head occupied by a section of the population, or how densely packed a group is in spatial (rather than organisational) terms. None of these meanings of segregation captures the meaning used in this paper, although as noted below all such measures are correlated to some extent. Concentration is commonly measured in terms of the Hoover Index:

$\Delta = 0.5 * \Sigma(|Xi/X-Ai/A|)$

where Xi is the number of people with characteristic X in organisational unit i, X is the total for all units, Ai is land area of unit i, and A is the total land area.

The formula is presented here because of its similarity to those used below. However, for the rest of this paper there is no consideration of the surface area of each organisational unit. Our work concerns the socio-economic characteristics of the intake to schools, and although it is concerned with the number of students in each school, it is not concerned with the land area covered by each.

One of the robust components of segregation in the Massey and Denton study is exposure (or its inverse - isolation), a measure of the extent to which members of a minority group interact with the majority group or with each other (see also Lieberson and Carter 1982). This definition has been used by researchers who are more interested in interactional probability than in spatial dissimilarity (Boal 1987). We consider this approach below because of its relative popularity compared to the other three.

For our study however, the key element of segregation is evenness. Segregation is a measure of the evenness or unevenness of distribution of individual characteristics between organisational units (Meuret 2000). In this we are in agreement with many of the major writers in the field. James and Taeuber (1985), for example, describe segregation as referring 'to the differential distribution of social groups among social organizational units' (p. 24), and writing of occupational segregation Blackburn et al. (1995) state that 'segregation is the tendency for women and men to be employed in different occupations. Such segregation creates gendered occupations which are disproportionately "female" or "male"' (p. 320). Very much the same logic of differential distribution applies to patterns of school attainment (Gorard 1999a), to polarisation where high and low-scoring students are disproportionately distributed between schools, genders, ethnic groups and so on. It is considered synonymous with stratification (in the social science sense, Meuret 2000), and can also be conceived as a measure of association, of inequality, and of social mobility. Le Grand (1999) points out, in relation to the definition of social justice, that popular notions cannot be denied in conceptualising such a process. He shows, for example, that both extreme redistribution of wealth without merit and complete laissez-faire can be considered unfair and would be publicly perceived as such. A similar point can be made here. Popular notions of segregation, and its unfairness, would stem primarily from evenness not exposure (or concentration etc.).[1]

To illustrate this point, imagine a society with two schools, each having the same number of students (100 for example). If one school contains all of the female students and the other contains all of the male students, then we could describe this school system as totally segregated in terms of gender. Now scaling the number of female students in both schools (by doubling to 200 for example) does not change this total segregation, nor does scaling the number of males in both schools (by decreasing to 50 for example). These situations are summarised in Table 1. Segregation is unaffected here by the actual numbers of female or male students in the society, and any measure of this segregation should therefore be 'composition invariant'.

Table 1 - alternative versions of total segregation

School	Females	Males	Total
A	100 (200)	0 (0)	100 (200)
B	0 (0)	100 (50)	100 (50)
Total	100 (200)	100 (50)	200 (250)

On the other hand, the situation of no segregation by gender occurs when each school has its 'fair share' of both groups of students. In our example, if both schools contain 50 female and 50 male students then there is no segregation (in this respect). Doubling the number of female students in both schools leads to no change in this pattern, nor does halving the number of males for example. These situations are summarised in Table 2. Again, segregation is unaffected here by the actual numbers of female or male students in the society, and any measure of this segregation should therefore be 'composition invariant'.

Table 2 - alternative versions of no segregation

School	Females	Males	Total
A	50 (100)	50 (25)	100 (125)
B	50 (100)	50 (25)	100 (125)
Total	100 (200)	100 (50)	200 (250)

In making this statement of invariance, of course, there is no suggestion that nothing has changed in the two schools systems. The situation in Tables 1 and 2 changes from one of gender balance overall, but with total (or no) segregation, to gender imbalance with more female students in the system, still with total (or no) segregation. It is also true that in the second version of Table 2 males are now less likely to meet other male students than they were before. These differences could be important social science findings but they are not ones that are relevant to our pursuit of measures of segregation.

In real-life, of course, both of the above extremes are unlikely. Any system of allocating students to schools, in Britain at least, is likely to lead to some but not total segregation. In fact, Cortese et al. (1976) suggest that evenness is an unrealistic demand and that segregation should be assessed against a random distribution of population characteristics. There is much merit in this proposal, although it leads to rather cumbersome calculations, but it is important to realise that the two approaches are entirely compatible. If it is assumed that there is a random component in any distribution, it follows that small amounts of 'segregation' assessed by an unevenness index, and small differences between amounts of segregation, may not be significant. What the analyst is seeking for social science comment are strong and consistent trends, or large consistent differences rather than small fluctuations over time or place.

How do we assess segregation?

In calculating patterns of distribution, it is essential to look at the figures involved as proportions of each other (Gorard 2000a). Clearly, we cannot estimate patterns of segregation using 'raw' figures. What are the alternatives?

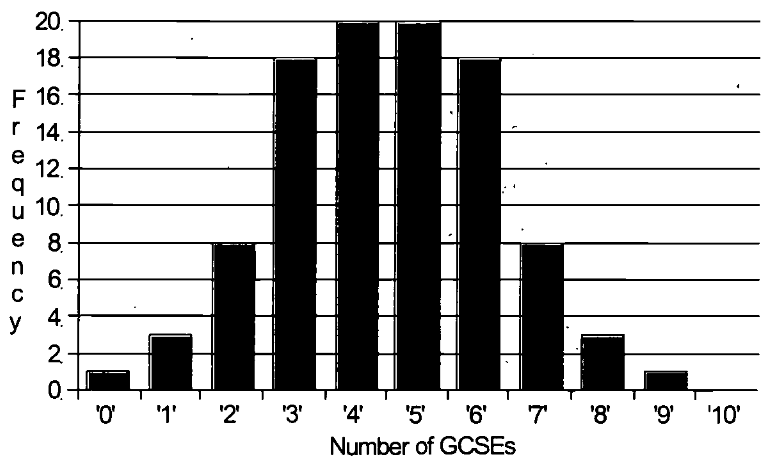
Percentage differences

One of the least conceptualised, but most common, methods of assessing inequalities between groups is the use of percentage point differences. They are in common usage in media reporting, policy-making, and even some academic writing (see Gorard 1999b). In summary, the method consists of subtracting percentages to estimate inequalities. For example, if 30% of boys and 40% of girls gain a C grade in Maths GCSE in one year the difference between them is said to be 10%. If 35% of boys and 46% of girls gain the equivalent a year later, the difference is said to be 11%, so the improvement among girls is said to be greater. This approach has been used in studies of segregation (Gibson and Asthana 1999), as well as of inequalities in school outcomes. The problems with this approach are manifold. First: the difference between two percentages is not necessarily a percentage and therefore a comparison of two differences in percentage points between four percentages (as above) may not be comparing like with like. Second: it is not even clear that percentages lie on an equal interval scale, as they would need to for such calculations. Third: 'additive' comparisons such as this take no account of changes in the scale of the figures involved. Fourth: superior but equally simple 'multiplicative' methods are available instead.

For example, a school manager might note that the number of children with dyslexia in the school has increased in the last five years. The manager might correctly calculate this number as a proportion of the school roll over five years, which is necessary since the school may also have changed in size. If the number of students with dyslexia has increased as a percentage of those in the school, the manager may believe that this calculation *proves* that the school is taking an increasing share of dyslexics compared to other schools. Such a conclusion should, of course, actually involve a consideration of the picture in those other schools, or at least a consideration of changes in the frequency of dyslexia in the population. It is possible that the manager is totally in error, and that the school is actually taking a *decreasing* share of the dyslexics from the population since the growth of dyslexia in the population may be out-stripping the growth in that school. To put it simply, when examining changes over time it is important to recall that all of the numbers involved are liable to change. The question for the school manager interested in student share is not whether the percentage of dyslexics in the school has grown, but whether it has grown faster than the percentage in the population from which the school derives its students. Put in arithmetic terms, the index for any year would be: $(d/n)/(D/N)$ where d is the number of dyslexics in the school, D is the number in the population, while n is the number of all students in the school, and N is the number in the population. If this index grows over time, then the manager can justly 'complain' of increasing share, but it should be noted that the index could decrease even where d/n increases (and vice versa). This is one of the main reasons why small-scale studies of changes over time may be of little worth.

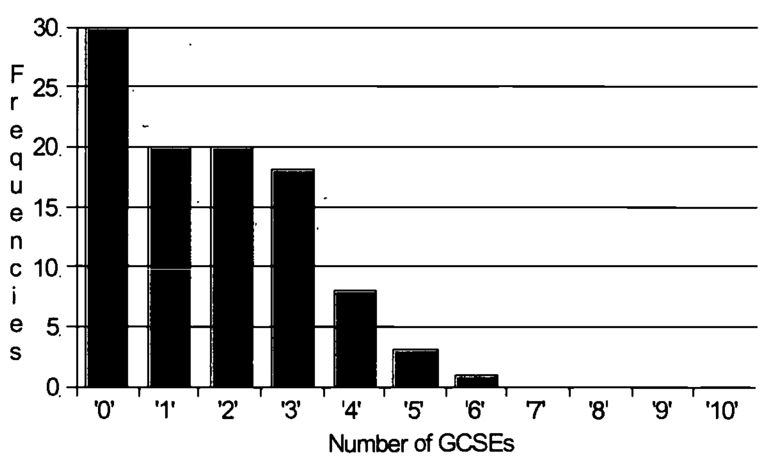
According to Linacre (2000), even if percentages are interval in nature they are not usually linear in function, therefore their arithmetic manipulation does not maintain the substantive implications related to their underlying variable (the raw figures). If the latent variable and its percentages have an ogival-shaped relationship for example, then it is easy to overestimate the rate of change at mid-range values at the expense of changes near 0 or 100. For example, the GCSE benchmark is a widely used figure in educational policy which is expressed as a percentage (of candidates gaining five or more GCSEs at grades A*-C). Since the number of GCSEs per student in any sizeable group may be approximately normally distributed, it can be argued that changes in the percentage gaining five (or any other arbitrary number) would be less significant near the 50% mark. In fact, it is almost certain that if the distribution of qualifications is not uniform in the population, then changes over time will not be linear. At the 50% mark, where the distribution is taller, a small movement along the x axis (representing a change in the number of GCSEs per student) would produce a disproportionately large change in the percentage attaining the benchmark. At either end (near 0 and 100%), a much larger change on the x axis would be needed to produce the same effect. This may be illustrated by the imaginary data presented in Figures 1 and 2, which represent the frequency in one school of students gaining 0 to 10 GCSEs (at grade C for example). In Figure 1, exactly 50 of 100 students gain five or more GCSEs. If all students gained one more GCSE, the benchmark figure for the school would rise to 70%.

Figure 1 - Distribution of GCSEs among candidates (high score)



In Figure 2, only 4 of 100 students gain five of more GCSEs (the actual pattern of frequencies for Figure 2 is the same as Figure 1 but moved three places towards the origin of the x axis). In this case, if all students gained one more GCSE, the benchmark figure for the school would rise to 12%. Increases in terms of absolute percentage scores are much more difficult for low-attaining schools.

Figure 2 - Distribution of GCSEs among candidates (low score)



Finally, although arguments can and have been made for using either multiplicative (see below) or additive (e.g percentage points) methods as measures of association in *one* table, the chief problem lies in their different results when comparing the patterns in two or more tables. Using percentage points does not take into account the proportion (a+b)/(c+d) which accounts for rises and falls in the frequency of the phenomenon being observed. Using this method a commentator can have no genuine idea of the significance of the resulting points difference. This can be emphasised in two ways. First: if 1% of men were MPs but 0% of women were, this would be an enormous difference and one that social science commentators would be right to draw attention to. On the other hand, if 75% of men and 76% of women were in paid employment, the difference may be of little account. However, both examples yield a score of 1 point in the additive method, suggesting that this method is fine for a rough guide to the presence or absence of a pattern, but of little value as a scaled measure of achievement gaps.

Despite these now well-publicised flaws, politicians, media commentators and even some academic writers continue to make asinine comparisons over time or space based on percentage point differences. For example, Gibson and Asthana (1999) claim that the gap in terms of GCSE performance between the top 10% and the bottom 10% of English schools grew significantly from 1994 to 1998. Their figures are reproduced in Table 3, which shows the proportion of students attaining five or more GCSEs at grade C or above (the official benchmark), for both the best and worst attaining schools in England. It is clear that the top 10% of schools has increased its benchmark by a larger number of percentage points than the bottom 10%. The authors conclude that schools are becoming more socially segregated over time, since 'within local markets, the evidence is clear that high-performing schools both improve their GCSE performance fastest and draw to themselves the most socially-advantaged pupils' (in Budge 1999, p.3).

Table 3 - Changes in GCSE benchmark by decile

Decile	1994	1998	Gain 94-98
Top	65.0%	71.0%	6.0
Bottom	10.6%	13.1%	2.5

This conclusion would be supported by a host of other commentators using the same method (including Robinson and Oppenheim 1998, and Chris Woodhead, in the Times Educational Supplement 12/6/98, p.5). Similar conclusions using the same method have been drawn about widening gaps between social classes (Bentley 1998), between the attainment of boys and girls (Stephen Byers, in Carvel 1998a, Bright 1998, Independent 1998), between the performance of ethnic groups (Gillborn and Gipps 1996), and between the results of children from professional and unemployed families (Drew et al., in Slater et al. 1999). In fact, it is almost impossible to overestimate the pervasive impact and widespread nature of this 'politicians error' on educational policy in the UK.

A multiplicative method on the other hand, using the same figures, might produce a result like Table 4. Although the difference between the deciles grows larger in percentage points over time, this difference grows less quickly than the scores of the deciles themselves. On this analysis, the achievement gaps are getting smaller over time (Gorard et al. 1999a, 2000). This finding is confirmed by the figures in the last column showing the relative improvement of the two groups. The rate of improvement for the lowest ranked group is clearly the largest. The bottom decile would, in theory at least, eventually catch up with the top decile (Gorard 1999a).

Table 4 - Changes in GCSE achievement gaps by decile

Decile	1994	1998	Ratio 1998/1994
Top 10%	65.0%	71.0%	1.09
Bottom 10%	10.6%	13.1%	1.24
Achievement gap	72.0%	68.8%	

Of course, if percentages are not linear in form (see above) then neither of these comparisons can take place and it is not possible on these figures to say whether a gap is increasing or decreasing. It should, however, be noted that in this example the group with the largest additive growth is nearer the more sensitive middle area of the distribution (65%), while the supposedly weaker group is nearer one of the margins (10%). The disconcerting thing about the crisis commentators is not that they are wrong (errors of analysis are common to all of us), but that they are so influential despite being wrong.

While rates are expressed in simple percentage terms are 'absolute', relative rates (such as odds ratios) are margin-insensitive in that they remain unaltered by scaling of the rows or columns (as might happen over time for example). This difference is visible in changes to the class structure and changes in social mobility. In Table 5, 25% of those in the middle class are of working class origin, whereas in Table 6 the equivalent figure is 40% (from Marshall et al. 1997, pp. 199-200). However, this cannot be interpreted as evidence that Society B is more open than Society A as the percentages do not take into account the differences in class structure between Societies A and B, nor their changes over time ('structural differences').

Table 5 - Social mobility in Society A

	Destination middle class	Destination working class
Origin middle class	750	250
Origin working class	250	750

Table 6 - Social mobility in Society B

	Destination middle class	Destination working class
Origin middle class	750	250
Origin working class	500	1500

Relative rates are commonly calculated as odds ratios [(a/c)/(b/d)], cross-product ratios [ad/bc], or disparity ratios [a/(a+c)/b/(b+d)]. Disparity ratios are identical to the segregation ratios used by Gorard and Fitz (1998a, 2000a). Odds ratios estimate comparative mobility changes regardless of changes in the relative size of classes, and have the practical advantage of being easier to use with loglinear analysis (Gilbert 1981, Goldthorpe et al. 1987, Gorard et al. 1999b). 'From the point of view of social justice... this is of course both crucial and convenient, since our interest lies precisely in determining the comparative chances of mobility and immobility of those born into different social classes - rather than documenting mobility chances as such' (Marshall et al. 1997, p.193). The cross-product ratio for Table 5 is 9, and for Table 6 it is also 9. This finding suggests that social mobility is at the same level in each society, despite the differences in class structure between them. In many ways this crucial difference in approach and findings is being repeated in the more recent dispute over changes in segregation between schools in the UK.

Some previous work has confounded changes in social fluidity with changes in the class structure. Nevertheless, disagreement about the significance of absolute and relative mobility rates continues (e.g. Clark et al. 1990, pp. 277-302). Gilbert (1981) concluded that 'one difficulty with having these two alternative methods of analysis is that they can give very different, and sometimes contradictory results' (p.119). The similarities to the issue concerning achievement gaps are fairly obvious. In each case, different commentators use the same figures to arrive at different conclusions. One group is using additive and the other is using multiplicative models. Very similar analyses also occur in social science more generally, and similar problems have arisen in health research (Everitt and Smith 1979), in studies of socio-economic stratification and urban geography (Lieberson 1981), in occupational gender segregation (Blackburn et al. 1995), in social mobility work (Erikson and Goldthorpe 1991), and in predictions of educational pathways (Gorard et al. 1999b). Results are disputed when an alternative method of analysis produces contradictory findings. Some of these debates are still unresolved, dating back to what Lieberson (1981) calls the 'index wars' of the 1940s and 1950s, perhaps even to the work of Wright (1937). In each case the major dispute is between findings obtained using absolute rates ('additive' models) and those using relative rates ('multiplicative' models).

Indices of segregation

In contrast to the method of percentage points, most multiplicative approaches to measuring segregation are based on standard indices. As might be expected, given the existence of at least five dimensions of segregation, the number of indices proposed is high. Most of these indices are based on (usually dichotomous) categorical variables and these are the type discussed here. However it should be recalled that an entirely different approach may be appropriate for measures of inequality based on real numbers, such as the technique of matching marginal totals pioneered by Blackburn and Jarman (1997). Although matching marginals are used with categories they can only be used where the categories are divisions on a continuum (as may be considered the case with occupational class or income but not pass/fail or sex). This technique and many others are also limited to two cases or organisational units. For example, Yule's Q is the standardised odds ratio for a two-by-two table: (bc-ad)/(bc+ad). It has most of the desirable criteria described below, but it limited to only one size of table. Another assumption made for the purposes of this discussion is that the indices are used with official statistics and other secondary census-type data. Thus, there is no need to consider errors in sampling or tests of significance. Where a sampling approach is used, Ransom (2000) suggests a technique to assess whether apparent changes in segregation are statistically significant.

Given that the number of candidate indices is large, we need to consider possible criteria for making judgements about their relative value (and choice of an index implies a value judgement about the nature of inequality, Ransom 2000). James & Taeuber (1985), for example, suggested at least three such criteria for successful indices to satisfy:

Composition invariance – The index should be unaffected by scaling of columns or rows, through increases in the 'raw' figures which leave the proportions otherwise unchanged.

Organisational equivalence – The index should be unaffected by changes in the number of sub-areas, by combination for example of two sub-areas on the same 'side' of the line of no segregation.

Principle of transfers – The index should be capable of being affected by the movement of one individual from sub-area to sub-area.

For any analysis of segregation over time both composition invariance and occupation invariance are key to our understanding of a useful measure. 'Compositional invariance refers to the invariance of the index, following uniform changes in the number of males and females in each occupation reflecting the overall, but typically unequal, percentage changes in male and female employment' (Watts 1998, p.490). Thus, if the minority proportion (female employees for example) were to double in all occupations any good measure of segregation would remain unchanged. 'Occupations invariance requires that the measure of segregation be invariant to changes in the relative size of occupations if the gender composition of these occupations remains constant' (Watts 1998, p. 490).

Thus, if the number of males in one occupation were to double and the number of females also doubled in the same occupation, again any good measure of segregation would remain unchanged. These two criteria would ensure that the measure of segregation would not be affected by either a proportionate increase in the absolute levels of a particular group across all sub-areas, or a proportionate increase in the absolute levels of all groups in a particular sub-area.

It is important to recall that compositional invariance is not a simple concept (Kalter 2000). Blackburn and Jarman (1997) suggest that all statistics of association are dependent on the absolute size of the marginal totals in any table. Some indices have marginal independence with respect to one pair of marginal totals, but only in the unlikely social event of a row or column being multiplied by a constant. In addition, most measures of segregation, including matching marginals tend to vary with the number of organisational units (Blackburn et al. 1999). Thus, it is extremely difficult to find any indices which meet a strict version of these criteria.

Nevertheless, using concepts such as these we can begin to classify indices in terms of their match. For example, the Sex Ratio (or Hakim) Index, once popular, is now generally seen as flawed on three main counts (Tzannatos 1990). It is not easily interpretable, it is theoretically unbounded (never able to reach complete segregation for example), and most importantly it is very sensitive to changes in population composition (leading to confusion between simple population changes and changes in the distribution of population elements between organisational units). Similar comments apply to the Entropy or Information [Theory] Index (H), and to the Variance Ratio. The Atkinson Index (A) is more complex to use than most indices, is difficult to interpret, and does not allow direct comparison between studies even where two studies have used it (Massey and Denton 1988). This is because A is really a family of indices (Kalter 2000), whose use depends on the researcher's judgement in weighting of different parts of the segregation (Lorenz) curve. The complexity of several indices appears to stem from their era of origin where it was considered easier to square values to eliminate negatives than to use their modulus (despite the fact that the act of squaring may lead to some distortion of results).

The remainder of this paper considers three indices in more detail, the isolation index (proposed as best for exposure by Massey and Denton 1988), the dissimilarity index and its close relative the Gini coefficient (proposed as best for evenness by Massey and Denton 1988), and the segregation index (proposed here as a viable alternative to all of the above).

Isolation Index (and other P*-type measures)

There has recently been something of a revival in the use of measures of exposure, and to a large extent these measures are not competing with measures of evenness. They are measuring something different. They do, however, have two characteristics that mark them out from the more usual indices of inequality. They are asymmetrical and they are not composition invariant (Lieberson 1981). Asymmetry means, for the typical case of two population categories, that the measure gives different values for each category. This, in itself, is not a problem as long as the asymmetry is clear and the two values are not substantively contradictory. The composition invariance is also not a problem, according to advocates, since these 'P*-type' measures were designed in such a way, to describe the relative isolation of a group by incorporating both the uneven distribution of groups and their relative composition (Lieberson and Carter 1982).

Massey and Denton (1988) identify two key measures as the 'best' available for assessing segregation as exposure. The Interaction Index (I) is a measure of the extent of interaction between two groups. The Isolation Index (I*) is the inverse - a measure of extent of isolation of one group. For a table of the form:

	Minority	Majority	Total
Unit 1	A_1	B_1	T_1
Unit 2	A_2	B_2	T_2
....			
Unit n	A_n	B_n	T_n
Total	A	B	T

and a region with n sub-areas in which segregation may take place, and i varies from 1 to n, the interaction/isolation index may be defined as:

$I = \Sigma (Ai/A).(Bi/Ti)$

$I^* = \Sigma (Ai/A).(Ai/Ti)$

These measures have one key advantage over more common measures of distribution, such as the Dissimilarity Index (D, see below). They are sensitive to transfers even between organisational units on the same side of the line of no segregation (that is they adhere to a strong interpretation of the principle of transfers). Whereas I and I* change to reflect movement between schools, D produces the same value for the school systems in Tables 7 and 8, even though it can be argued that the schools in Table 8 are more segregated (since SchoolA now has no disadvantaged students at all).

Table 7 - Comparison of indices on principle of transfers

School	FSM	non-FSM	Total	I	I*	D	S	S*
A	10	90	100	0.01	0.13	0.25	0.19	0.06
B	10	90	100	0.01	0.13	0.25	0.19	0.06
C	50	50	100	0.36	0.36	0.50	0.38	0.12
Total	70	230	300	0.39	0.61	0.50	0.38	0.12

[note: the derivation of, and comparison with D, S and S* are described below]

Table 8 - Comparison of indices on principle of transfers

School	FSM	non-FSM	Total	I	I*	D	S	S*
A	0	100	100	0	0	0.43	0.33	0.10
B	20	80	100	0.06	0.23	0.06	0.05	0.01
C	50	50	100	0.36	0.36	0.50	0.38	0.12
Total	70	230	300	0.41	0.59	0.50	0.38	0.12

This could be a clear and convincing advantage under certain circumstances, but is more than over-shadowed by the obvious disadvantages of P*-type measures. For example, Tables 9 and 10 produce the same value for I (and I*) even though there is clearly no segregation, injustice, inequality, or unevenness in the school system of Table 9, and there is equally clearly some segregation in Table 10. These measures of exposure are therefore absurd as measures of segregation, bearing no relation to the Le Grand (1999) notion of defining fairness by at least taking into account popular conceptions.

Table 9 - No segregation

School	FSM	non-FSM	Total	I	I*	D	S	S*
A	20	80	100	0.25	0.08	0	0	0
B	20	80	100	0.25	0.08	0	0	0
C	20	80	100	0.25	0.08	0	0	0
Total	60	240	300	0.75	0.25	0	0	0

Table 10 - Some segregation

School	FSM	non-FSM	Total	I	I*	D	S	S*
A	0	100	100	0	0	0.36	0.33	0.03
B	0	100	100	0	0	0.36	0.33	0.03
C	20	80	100	0.75	0.25	0.73	0.67	0.06
Total	20	280	300	0.75	0.25	0.73	0.67	0.06

In addition, consider the situation where all students grew alike over time (for example if poverty was eradicated) then while standard measures of segregation would produce a value of zero (no segregation by definition), the Isolation Index produces a value for maximum segregation. Despite these absurdities, Noden (2000) advocates the use of exposure indices as a measure of socio-economic segregation between schools, and is taken seriously by what we have termed 'crisis commentators' on British education since the method appears to show that segregation has grown over the past decade. In fact, Noden's analysis is to large extent determined by the growth of indicators of poverty not the unevenness of their distribution (Gorard and Fitz 2000).[2]

Dissimilarity Index

The Dissimilarity Index (D), similar to the standardised segregation index (Coleman et al. 1982), is the most commonly used for all types of areal social segregation (Taeuber et al. 1981). While there was a period of what have been termed 'index wars' by Peach (1975) which started around the time of Wright (1937), the 'Pax Duncana' apparently crowned the Dissimilarity Index as the premier of all measures. Duncan and Duncan (1955a) showed that all other measures can be encapsulated by a judicious use of the population composition figures and D. Their paper presented a number of segregation indices and showed that they were all related to the segregation or Lorenz curve and, hence, to each other. However, it was another article by the same authors (1955b), which made explicit use of the Dissimilarity Index for their own research, that may have proved the catalyst for the current extensive use of D as a measure of segregation (see Lieberson 1981). Its advantages are that it is essentially independent of population composition (Taeuber et al. 1981), and that it is easy to comprehend while covering the same empirical ground as its 'competitors' (Massey and Denton 1988).[3]

Potentially D varies from 0 to 1, and represents the proportion of minority members who would have to move between units to achieve a perfectly even distribution. It 'describes the percentage of one group or the other which would have to move if there was to be no segregation between the groups' (Lieberson 1981, p.62). For a region with n sub-areas in which segregation may take place, and i varies from 1 to n, the index of dissimilarity may be defined as:

$D = 0.5 * \Sigma |Ai/A - Bi/B|$.

Using gender segregation by occupation as an example: Ai and Bi are the number of mutually exclusive cases in occupation i, giving a total of Ti cases in occupation i, A is the sum of Ai where i varies from 1 to n (the number of occupations), B is the equivalent sum for Bi, and T is the equivalent sum for Ti.

	Women	Men	Total
Occupation 1	A_1	B_1	T_1
Occupation 2	A_2	B_2	T_2
....			
Occupation n	A_n	B_n	T_n
Total	A	B	T

Despite its several advantages, reinforced by the further advantage of comparability between studies stemming from its widespread use, D has not been immune from criticism. It does not pick up all transfers (although it is capable of picking up transfers), and it needs a high ratio of minority population to the number of units of organisation. Neither of these slight weaknesses outweighs its reported strengths. More importantly perhaps is the doubt cast over its major claimed advantage of comparability over time and space by the work of Cortese et al. (1976). They claim that since it is not always organisationally invariant it is not reliable for use in comparing regions of different scale (for example).

The Dissimilarity Index, unlike many of the 'losers' in the war, has long been considered as composition invariant, for even though Duncan and Duncan (1955a) acknowledge that the proportion of both subgroups is present in the calculation they argue that D is unaffected by scaled changes in either group. Lieberson (1981) agrees that D is not affected by population composition, and gives as an example 'if the number of whites in each subarea was divided by ten, then the index of dissimilarity would remain unchanged' (p.63). By 1982, Lieberson and Carter state that D is 'affected by group size under special circumstances... but the conditions under which a problem arises are quite extreme and are unlikely to occur in real-life circumstances' (p. 296). One of the primary purposes of this paper is to argue that on a *strong* interpretation of composition invariance this is not, in fact, so (or that at least D does not meet both of the two requirements of composition invariance as described by Watts above). In this respect we agree with the critique of Cortese et al. (1976) and the later comments of Blackburn et al. (1995) that D is not entirely free of 'unwanted influence'.

Table 11 presents a hypothetical example of the number of students in four schools who are eligible for free school meals (FSM is an indicator of families defined as in poverty). D for this set of schools is 0.27. The fifth column shows what proportion of the total number of children in poverty are in each school. The sixth column shows what proportion of the total number of children in each school are in poverty. These related figures are presented to clarify the two aspects of compositional invariance described by Watts (1998) and others (see above).

Table 11 - Comparison of D and S for composition invariance I

School	FSM	non-FSM	Total	% of Total FSM	FSM as % of Total popn.	D	S
A	10	90	100	10	10	0.20	0.15
B	20	80	100	20	20	0.07	0.05
C	30	70	100	30	30	0.07	0.05
D	40	60	100	40	40	0.20	0.15
Total	100	300	400	100	25	0.27	0.20

Obviously, in the trivial case where all of the numbers in Table 11 are scaled (so that School A takes 20 FSM students from a total of 200 for example), D remains the same. Additionally, as Lieberson and others have pointed out, if the number of students eligible for free school meals is doubled in each school, perhaps reflecting a period of economic recession, then D remains the same (Table 12). This is so despite changes in the proportion of students in poverty in each school (column 6) since the proportion in each school of the total in poverty remains the same as in Table 11 (column 5). However, it should be noted that this invariance only applies if the number of students *not* eligible for free school meals is held constant (and this proviso is seldom acknowledged in verbal descriptions of the index properties). The total school population and the number of children in poverty increase by the same amount while the number of non-poor children remains the same (therefore the sixth column changes while the fifth column remains the same). This is what we have termed 'weak' composition invariance (Taylor et al. 2000a).

Table 12 - Comparison of D and S for composition invariance II

School	FSM	non-FSM	Total	% of Total FSM	FSM as % of Total ppon.	D	S
A	20	90	110	10	18	0.20	0.12
B	40	80	120	20	33	0.07	0.04
C	60	70	130	30	46	0.07	0.04
D	80	60	140	40	57	0.20	0.12
Total	200	300	500	100	40	0.27	0.16

If, instead, the number of students in poverty rises as a proportion of an existing school population but in such a way that the relative distribution of students in poverty remains unchanged between schools, then D varies. In Table 13, D increases to 0.35, which suggests that segregation has increased even though the proportion of the total students eligible for free school meals is the same for each school as it was in Tables 11 and 12. Put simply, a doubling of the figures for column 6 leads to an increase in D, yet it is far from clear that the schools in Table 13 are any more segregated (i.e. with FSM more unevenly distributed between schools) than those above. What D is picking up here is simply an increase in poverty across *all* schools.

Table 13 - Comparison of D and S for composition invariance III

School	FSM	non-FSM	Total	% of Total FSM	FSM as % of Total popn.	D	S
A	20	80	100	10	20	0.30	0.15
B	40	60	100	20	40	0.10	0.05
C	60	40	100	30	60	0.10	0.05
D	80	40	100	40	80	0.20	0.15
Total	200	200	400	100	50	0.35	0.20

These three hypothetical examples illustrate one potential misinterpretation of figures of segregation whether in school intakes, as represented here, or in ethnicity of cities or the gendered division of labour, in situations with differing composition. To be 'strongly' composition invariant an index must be unaffected by changes in the relative frequency of the groups being measured. As an example, an occupation containing 20% of the total workforce but only 10% of the women in the workforce cannot be said to be more or less segregated than it was simply because the overall number of women in the workforce changes. Segregation would change when either more or less than 20% of the workforce were in that occupation, or more or less than 10% of the women in the workforce were. The point is similar in many respects to that made about achievement gaps in Gorard (1999a). Simple scaling of the numerator should not lead to changes in either achievement gaps or measures of segregation. The principle of composition invariance states that a proportional change of size of one group which leaves its distribution over the categories unchanged leaves segregation unchanged (Kalter 2000). Yet this apparently simple rule leads to a paradox whereby either the figures in Table 12 *or* the figures in Table 13 are seen as differently segregated to those of Table 11.

Segregation index

What we have termed the 'segregation index' (S), like the other indices discussed so far, has relative advantages and disadvantages. To a large extent these are the same as those for D, for as will be seen below the two measures are related such that D is the proportion of the minority group needing to be moved to produce evenness whereas S is the proportion needing to be replaced. The calculation of S uses the difference between the proportion of a particular group in a single sub-area and the proportion of all group members in the same sub-area. Using the same terms as above:
$S = 0.5 * \Sigma (|Ai/A - Ti/C|)$.

The major objection to S over D in principle is that it is not symmetric, so that if group A is the minority (or disadvantaged) then the formula gives a different answer for the majority group B (which we have termed 'S*' throughout). However, for some commentators, such as Lieberson (1981), asymmetry is not intrinsically problematic. The key issue with asymmetry is not having the two different values for the two groups but whether they give contradictory results (Watts 1998), which for S they do not (see below). In addition most definitions of segregation (see above) are primarily concerned with the distribution of one group only. For example, the definition of unevenness used by Massey and Denton (1988) is where 'minority members may be distributed so that they are overrepresented in some areas and underrepresented in others' (p.283). The definition is asymmetric since it is only concerned with the distribution of the minority group. This is what S encapsulates.

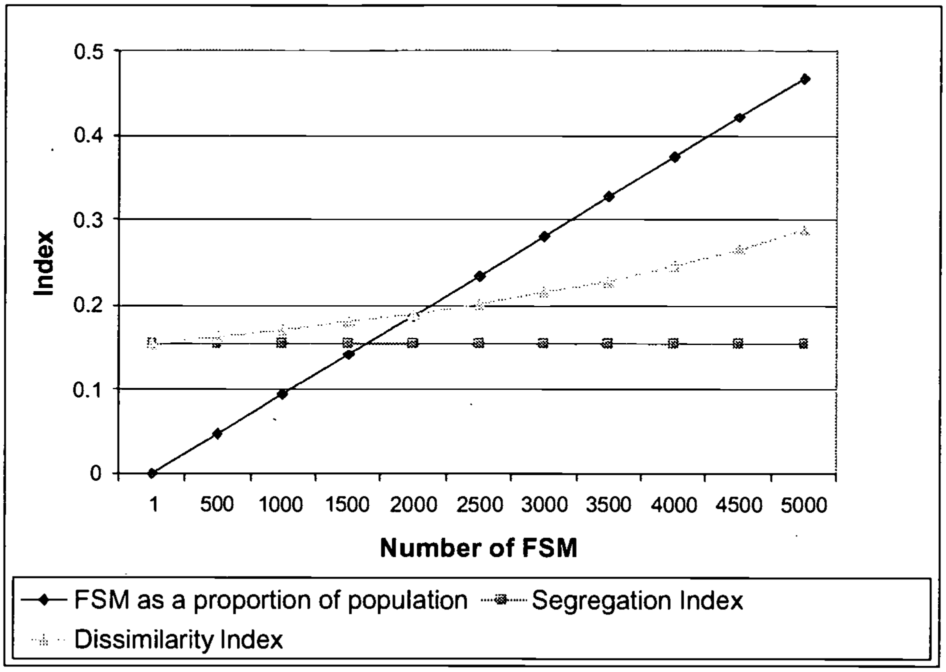
The key difference is in the base figure used to compare the distribution of any particular group. Hence, while D compares the proportion of two groups with each other by sub-area, S compares the proportion of one group with the total for that sub-area. This means that even if the proportion of students eligible for free school meals is altered, S remains unchanged as long as they are distributed to each of the schools in the same proportions as the original figures. This is illustrated in Figure 4 which shows the effects on both indices of artificially changing (or scaling) the overall proportion of students eligible for free school meals across the whole of one local education authority (Camden in 1994), while retaining the initial proportion of students eligible for free school meals in each school. As can be seen, S remains constant irrespective of changes to the absolute levels of students eligible for free school meals. However, the effects of such changes on D are clearly evident and curvilinear.

To examine this further consider the simplest case of total segregation with only two organisational units (Table 14).

Table 14 - Total segregation of FSM I

	FSM	non-FSM	Total	S	S*	D
School A	0	100	100	0.5	0.5	1
School B	100	0	100	0.5	0.5	1
Total	100	100	200	0.5	0.5	1

Figure 4 - Comparison of indices using schools in Camden



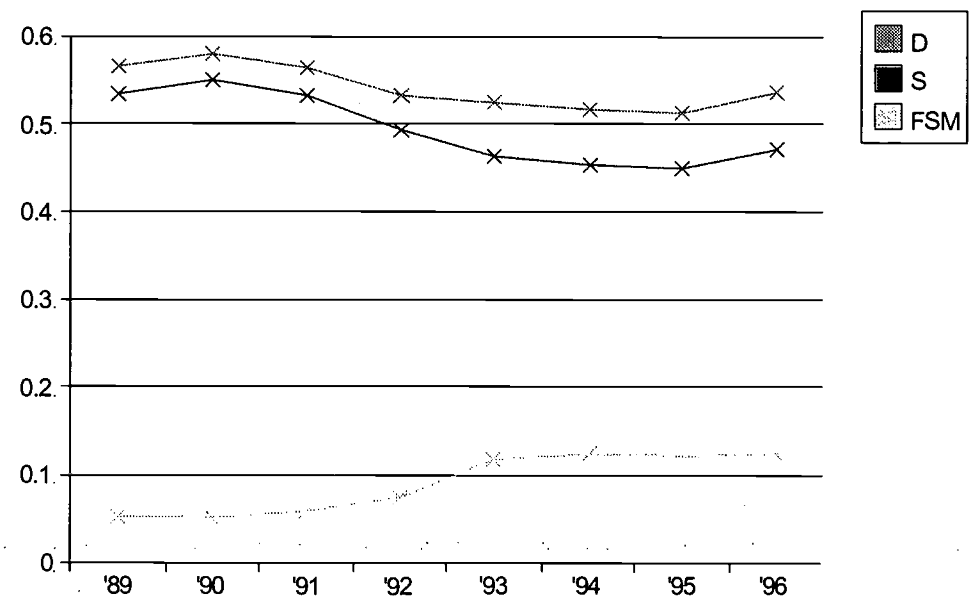
Now assume that poverty in this system has halved, but that all poor children are still in one school. According to our definition of evenness, segregation remains unchanged, but D decreases (Table 15). Therefore the dissimilarity index, like the index of isolation and others but unlike the segregation index, is measuring two different components of the composition and distribution of cases. Both S and D change as the proportion of existing FSM is altered between schools, and both also change when the overall proportion of FSM changes and is allocated differentially to schools. However, only D changes when the proportion of FSM changes otherwise. This is the key advantage of using S.

Table 15 - Total segregation of FSM II

	FSM	non-FSM	Total	S	S*	D
School A	0	100	100	0.5	0.17	0.67
School B	50	50	100	0.5	0.17	0.67
Total	50	150	200	0.5	0.17	0.67

Figure 5 shows both D and S for FSM in all secondary schools in Wales 1989-1996, as well as the percentage of FSM students. While the substantive conclusions for D and S are broadly similar, it is clear that the lines diverge as the incidence of FSM increases.

Figure 5 - Segregation in Wales 1989-1996



[note: the figures shown here include fee-paying and special schools].

We have already shown elsewhere how studies of inequality of distribution use many of the same techniques as studies of inequality of outcome (Gorard 1999a, 2000c). The calculation and discussion of achievement gaps between different sub-groups of students ('differential attainment') has become common among policy-makers, the media, and academics. An 'achievement gap' is an index of the difference in an educational indicator (such as an examination pass rate) between two groups (such as males and females). In addition to patterns of differential attainment by gender, recent concern has also been expressed over differences in examination performance by ethnicity, by social class, and by the 'best' and 'worst' performing schools. The concerns expressed in each case derive primarily from growth in these gaps over time. In particular, the Equal Opportunities

Commission calculate an 'achievement gap' to measure changes over time in patterns of differential attainment at school by boys and girls (Arnot et al. 1997). This measure, developed by the Oxford and Cambridge Examinations Board has now become standard in the field (see Gorard et al. 1999a, 2000), and is particularly useful as it can be separated into an 'entry gap' (an index of differences in number of each sex attempting an examination) and the attainment gap proper (an index of differential outcome unweighted by differential examination entry rates). Using numbers of the form expressed in Table 16, the achievement gap is defined as the number of girls reaching a certain level of attainment (Gp) minus the equivalent figure for boys (Bp) divided by the sum of boys and girls attaining that level and minus the entry gap. The entry gap is defined as the number of girls taking the examination (Tg) minus the number of boys (Tb), divided by the total examination entry.[4]

More formally:

The achievement gap = (Gp-Bp)/Tp - (Tg-Tb)/T.

Table 16 - Differential attainment by sex

	Pass	Fail	Totals
Girls	Gp	Gf	Tg
Boys	Bp	Bf	Tb
Totals	Tp	Tf	T

Since the gap uses only the first and third columns of Table 16 it is clearly more closely related to the segregation index than to the dissimilarity index, or indeed any others. In fact, the gap is simply the two organisational unit version of the segregation index, decomposed into attainment and entry components. It is a measure of evenness of pass/fail among two groups of students (girls/boys).

The segregation index = $|(Gp/Tp - Tg/T)| + |(Bp/Tp - Tb/T)|$.

Therefore if $Gp/Tp > Bp/Tp$

$S = (Gp/Tp - Tg/T) - (Tb/T - Bp/Tp)$

$S = (Gp-Bp)/Tp - (Tg-Tb)/T$

While if $Gp/Tp < Bp/Tp$ then the gap is the inverse (see note 4).

On the other hand, the dissimilarity index = $|(Gp/Tp - Gf/Tf)| + |(Bp/Tp - Bf/Tf)|$.

Therefore if $Gp/Tp > Bp/Tp$

$D = (Gp/Tp - Gf/Tf) - (Bf/Tf - Bp/Tp)$

$D = (Gp-Bp)/Tp - (Gf-Bf)/Tf$

While if $Gp/Tp < Bp/Tp$ then the gap is the inverse (see note 4).

Therefore for the two-by-two case, S gives the same result using the same algebraic formula as the achievement gap while D does not. Both as a gap between two units and an index for two or more units, S is a very valuable and widely-used measure of difference from evenness. It also appears as the Hoover coefficient of income inequality (Kluge 1998), the Hoover index of concentration (Massey and Denton 1988), as the Women and Employment (WE) Index used by the OECD (1980), as the Concentration Index (Moir and Selby Smith 1979, OECD 1985), as the WAVE or Replacement Index (Meuret 2000), and as the 'Relative Citation Index' which is the conventional bibliometric technique to compensate for differences between sizes of national academic communities (Research Fortnight 2000). The method therefore has many of the same pragmatic advantages claimed by D. Being widely used allows numerous comparisons between time, areas, and even fields of research (Lieberson and Carter 1982).

The other key advantages claimed for D are its ease of interpretation, and fact that it limits the influence of population composition (Lieberson and Carter (1982). These same advantages are true of S, which is as easy (perhaps easier) to interpret and, unlike D which now has questionable claims to composition invariance, it is 'strongly' composition invariant. In fact, S is perhaps the only measure of association usable for more than two cases which appears completely free of the influence of population changes. It is also, indirectly, advocated by Cortese et al. (1976) because it resolves an ambiguity over the meaning of D. Although reportedly easy to interpret there has been conflict over whether it represents the number of minority cases that would have to be moved or the number to be exchanged with majority members. Indeed D is often misunderstood as the proportion of one group *or* the other that would have to moved (Tzannatos 1990). 'What is often desired is the proportion of minority population which would have to be exchanged while keeping the number of households per unit constant (Cortese et al. 1976, p.633). S overcomes this ambiguity since it represents precisely that 'exchange proportion'.[5]

Inter-relationships

The chief indices described here form a kind of hierarchical family. P*-type measures, such as the Isolation Index include a term for the proportion of the focus group in the total population. This is why they are not composition invariant, and they change as the relative size of the focus group changes even if its pattern of distribution remains the same. Stripped of this term, P*-type measures are similar to evenness measures, such as the Dissimilarity Index. D is weakly composition invariant, but unlike I* falls foul of a strict interpretation of the principle of transfers. S is strongly composition invariant, but falls foul of a strict interpretation of the principle of transfers, and is not symmetric. Its inverse S* has the same properties. D is actually the sum of S and S*, which are therefore not contradictory. The rough family relationship is proposed in Figure 6.

Figure 6 - Hierarchical typology of indices

Exposure Index =		1 - Isolation Index =			
	Composition	+	Dissimilarity Index =		
			Segregation Index	+ Segregation Inverse	

The segregation index or exchange proportion can be calculated for the minority (S) or the majority population (S*), although only the first is usually of substantive interest. Each of these is strongly composition invariant and follows the standard interpretation of both organisational invariance and the principle of transfers (i.e. able to be, but not always, affected by transfers between units). The dissimilarity index (D) is the sum of S and S* and follows the standard interpretation of both organisational invariance and the principle of transfers, but is only weakly composition invariant. Kalter (2000) argues, incorrectly in our opinion, that as S = D.(B/T) and B/T is the proportion of the majority population group, then the segregation index is clearly composition variant. This argument neglects the symmetry of algebraic identities. It is also the case that D = S/(B/T), and as B/T is the proportion of the majority population group, then the dissimilarity index is clearly composition variant. As shown above, empirical and logical considerations reveal that this second interpretation is the correct one. If a component reflecting the population composition is added directly to D, I* emerges. I* is therefore clearly not even weakly composition invariant, and while it meets the standard interpretation of both other criteria it is so far from a common understanding of what segregation is and leads to such ludicrous results in real-life comparisons that it is not advised.[6]

Real-life comparisons

As already noted by many writers in this field, to a large extent it makes little difference which of the available multiplicative techniques an analyst uses. Most give very similar results under real-life conditions (as opposed to extreme cases). As an example, using the dissimilarity index (D) with data on patterns of poverty (FSM) in the schools in Swansea LEA produces the following results (Table 17). These correlate perfectly with the results for the same data using the segregation index (S). While both indices give different actual figures these are, in a sense, arbitrary. What matters here is that allowing for rounding errors the two figures are in perfect agreement about the rise and fall of segregation between schools in Swansea.

Table 17 - Comparison of Dissimilarity and Segregation index

Index	1990	1991	1992	1993	1994	1995	1996	1997
D	0.81	0.78	0.74	0.74	0.71	0.68	0.69	0.68
S	0.35	0.33	0.31	0.30	0.28	0.26	0.27	0.26

In one sense then it may not matter for many purposes which index is used, or how large the numeric results are. A more important question may be, how are the results changing over time? However, social changes are generally small and incremental, and therefore even small differences between the results using alternative methods could be important.

In the Massey and Denton (1988) study, all twenty indices were significantly correlated with each other, with a range of 0.98 to 0.99 (Pearson's r coefficient) for dissimilarity, Gini and Atkinson. Other definitions of segregation lead to indices with weaker common relationships. Factor analysis of the twenty indices suggests that while evenness and exposure are sufficiently robust and separate concepts the other three dimensions of segregation are empirically insignificant, once these first two are accounted for. Using our database of all schools in England and Wales over eleven years, we find similarly high correlations for FSM between the focus indices discussed in this paper. As expected, S and D are the most highly correlated with each other, while I (followed by D) has the strongest relationship with the population composition.

Conclusion

As noted above, the choice between the segregation index and the dissimilarity index, and others, may make little practical difference in some real-life situations. However, the differences between them are important and worthy of further investigation, particularly in terms of what we have called strong and weak composition invariance. Given that in most social science investigations of segregation the differences between places or over time can be very small, it follows that even small differences between indices can be significant. It is quite clear that any empirical consideration of segregation by area or inequality between groups, however defined, requires analytical tools such as indices to summarise the complex patterns of change over time and place. It is also clear that the choice of an index must be subsidiary to the working definition of inequality to be used in the study, and that one index alone may not be able to encapsulate that definition. For these reasons, more than the technical pros and cons of each index, debates about the use of indices are likely to continue.

The segregation index proposed here was devised in just such an empirical manner. The original form in which it was published betrays its derivation from a verbal definition of what segregation between sub-areas actually is (see Gorard and Fitz 1998a). The original proposal also included another technique, described as the segregation ratio, which combined well with the index in measuring aspects of the process of segregation which the overall index is less sensitive to (for example identifying the sub-areas in which segregation is worst). The chief recommendation for the segregation index is that it is strongly composition invariant, making it particularly appropriate for a study of changes in FSM over time since poverty has increased faster over the last ten years than school population (Taylor et al. 2000b). It is also unaffected by the change in official records in

England from figures for FSM takeup to the significantly larger figures for FSM eligibility. This change prevents other analysts from taking such a long-term view of trends, since their metric would be distorted by the sudden apparent change in population composition. The segregation index is the only index we have encountered which is thus able to *separate* the overall relative growth of FSM from changes in the distribution of FSM between schools. On balance therefore we prefer it.

Notes

1. There is an argument that inequality stems not only from an uneven distribution of population elements between organisational units but also depends on a value judgement as to which units are preferable (see Blackburn et al. 1999). For example, it could be argued that a concentration of poor children in one school is not unjust if that school is of higher quality than the others. We consider this situation still to display considerable segregation, and we attempt no such value judgements here. Therefore the paper does not consider inequality in this sense. We *are* working on an analysis of the relative effectiveness of schools over time, but there is sufficient evidence already from 'failing' schools to suggest that the schools attended disproportionately by poor children are *not* preferable to all others.

2. Little account should in any case be taken of the arguments in Noden (2000). While the paper represents a praiseworthy attempt to drive the field forward, its simple arithmetic errors make the conclusions drawn totally invalid. For example, the national levels of segregation calculated by Noden using both D and I are the 'averages' of the local levels for each area. Noden has added the indices for each LEA and divided the total by the number of LEAs regardless of the size of LEAs (and it should be recalled that the smallest LEAs have one secondary school and the largest have hundreds of schools). Therefore, if Merthyr Tydfil (4 schools) had a segregation index of 0 (no segregation at all) while Essex (380 schools) had a segregation index of 1 (total segregation), then their 'average' according to Noden would be 0.5. According to our methods of calculation, the average of these two areas would be close to 1. If after a number of years, the index for Merthyr was 1 and for Essex it was 0, Noden would conclude that segregation had not changed from 0.5, while we would conclude that it had reduced considerably from near 1 to near 0. A simple

arithmetic slip such as this (and this is not the only one) leads to substantially different but clearly bogus 'findings'. It is remarkable that the peer review system for the journal in question did not pick these errors up.

3. It is important to the argument advanced in this paper that Massey and Denton (1988) did not include S among their 20 indices. Thus, when they conclude that D is the best overall measure of evenness they mean in comparison to the Gini Coefficient and others. Hutchens (1991), on the other hand, argues that the Gini index is a preferable measure of segregation to D since it is more sensitive to some changes in occupational distribution.

4. It is, of course, possible to calculate the symmetrical inverse of this gap by subtracting the figures for girls from boys. Apart from swapping plus and minus results, no difference ensues. The gap is calculated in this way since in the UK girls are currently reaching higher levels of attainment in general than boys.

5. Despite this simplicity, many commentators in the field of education appear confused about the difference between movement and exchange, and therefore about the two indices. Errors in formulae and therefore in calculation are quite common (see Gibson and Asthana 2000, Levacic and Woods 2000, Meuret 2000, Noden 2000 as recent examples of this phenomenon).

6. Continuing this theme of seeking relationships between indices, when restricting analysis to relationships in a simple two-by-two contingency table even more close ties emerge between methods (Gorard 2000c). For example, in addition to the indices used here, the cross-product (or odds) ratio commonly used to estimate social mobility, and the segregation (or disparity) ratio (or dissimilarity index) can all be used for the same purpose. The achievement gap is used to analyse differential attainment by sub-groups, but is also useful for defining differential access to public services. Despite their differences, there are many similarities between all of the methods and their variants (Darroch 1974). At the limiting case of no relationship (no interaction, or no change over time), and also for its complete opposite, the methods are identical. No relationship is defined as equal cross-products. Given a two-by-two table of the form:

a	b
c	d

For the cross-product ratio, no change is defined as: ad/bc = 1, equivalent to $ad = bc$.

For the segregation ratio, no difference is defined as: a/(a+c) / ((a+b)/(a+b+c+d)) = 1, equivalent to a/(a+c) = (a+b)/(a+b+c+d), equivalent to $ad = bc$.

For the achievement gap, no gap is defined as: (a-b)/(a+b) - ((a+c)-(b+d)) / ((a+c)+(b+d)) = 0, equivalent to (a-b).((a+b)+(c+d)) = (a+b).((a+b)-(c+d)), equivalent to $ad = bc$.

For the percentage point method, no difference is defined as: 100a/(a+c) - 100b/(b+d) = 0, equivalent to a/(a+c) = b/(b+d), equivalent to $ad = bc$.

References

Arnot, M., David, M. and Weiner, G. (1996) *Educational Reform and Gender Equality in Schools*, Manchester: Equal Opportunities Commission

Bentley, T. (1998) *Learning beyond the classroom*, London: Routledge

Blackburn, R., Siltanen, J. and Jarman, J. (1995) The measurement of occupational gender segregation: current problems and a new approach, *Journal of the Royal Statistical Society*, 158, 2, 319-331

Blackburn, R., Jarman, J. and Brooks, B. (1999) *The relation between gender inequality and occupational segregation in 32 countries*, University of Cambridge: Cambridge Studies in Social Research, 2

Blackburn, R. and Jarman, J. (1997) Occupational gender segregation, *Social Research Update,* 16, 1-4

Boal, F. (1987) Segregation, in Pacione, M. (Ed.) Social geography: progress and prospect,

Bright, M. (1998) The trouble with boys, *The Observer*, 4/1/98, p. 13

Budge, D. (1999) Gulf separating weak and strong increases, *Times Educational Supplement,* 30/4/99, p. 3

Carvel, J. (1998a) Tory market theory 'widened schools performance gap', *The Guardian*, 11/6/98, p. 5

Clark, J., Modgil, C. and Modgil, S. (1990) *John Goldthorpe: consensus and controversy*, London: Falmer

Coleman, J., Hoffer, T. and Kilgore, S. (1982) *High School Achievement: Public, private and Catholic schools compared*, New York: Basic Books

Cortese, C., Falk, F. and Cohen, J. (1976) Further considerations on the methodological analysis of segregation indices, *American Sociological Review*, 41, 630-637

Darroch, J. (1974) Multiplicative and additive interactions in contingency tables, *Biometrika*, 61, 2, 207-214

Duncan, O. and Duncan, B. (1955a) A methodological analysis of segregation indexes, *American Sociological Review*, 20, 210-217

Duncan, O. and Duncan, B. (1955b) Residential distribution and occupational stratification, *American Journal of Sociology*, 60, 5, 493-503

Erikson, R. and Goldthorpe, J. (1991) The constant flux: a study of class mobility in industrial societies, Oxford: Clarendon Press

Everitt, B. and Smith, A. (1979) Interactions in contingency tables: a brief discussion of alternative definitions, *Psychological Medicine*, 9, 581-583

Gibson, A. and Asthana, S. (1999) *Schools, markets and equity: access to secondary education in England and Wales*, presentation at AERA Annual Conference, Montreal

Gibson, A. and Asthana, S. (2000) What's in a number?, *Research Papers in Education*, 15, 2

Gilbert, N. (1981) *Modelling society: an introduction to loglinear analysis for social researchers*, London: George Allen and Unwin

Goldthorpe, J., Llewellyn, C. and Payne, C. (1987) *Social mobility and class structure in modern Britain*, Oxford: Clarendon Press

Gorard, S. (1999a) Examining the paradox of achievement gaps, *Social Research Update*, 26

Gorard, S. (1999b) Keeping a sense of proportion: the "politician's error" in analysing school outcomes, *British Journal of Educational Studies*, 47, 3, 235-246

Gorard, S. (2000a) *Education and Social Justice*, Cardiff: University of Wales Press

Gorard, S. (2000b) Here we go again: a reply to 'What's in a number?' by Gibson and Asthana, *Research Papers in Education*, 15, 2, 155-162

Gorard, S. (2000c) One of us cannot be wrong: the paradox of achievement gaps, *British Journal of Sociology of Education*, 21, 3

Gorard, S. (2000d) Questioning the crisis account: a review of evidence for increasing polarisation in schools, *Educational Research*, 42, 3

Gorard, S. and Fitz, J. (1998a) The more things change.... the missing impact of marketisation, *British Journal of Sociology of Education*, 19, 3, 365-376

Gorard, S. and Fitz, J. (1998b) Under starters orders: The established market, the Cardiff study and the Smithfield project, *International Studies in Sociology of Education*, 8, 3, 299-314

Gorard, S. and Fitz, J. (2000a) Investigating the determinants of segregation between schools, *Research Papers in Education* , 15, 2, 115-132

Gorard, S. and Fitz, J. (2000b) Markets and stratification: A view from England and Wales, *Educational Policy* , 14, 3, 405-428

Gorard, S., Rees, G. and Salisbury, J. (2000) The differential attainment of boys and girls at school: investigating the patterns and their determinants, *British Educational Research Journal* , 26, 5

Gorard, S., Salisbury, J. and Rees, G. (1999a) Reappraising the apparent underachievement of boys at school, *Gender and Education*, 11, 4, 441-454

Gorard, S., Rees, G. and Fevre, R. (1999b) Two dimensions of time: the changing social context of lifelong learning, *Studies in the Education of Adults* , 31, 1, 35-48

Hutchens, R. (1991) Segregation curves, Lorenz curves, and inequality in the distribution of people across occupations, *Mathematical Social Sciences*, 21, 31-51

Independent (1998) Classroom rescue for Britain's lost boys, *The Independent*, 5/1/98, p. 8

James, D. and Taeuber, K. (1985) Measures of segregation, in Tuma, N. (Ed.) *Sociological Methodology*, San Fransisco: Jossey-Bass, pp.1-32

Kalter, F. (2000) *Measuring segregation and controlling for independent variables*, Mannheimer Centre for European Social Science, Working Paper 19

Kluge, G. (1998) *Wealth and people: Inequality measures, Entropy and Inequality Measures*. http:\\www.ourworld.compuserve.com (as at 21/10/99)

Le Grand, J. (1999) Conceptions of social justice, in Walker, R. (Ed.) *Ending child poverty: popular welfare for the twenty-first century?*, Bristol: Policy Press

Levacic, R. and Woods, P. (2000) *The impact of quasi-market and performance regulation on socially disadvantaged school communities*, presentation at AERA Conference, New Orleans April 2000

Lierberson, S. (1981) An asymmetrical approach to segregation, in Peach, C., Robinson, V. and Smith, S. (Eds.) *Ethnic Segregation In Cities,* Croom Helm, London

Lieberson, S. and Carter, D. (1982) Temporal changes and urban differences in residential segregation: a reconsideration, *American Journal of Sociology*, 88, 296-310

Linacre, J. (2000) *Modelling percentages with continuous Rasch models*, presentation at AERA, New Orleans, April 2000

Marshall, G., Swift, A. and Roberts, S. (1997) *Against the odds? Social class and social justice in industrial societies*, Oxford: Clarendon Press

Massey, D. and Denton, N. (1988) The dimensions of residential segregation, *Social Forces*, 67, 373-393

Massey, D., White, M. and Phua, V. (1996) The dimensions of segregation revisited, *Sociological Methods and Research*, 24, 2, 172-206

Meuret, D. (2000) *Equity issues and stratification in education*, University of Bourgogne (mimeo)

Moir, H. and Selby Smith, J. (1979) Industrial segregation in the Australian labour market, *Journal of Industrial Relations*, 21, 281-292

Noden, P. (2000) Rediscovering the impact of marketisation: dimensions of social segregation in England's secondary schools, *British Journal of Sociology of Education* (forthcoming)

OECD (1980) *Women and employment*, Paris: Organisation for Economic Cooperation and Development

OECD (1985) *The integration of women in the economy*, Paris: Organisation for Economic Cooperation and Development

Peach, C. (1975) *Urban Social Segregation*. Longman, New York

Ransom, M. (2000) Sampling distributions of segregation indexes, *Sociological Methods and Research*, 28, 4, 454-475

Research Fortnight (2000) A citation measure that fits all sizes, *Research Fortnight*, 6, 8, 18-19

Robinson, P. and Oppenheim, C. (1998) *Social Exclusion Indicators*, London: Institute of Public Policy Research

Slater, J., Dean, C. and Brown, R. (1999) Learning gulf divides rich and poor pupils, *Times Educational Supplement,* 2/4/99, p.1

Taeuber, K., Wilson, F., James, D. and Taeuber, A. (1981) A demographic perspective in school desegregation in the USA, in Peach, C., Robinson, V. and Smith, S. (Eds.) *Ethnic Segregation In Cities,* Croom Helm, London

Taylor, C., Gorard, S. and Fitz, J. (2000a) A re-examination of segregation indices in terms of compositional invariance, *Social Research Update* , 30, 1-6

Taylor, C., Gorard, S. and Fitz, J. (2000b) *Size matters: does school choice lead to 'spirals of decline'? Measuring Markets: the case of the ERA 1988 Working Paper 36*, ISBN 1872330339, 43 pages

Tzannatos, Z. (1990) Employment segregation: can we measure it and what does the measure mean? *British Journal of Industrial Relations*, 28, 105-111

Watts, M. (1998) Occupational gender segregation: Index measurement and econometric modelling, *Demography* , 35, 4, 489-496

White, P., Fitz, J. and Gorard, S. (1999a) *The new legislation on school admissions: Working Paper 32*, Cardiff: School of Social Sciences, ISBN 1 872330 29 0, 17 pages

White, P., Gorard, S. and Fitz, J. (1999b) *An analysis of local school admissions arrangements: Working Paper 35*, Cardiff: School of Social Sciences, ISBN 1 872330 26 6, 23 pages

Wright, J. (1937) Some measures of distribution, *Annals of the Association of American Geographers*, 27, 177-211

NOTICE

REPRODUCTION BASIS

☑ This document is covered by a signed "Reproduction Release (Blanket) form (on file within the ERIC system), encompassing all or classes of documents from its source organization and, therefore, does not require a "Specific Document" Release form.

☐ This document is Federally-funded, or carries its own permission to reproduce, or is otherwise in the public domain and, therefore, may be reproduced by ERIC without a signed Reproduction Release form (either "Specific Document" or "Blanket").