What *does* an index of school segregation measure? A commentary on Allen and Vignoles

Stephen Gorard

Department of Educational Studies
Research Paper 2006/04

The University of York

What *does* an index of school segregation measure? A commentary on Allen and Vignoles

Stephen Gorard

ISBN: 1 85342 632 6

Department of Educational Studies
University of York
York
YO10 5DD

I read the recent paper by Allen and Vignoles (2006) with great interest, because it is the first that I have seen to re-analyse my own results on the socio-economic segregation between all state-funded secondary schools in England from 1989 onwards (Gorard et al. 2003). Several other authors have claimed such a re-analysis. But all of these have, on reading, been found to have used different years (not re-analysing the key years from 1989 to 1992), geographical regions (not using all schools), and even completely different measures (such as achievement gaps) or making simple errors of analysis (such as finding a national figure by averaging the figures for LEAs without regard to their number of schools). See Gorard and Fitz (2006) for a summary of these problems. Perhaps the main reason why a complete re-analysis has not been attempted before is due to the change in recording by the Annual School Census after 1992. In 1989, the earliest year in which the census collected data on pupil backgrounds, the only available measure of poverty was the number of pupils taking free school meals. This measure has the disadvantage that take-up of free meals is incomplete. Only about 80% of pupils below the official poverty line actually take the free school meals. This figure varies between LEAs, and is likely to be affected by religious dietary factors. From 1993, this measure was supplemented by the more widely applicable, and so substantially larger, number of pupils eligible for free school meals. Analysts were rightly concerned that persevering with measures based on the inferior take-up figures would be criticised, but that the abrupt change in coverage and scale from 1992 to 1993 might be reflected in, and so skew, the results.

When I first conducted an analysis of the segregation between schools in terms of free school meals (FSM), I created an index of segregation with a special property (Gorard 1997). It was invariant to the change in scale from 1992 to 1993 in a way that other indices are not (Gorard and Taylor 2002). I called this the segregation index (S), but Allen and Vignoles (2006), and others, have termed it Gorard's segregation index (or GS). As a simple illustration of a property of this index, consider the artificial situation in which there are two schools, and all of the FSM pupils are in only one school (Table 1). The GS index is calculated for the set of schools, as the proportionate deviation of its FSM pupils from a completely even distribution of poverty. In Table 1, School A has 50 FSM pupils more than its fair share, and School B has 50 fewer. There is, therefore, total segregation in this system, with the deviation from even distribution (50+50 pupils) equal to the total number of FSM pupils in the system (100). For ease of understanding, GS is usually halved, so that it represents the proportion of FSM pupils who would have to exchange schools with non-FSM pupils in order to obtain a perfectly even distribution. For Table 1, GS is 0.5 or 50%, meaning that half of the FSM pupils in School A would have to exchange places with pupils in School B for there to be no segregation.[1]

Table 1 – Segregation of all FSM pupils in one school (case one)

	FSM pupils	Non-FSM pupils	Total
School A	100	100	200
School B	0	200	200
Total	100	300	400

In Table 2, everything is the same as above except that there has been an abrupt increase in the overall number of FSM pupils. In this example, School A has 75 FSM pupils more than its fair share, and School B has 75 fewer. There is, still, total segregation in this system with the deviation from even distribution (75+75 pupils) equal to the total number of FSM pupils in the system (150). The proportion of FSM pupils who would have to exchange schools with non-FSM pupils in order to obtain a perfectly even distribution (GS) is still 0.5, or 50%. Analysed from the perspective of the 'out-group', the situation in Tables 1 and 2 is the same. School places are being allocated in such a way that the poorer FSM students are being sorted and clustered into only one school.

[1] More formally, $GS = 0.5 * (\sum|F_i/F - T_i/T|)$
Where:
F_i is the number of disadvantaged children in school i
T_i is the total number of children in school i
F is the total number of disadvantaged children in the chosen area
T is the total number of children in the chosen area

Table 2 – Segregation of all FSM pupils in one school (case two)

	FSM pupils	Non-FSM pupils	Total
School A	150	50	200
School B	0	200	200
Total	150	250	400

Among the dozens of indices of segregation that I have tested, GS is unique in maintaining this 'strong compositional invariance' (Gorard and Taylor 2002), and therefore in correctly identifying the situation in both tables as being the same (i.e. total segregation). It is also special in having such an easily understandable meaning. It is the strict exchange proportion or precise deviation from evenness in any school system, using any indicator of disadvantage (and I have presented analyses elsewhere using special educational needs, English as a second language, and ethnic minority status, instead of FSM). Many indices, by squaring the deviations from evenness rather than using absolute values, lose any simple meaning for their reader (Gorard 2005).

The index championed by Allen and Vignoles (2006) is the dissimilarity index (or D), which is related to GS, but found by comparing the values in the first and second data columns of the Tables above, rather than the first and third rows.[2] In fact, D is simply the sum of GS and GS', where GS' is calculated for the majority group (non-FSM pupils in this example). For Table 1, D is 0.67. For Table 2, D is 0.8. According to D, the situation in Table 2 is more segregated than Table 1. But is this true? Well, it depends to a very great extent on what one means by 'segregation'. My purpose here is not to argue the position either way, but to demonstrate that using D over the years 1992 to 1993 with an abrupt increase in the apparent number of FSM pupils would show an illusory increase in segregation even where the actual distribution of poor pupils was unchanged, like moving from the situation in Table 1 to Table 2. For this purpose, among others, GS is to be preferred.

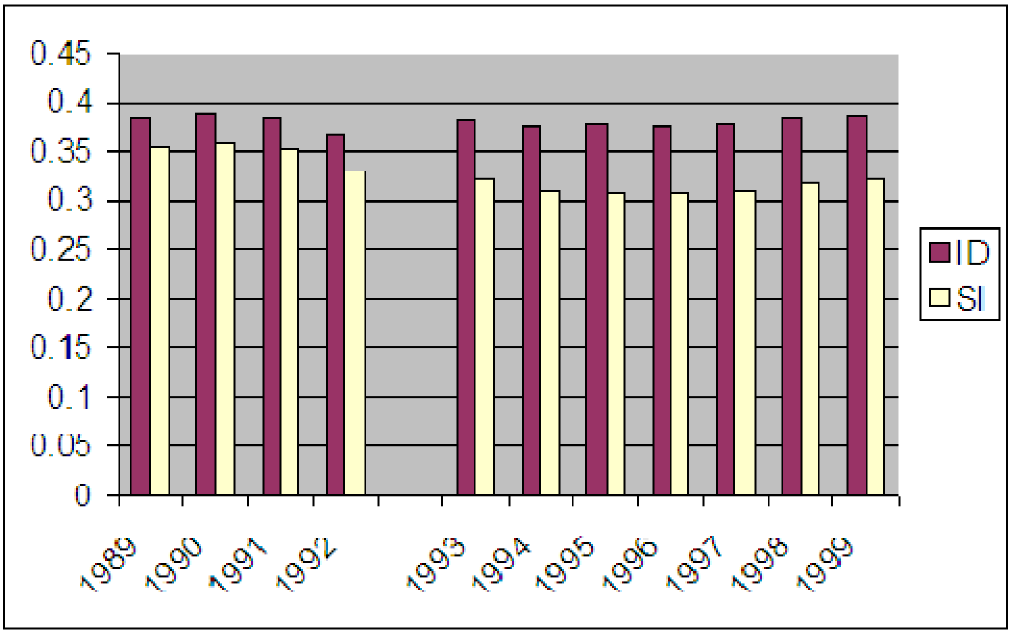
Allen and Vignoles (2006) have repeated my analysis with the same figures and the GS index of segregation. Unsurprisingly, they produce the same figures and the same pattern of change over time. Allen and Vignoles (2006) also ran the analysis with the same figures over the same time period, but using D instead of GS. They find, again unsurprisingly to me, the same substantive results. Segregation between schools did not increase after 1989, using either GS or D, despite the predictions of some theorists and the claims of some UK researchers conducting small-scale snapshot studies. In fact, using either GS or D, segregation actually declined after 1989. The claim by Allen and Vignoles (2006) that I have over-stated 'the size of the fall in segregation' (Executive Summary) cannot be upheld. The two different indices give different figures even for the same set of schools (see above), in rather the same way that Centigrade and Fahrenheit give different figures for the same underlying temperature. One cannot use the Centigrade scale to claim that a temperature drop has been over-stated because it is in Fahrenheit. The key point here is that Allen and Vignoles have shown using their preferred approach that I was right and other commentators (such as Gibson and Asthana 2000, Goldstein 2001, Noden 2002) were wrong in what Allen and Vignoles (p.2) refer to as 'a vigorous, and at times heated debate' about whether segregation rose or fell after 1989. For this, I thank them, even though they are rather shy about making their support clear in the paper (perhaps because many of the critics are linked to the same London institution, and acknowledged as helping in the production of their paper).

The reason that their results are unsurprising to me is that I had, as a matter of course, checked my results using a large number of alternative approaches (including Lorenz curves). I showed, among other things, that in real-life situations without a change in the definition of measurements, GS and D produced the same substantive conclusions. For example, in the Appendix of Gorard (2000, pp.201-202), I present an example using both GS and D. Figure 1, below, is based on the same analysis. I

[2] More formally, $D = 0.5 * (\sum|F_i/F - N_i/N|)$
Where:
F_i is the number of disadvantaged children in school i
N_i is the number of non-FSM children in school i
F is the total number of disadvantaged children in the chosen area
N is the number of non-FSM children in the chosen area

have used this example regularly in publications/presentations to make the point that the two indices are measuring the same underlying phenomenon. The two indices track each other over time, showing a rise in segregation in 1990, a decline in segregation after 1990, a plateauing of segregation from 1994 to 1997, and then a rise from 1998 that continues to this day. However, the surface difference between the two indices increases after 1992, and the figures for D must be read as two separate series. This is why, on balance, I preferred GS for summaries of the historical evidence (see above). GS and D correlate at around +0.98, or better, over repeated simulations of realistic school systems without abrupt changes in the definition of measures, suggesting that the two are otherwise generally interchangeable.[3]

Figure 1 - Comparison of dissimilarity (ID) and segregation (SI) indices – all secondary schools in England, 1989 to 1999



The remainder of the substantive results in Allen and Vignoles (2006) are also replications of my own. So, for example, my previous analyses (Gorard et al. 2003) have shown that segregation has been rising in England since 1997, and that the situation is worse in areas with high proportions of schools that do not use the same admissions arrangements as the rest of the LEA – such as voluntary-aided schools.

The first part of the paper by Allen and Vignoles (2006) discusses the relative merits of D and GS in terms of desirable properties for an index of segregation. Their thesis is clear, making a difficult and somewhat esoteric subject readable. As with the substantive results, I do not differ with most of what they say in this part of their paper, and for the same reasons. I have been over the same ground before. There is no ideal index, and all summary statistics are unable to do full justice to the raw figures from which they are calculated. The choice of an index must be as reasoned as any other part of scientific research. If the relative advantages of GS are its everyday meaning and compositional invariance, what are its disadvantages? Perhaps its biggest inconvenience is its asymmetry, giving a different

[3] More precisely, in 100 repeated trials of a school system with 100 schools, each of which had a composition consisting of a random number (between 1 and 100) of FSM pupils and a random number (between 1 and 300) of non-FSM pupils, the correlation between GS and D was 0.98 or greater. Importantly, for the purposes of this paper, the correlation between GS and GS' was 0.74 or greater – showing that these two components of D are far from contradictory. More importantly, for each school the deviations from evenness measured by D, GS, and GS' all formed a perfectly straight line when cross-plotted, and each correlated with the other at 1.00 for all trials. This shows that these three approaches to evenness are all measuring *exactly* the same thing. This forms the empirical basis for my analogy with measuring temperature using Fahrenheit and Centigrade.

value for the out-group and the in-group. However, Allen and Vignoles (2006) are quite wrong to say (Executive Summary) that GS is 'capable of showing that FSM segregation is rising and non-FSM segregation is falling simultaneously'. They are wrong because both GS and GS' are estimating the same thing, and cannot move in opposite directions. All that they would have had to do to test their statement would be to try, and fail, to provide an example where this happens. Iin fact where they discuss this point in detail (p.13) their summary claim that GS *is* contradictory is no longer made. In any school system, the proportion of the minority group who have to exchange places with the majority group in order to be evenly spread in schools is higher than the proportion of the majority group who would have to exchange to achieve the same effect. Obviously – because this is what 'minority' and 'majority' imply. This is not an error, nor a contradiction, although it is confusing for some commentators.

There are a number of other minor errors in the paper by Allen and Vignoles (2006). Their claim (footnote 6) that D represents 'the share of either group which must be removed, without replacement, to achieve zero segregation' is clearly incorrect. In Table 1 (above), for example, if no exchanges are allowed, then the only way to achieve no segregation is to remove all FSM pupils because they are all in one school. But D is 67% not the 100% that this 'meaning' requires. Similarly, one cannot achieve desegregation in Table 1 by removing only non-FSM pupils. Try it! So their attempted 'meaning' for D does not work.

Allen and Vignoles (2006) also explain that GS=D*(1-p), where p is the proportion of FSM pupils in the system. This is true, and it is also true, although they ignore it, that D=GS/(1-p). Neither of these equations is directional, and neither can establish the superiority of one index over the other, in the way that Allen and Vignoles imply (pp.10-11). It is not true that the upper limit of GS is (1-p), and again this is clear from a consideration of Table 1. In this example, GS=0.5 and (1-p)=0.75. But the segregation in Table 1 is complete, and GS cannot get any bigger without us changing the proportion of FSM pupils, and so changing (1-p). So clearly, (1-p) is not the upper bound of GS. However, the upper bound of GS does vary, in ways I have described before (only not in the way Allen and Vignoles state), and this does cause some problems for interpretation.

Whether these somewhat undesirable characteristics of GS are seen to over-ride its relative advantages will depend on the dataset and the judgement of the analyst. However, it is worth stressing that a finding so sensitive that it would be reversed simply by the use of a different, sensible, index is not really any finding at all. It behoves all researchers to seek alternative explanations for any finding, including methodological ones (Gorard 2002a, Gorard 2006), and only to convert their findings into policy or practice relevance if they are robust enough to be justified on ethical grounds (Gorard 2002b). I believe that the fall in segregation after 1989 (however its magnitude is measured), and its subsequent rise after 1997 linked to increased own-admissions schools, *are* robust findings that have stood the test of time, critique and replication even with different methods of analysis. I am glad that Allen and Vignoles agree.

References

Allen, R. and Vignoles, A. (2006) *What should an index of school segregation measure?,* London: Centre for Economics of Education, http://cee.lse.ac.uk/pubs/default.asp

Gibson, A. and Asthana, S. (2000) What's in a number? Commentary on Gorard and Fitz's 'Investigating the determinants of segregation between schools' *Research Papers in Education*, 15, 2, 133-154

Goldstein, H. (2001) Review of 'Education and social justice', *British Journal of Education Studies*, 49, 3, 354-357

Gorard, S. (1997) The missing impact of marketisation, *School Leadership and Management*, 17, 3, 437-438

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

Gorard, S. (2002a) Fostering scepticism: the importance of warranting claims, *Evaluation and Research in Education*, 16, 3, 136-149

Gorard, S. (2002b) Ethics and equity: pursuing the perspective of non-participants, *Social Research Update*, 39, 1-4

Gorard, S. (2005) Revisiting a 90-year-old debate: the advantages of the mean deviation, *The British Journal of Educational Studies*, 53, 4, 417-430

Gorard, S. (2006) Towards a judgement-based statistical analysis, *British Journal of Sociology of Education*, 27, 1, 67-80

Gorard, S. and Fitz, J. (2006) What counts as evidence in the school choice debate?, *British Educational Research Journal* (forthcoming)

Gorard, S. and Taylor, C. (2002) What is segregation? A comparison of measures in terms of strong and weak compositional invariance, *Sociology*, 36, 4, 875-895

Gorard, S., Taylor, C. and Fitz, J. (2003) *Schools, Markets and Choice Policies,* London: RoutledgeFalmer

Noden, P. (2002) Education markets and social polarization, *Research Papers in Education,* 17, 4, 410-412

ISBN: 1 85342 632 6

THE UNIVERSITY *of York*