Does the index of segregation matter? The composition of secondary schools in England since 1996

Stephen Gorard
The School of Education
The University of Birmingham
s.gorard@bham.ac.uk

Abstract

This paper presents a new analysis of segregation between schools in terms of pupils living in poverty, for all secondary schools in England from 1996 to 2005. This shows that the clustering of similar pupils in specific schools increased noticeably from 1996 to 2001, but then settled at a level still below that of 1989 when official records began. The analysis uses four estimates of segregation using figures for take-up of, and eligibility for, free school meals compiled to create both the dissimilarity index and what has been termed the Gorard index of segregation. All four estimates give the same substantive results, and the findings for the dissimilarity index and the Gorard index of segregation using either measure of FSM are indistinguishable. The two indices are, therefore, measuring the same thing. However, the Gorard index of segregation is again shown to be more tolerant of the precise measure being used, and so more strongly composition invariant than the dissimilarity index. This has important implications both for the past debate on how to measure segregation between schools, and for how education authorities go about estimating segregation in the future.

Introduction: the dangers of SES segregation

Since at least 1997, when initial analyses showed no increase in socio-economic segregation between secondary schools following the Education Reform Act 1988 (Gorard 1997), there has been what Allen and Vignoles (2006) call 'a vigorous debate that continues unabated' about the findings. This paper re-visits that debate through a new analysis of the figures for all secondary schools in England over a decade, using a variety of techniques and measures. But before moving onto how segregation is measured, and what the actual trends over time have been, the paper briefly considers why it matters who goes to school with whom.

Discussing how to measure segregation between schools, and so the trend in segregation over time and its cause, is a fascinating exercise. However, it must not be forgotten that the level of socio-economic segregation in any national school system is more than just a curiosity (Gorard and Smith 2004). It is possible to argue that the concentration of disadvantaged pupils within areas or schools is a practical advantage when seeking to administer some ameliorative package (Simpson 2004). In addition to making focussed interventions easier, the separation of more affluent and more deprived pupils between schools might reduce any resentment produced by the scale of income inequality (Gordon and Monastiriotis 2006). On the other hand, focussing disadvantage into areas or particular schools may also polarise information about

future opportunities, remove role models, and can lead to the inefficient distribution of ameliorative resources over time (Gorard 2005a).

In general, pupil achievement in integrated rather than segregated school systems is thought to depend less on the social and cultural resources of their family, delaying for as long as possible the separation of pupils by attainment, and allowing the maximum time for schools to counteract resource differences between pupil backgrounds. Such systems can, therefore, be deemed fairer in terms of a reduced association between pupil origin and outcome (Dupriez and Dumray 2006). A number of large-scale international tests and surveys have also suggested that equality between schools, far from being antithetical to high quality, improves low attainment and so can increase both the quality and equality of learning outcomes (Haahr et al. 2005). Dividing students into tracks by attainment or aptitude, or promoting diversity through the institutional differentiation of schools, on the other hand, tends to be associated with an increase in the disparity of test scores without necessarily improving the overall scores. Such segregation is also linked with a larger association between the socio-economic backgrounds of pupils and their test scores (EGREES 2005). Perhaps this is because pupils growing up in segregated settings tend to receive poorer instruction at school, less qualified teachers, substandard materials, more dilapidated plant, and experience higher crime, and generally poorer local services (Massey and Fischer 2006).

However, it is not necessary to be convinced by this evidence of a link between comprehensive intakes to schools and better, fairer, outcomes in terms of test scores (Gorard 2006). There are at least two more reasons why SES segregation between schools should be taken seriously as a national and local indicator of the (ill-) health of any school system. The first of these reasons is that the school mix could affect more than the academic outcomes assessed by formal tests. Who goes to school with whom can affect patterns of residential integration (the Belfast model – see Gorard et al. 2003), and might affect pupils' subsequent attitudes and aspirations (Burgess et al. 2005, Casey et al. 2006).

People growing up in segregated settings may be less prepared for the academic challenges of subsequent education (Gorard and Rees 2002). They may be less prepared to face diversity when they meet it, even leading to a feeling of not 'belonging', increasing their anxiety and so inhibiting performance (Massey and Fischer 2006). Segregation is strongly linked to wider social ills, such as ill-health and delinquency (Clotfelter 2001). A high level of segregation also erodes the chances of citizenship education being effective. For in the same way that schools promoting healthy eating have been found to be ineffective if they adopt a merely pedagogic stance, rather than integrating healthy eating into their processes and meals for pupils *and* staff, so citizenship needs to be adopted as a school-wide phenomenon if it is to be taken seriously. Their experience of school is probably the fundamental influence on pupils in developing their perceptions of what constitutes a fair and equitable society (as required under the National Curriculum in England, DfES 2002). The teaching programmes about general principles will be quickly seen as hypocritical, if these principles are not evident in the makeup of the school. The ethos of the school and the inter-relationships between pupils are important contexts for citizenship teaching (Davies and Evans 2002). In divided schools, citizenship education can actually generate negative results (Print and Coleman 2003).

The second non-academic reason for trying to avoid segregating pupils between schools by social class or ethnic origin is simply that it is an affront (see Massey and Denton 1998). So what has been happening to patterns of segregation in England, after a decade of increased diversity of schooling?

How should we measure segregation?

Perhaps the first problem we face when trying to analyse patterns of SES segregation between schools lies in deciding how to measure it. This problem has two components – the relevant indicator (such as poverty, ethnicity, or attainment) and the summary index used (such as Gini coefficient, isolation index, or Lorenz curve). Different choices for either component could lead to different results emerging from the same dataset, and so lead to analysts unwittingly arguing about analytical choices as though they were differences of empirical substance.[1] To some extent, this has been happening in the UK since the publication of a paper by Gorard and Fitz (1998). Gorard et al. (2003) showed that segregation measured in terms of free school meals declined substantially from 1989 to 1995. They showed that segregation measured in terms of ethnicity, first language, and special educational need also declined, once these measures were available as part of the Annual Schools Census (ASC) for England and Wales. The summary index used was the segregation index – termed the Gorard segregation index by others – although the study also showed through various publications (e.g. Gorard 2000) that exactly the same pattern emerged when using any sensible approach to capturing segregation including the dissimilarity index, and visually though Lorenz curves.

The mixture of large-scale and in-depth evidence presented by Gorard et al. (2003) suggested a number of likely determinants of local and national levels of school segregation, including patterns of residential segregation, contrasting approaches to allocating schools places (such as banding or catchment areas), and the proportion of schools not sharing their local authority approach to allocating places (such as faith-based or selective schools). The study also showed that segregation began to rise again from 1997 to 2001, and that this rise was linked both temporally and geographically with growing diversity of schooling, especially in the proportion of schools not sharing their local authority approach to allocating places. Much of this work has been confirmed through direct replication by Allen and Vignoles (2006) leading, unsurprisingly, to exactly the same results when using the same measures and indices. Allen and Vignoles (2006) state in their introduction that '…the findings of Gorard et al. do hold regardless of measure used…'.[ii] My original findings have also been indirectly confirmed by other analysts, including (Johnston et al. 2006), and Croxford and Paterson (2006) using different datasets (sample surveys rather than a school census) who say 'the second contribution we have made to this debate concerns our substantive conclusions. These do, on the whole, tend to lend empirical support to the conclusions which Gorard and Fitz drew from their data (p.401).

These replications and confirmations are important because several other earlier analysts and commentators had given the mistaken impression that they had performed similar analyses but come to different substantive conclusions. This, and perhaps an ideological desire among some commentators to see segregation rise after

the Education Reform Act 1988 rather than celebrate its fall, led to the 'vigorous debate' mentioned by Allen and Vignoles (2006). All of these other analyses have, on reading, been found to have used different years (not re-analysing the key years from 1989 to 1992), different 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). Allen and Vignoles (2006) are the first analysts to do exactly what I did and, as they somewhat grudgingly admit, they found exactly what I found.

Croxford and Paterson (2006), for example, point out that Gibson and Asthana (2000), who claimed to dispute the Gorard and Fitz (2000) paper, had not only not measured SES segregation at all but had confused segregation as a process with the measured outcomes of segregation in schools. Harrison (2004) and Hames (2003) both show how the Smithfield study in New Zealand, often cited in opposition to Gorard and Fitz (2000), was riddled with shortcomings, had major flaws in its empirical work, and proposed recommendations that did not follow from its analysis. Hames (2003) said 'The Harrison critique [a report to the NZ Ministry of Education] constitutes a fairly complete demolition of any intellectual pretensions the Smithfield reports might have… it was astonishing the government had spent good money on it' (p.9-10). Reardon and Firebaugh (2002b) warn of the danger of arbitrarily matching a convenient or fashionable procedure with the verbal concept of segregation, and so ending up using an index which behaves very strangely when the relative share of groups changes. This is what Noden (2000) did, in another analysis that is frequently cited as being in opposition to that of Gorard and Fitz (2000). In fact, there were so many errors in the analysis by Noden (2000) that he took the unusual step of retracting it (Noden 2002).

See Gorard and Fitz (2006) for fuller discussion of the problems in these and other analyses of segregation. The serious point remains that every analyst needs to decide on and justify the indicator used to measure segregation, and on the index used to summarise it.

The battle of the indices

Measuring between-school segregation in England requires at least two important decisions – about the indicator and the index used. Free school meals (FSM) have been frequently used as an indicator of disadvantage for a number of reasons. The data are officially collected from all maintained schools, have been collected continuously since 1989, are legally required, and have a very simple binary legal definition. Eligibility for FSM represents all pupils that are known to be living below the official poverty line, from any family entitled to Income Support (and more recently Income-based Jobseeker's Allowance or support under Part 6 of the UK Immigration and Asylum Act 1999). Take-up of FSM represents all pupils actually in receipt of free meals. Compared to eligibility, take-up is perhaps more likely to be biased by religious or other dietary requirements. But take-up is a relatively sure count based on the taking of meals, and it may also represent more nearly those families in particular need of this help, who are less likely to allow religious and other dietary requirements to prevent them from taking up their entitlement.

The obvious limitations of using FSM are that they only tell us something about the economic circumstances of around 20% of the school population, and the schools may not know about eligibility if the families concerned do not tell them. However, when concerned with social justice, it is usually the distribution of the most disadvantaged that is of most concern. The potential segregation between schools for the rich and the super rich, for example, appears less pressing as a social issue than that between the poorest 20% and everyone else. Also, it not at all clear that there is any more complete dataset than the ASC. Cohort studies and sample surveys of the kind used by Croxford and Paterson (2006), among others, have much higher non-completion than the ASC, yet represent only a fraction of the relevant school population. The potential for sample bias is much higher. And variables such as occupational or social class might appear attractive in covering all cases, but they are based on arbitrary and disputed concepts, impose many threshold effects on an underlying pattern of continuous distribution, require considerable subjective judgement to code, and lead to high non-response (Gorard and Smith 2006). FSM, in contrast, has a relatively simple legal definition yielding a binary variable of high reliability.

A third practical problem when using FSM is that the early figures from 1989 to 1992 were based solely on take-up. From 1993, this measure was supplemented by the more widely applicable, and so substantially larger, number of pupils who were eligible for free school meals whether they took the meals or not. So the analytical problem is that if one only uses FSM eligibility, then the analysis cannot go back before 1993, thus missing the crucial years immediately after the 1988 Education Reform Act. If one uses FSM take-up instead then one can go back to 1989, but only with an inferior measure. A good compromise might be to use take-up until 1992, and eligibility from 1993. But this means that there will be an abrupt change in coverage and scale from using take-up of FSM in 1992 to eligibility in 1993, which might be reflected in the computed results. One solution is to use an index with both of these indicators which is strongly composition invariant (Gorard and Taylor 2002), and so unaffected by the apparent change in scale from 1992 to 1993. Gorard (1997) proposed a segregation index which has this property of invariance (see below for the formula), and which represents the proportion of FSM pupils who would have to exchange schools with non-FSM pupils for there to be a totally even spread of FSM pupils. Using other indices, such as the dissimilarity index, 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

Although Allen and Vignoles (2006) and others have termed the index proposed in Gorard (1997) 'the Gorard segregation index' it actually has a pedigree from long before that in a number of guises (Gorard and Taylor 2002). For example, Duncan et al. (1961) introduce an index which they call 'delta', which is in turn adapted from the Hoover coefficient used in populations studies (Hoover 1941). All three of these are effectively the same index, found by comparing the distribution of the out-group with the distribution of the population. This distinguishes it from indices, such as the dissimilarity index, found by comparing the distribution of the out-group with the distribution of the in-group. It is also distinguished from all of those indices, such as the Gini coefficient, that use squaring of numbers rather than absolute values to eliminate the negative residuals. Squaring the residuals before aggregation, even when

the result is then square-rooted, distorts the index by emphasising larger deviations, making it harder to calculate and understand, and losing its everyday meaning (Gorard 2005b).

There is now considerable agreement about the desirable properties of any index of segregation, and also agreement that no index is perfect. To use any one index in preference to another requires a kind of cost-benefit analysis of its strengths and weaknesses in relation to the analytical problem at hand. Different fields of public policy might express these characteristics in different ways, but the traditional four desirable attributes for an index are that it is:

- Organisationally invariant, such that if a school is broken into two, or if two schools merge, with the same proportion of FSM pupils in all, then the value of the index remains the same
- Size or scale invariant, such that if the number of both FSM and non-FSM pupils is multiplied by a constant in all schools, then the value of the index remains the same
- Compositionally invariant, such that if the number of FSM pupils is multiplied by a constant in all schools, then the value of the index remains the same (equivalent to the margin-free criterion in sex segregation analysis)
- Affected by transfers, such that if an FSM pupil moves from a school with more FSM pupils to a school with less, then the value of the index goes down

Reardon and Firebaugh (2002a) add another, or perhaps clarify an alternative to the fourth one above:

- Affected by exchanges, such that if an FSM pupil from a school with more FSM pupils exchanges places with a non-FSM pupil in a school with less FSM, then the value of the index goes down

In addition, there are desirable aesthetic characteristics of different indices. It is better, on the whole, that they are easy to calculate, do not distort deviations through squaring to produce absolute values, are easy to comprehend with a real-world meaning, symmetrical between FSM and non-FSM pupils, and in a clearly defined range such as 0 for no segregation and 1 for total segregation (Gorard and Taylor 2002, Reardon and Firebaugh 2002b, Hutchens 2004, Gorard 2005b).

The two indices used in this paper (see below) meet these desirable characteristics to about the same extent. Both the dissimilarity index (D) and the Gorard segregation index (GS) are organisationally and scale invariant, affected by transfers and exchanges, easy to calculate, do not distort deviations, and lie within a clearly defined maximum range. GS also has the specific advantages of being strongly composition invariant, and with an easy-to-understand everyday meaning. D has the specific advantage of being completely symmetrical, giving the same result for both FSM and non-FSM pupils. Which of these specific characteristics is deemed more important is not an absolute decision, but one dependent upon the analytical problem faced.

Methods

The new analysis presented here is based on figures from the Annual Schools Census (ASC) for all maintained secondary and middle-deemed-secondary schools, in England from 1996 to 2005. There were 3,610 such schools in 1996, and 3,415 in 2005. The ASC provided the number of full-time equivalent pupils in each school, the number taking free school meals, and the number known to be eligible for free school meals. In a few schools, the number of free school meal (FSM) pupils was so small that the DfES (now DCSF) considered it a danger that individuals might be identifiable. The numbers were recorded as '#', and for the purposes of this analysis they have all been treated as equivalent to two pupils.

The three figures for each school were used to calculate the Dissimilarity Index (D) and also the Gorard Segregation Index (GS) for both FSM take-up and FSM eligibility. Thus, there are four estimates of national segregation by poverty between schools, and four residuals from perfect even distribution of poverty for each school. These four estimates are compared graphically, and in terms of Pearson correlations, and used to draw robust substantive conclusions about trends over time.

The residual for the Dissimilarity Index (or D) is the absolute value of the result of subtracting the population proportion of non-FSM pupils in each school from the population proportion of FSM pupils in each school. D itself is the sum of these residuals for all schools, then divided by two. More formally, $D = 0.5 * (\sum|F_i/F - N_i/N|)$
Where:
F_i is the number of FSM children in school i
N_i is the number of non-FSM children in school i
F is the total number of FSM children in England
N is the number of non-FSM children in England

The residual for the Gorard Segregation Index (GS), on the other hand, is the absolute value of the result of subtracting the population proportion of all pupils in each school from the population proportion of FSM pupils in each school. GS itself is the sum of these residuals for all schools, then divided by two. More formally, $GS = 0.5 * (\sum|F_i/F - T_i/T|)$
Where:
F_i is the number of FSM children in school i
T_i is the total number of children in school i
F is the total number of FSM children in England
T is the total number of children in England.

The measure of disadvantage used in these formulae can be take-up of, or eligibility for, FSM, as here, or any other figures from the ASC such as pupils with special needs, English as an additional language, or from minority ethnic groups. The local unit used here is each school, and the larger area is England, although the indices can be calculated for any area containing two or more schools. The two indices and the two measures of FSM yield four estimates of segregation, abbreviated below as GSt (take-up), GSe (eligibility), Dt and De.

Comparison of D and GS

When the four residuals – for GSt, GSe, Dt, De - are created for the 3,500 or so schools they are very highly correlated. In fact, the residuals for either indicator of disadvantage yield a perfect correlation (1.00) between D and GS.[iii] The residuals are slightly different in scale. For example, Figure 1 shows a crossplot of the FSM take-up residuals for all secondary schools in 2004. The y-axis shows the D residual for each school, and the x-axis shows the GS residual. In each case, the D residual is higher than the GS residual, but the cross-plot is a perfect straight line. This means that the two scores are both measuring exactly the same thing but are expressed in a different scale, in the same way as measuring distances in metres and yards. Any distance measured accurately in yards will yield a higher figure than one measured in metres, but the two measures are precisely equivalent and measuring the same thing. This is what Figure 1 illustrates is also true for D and GS.[iv] Figure 2 shows the equivalent figures for FSM eligibility summarised in both D and GS. De and GSe are measuring the same thing even though expressed in slightly different figures. This much is, anyway, discernible from their very similar formulae (see above). I stress this point because other commentators have suggested that D is a good index for measuring segregation between schools whereas GS is a poorer one (e.g. Allen and Vignoles 2006). These commentators are simply mistaken. D and GS are variants of each other with almost all of the same desirable, and some undesirable, characteristics (see above, and Gorard 2007).

Figure 1 – Crossplot of Dt and GSt residuals, all secondary schools in England, 2004
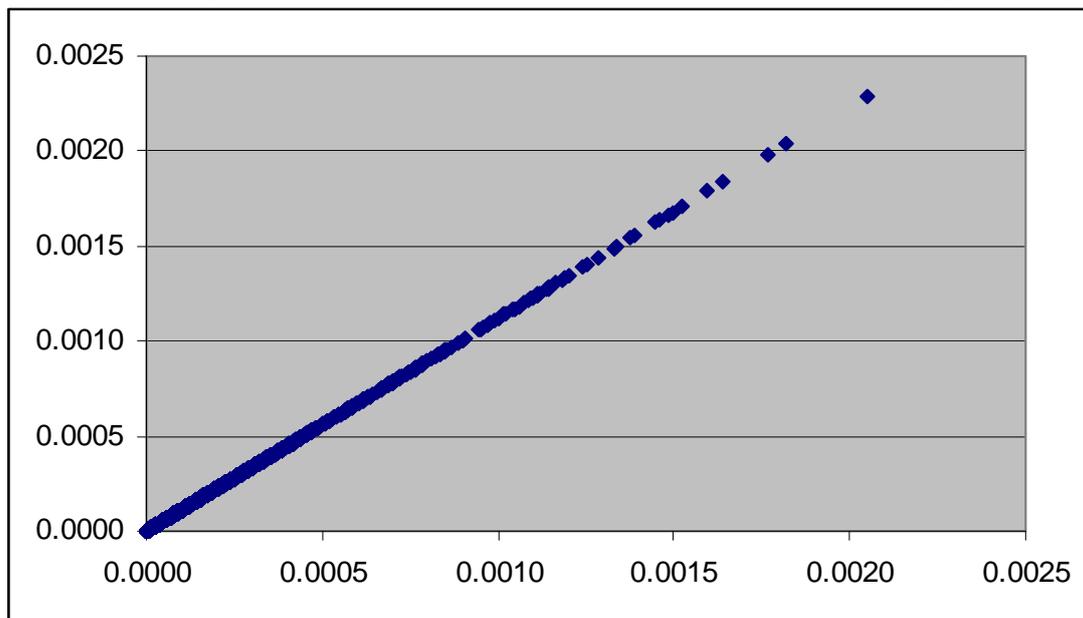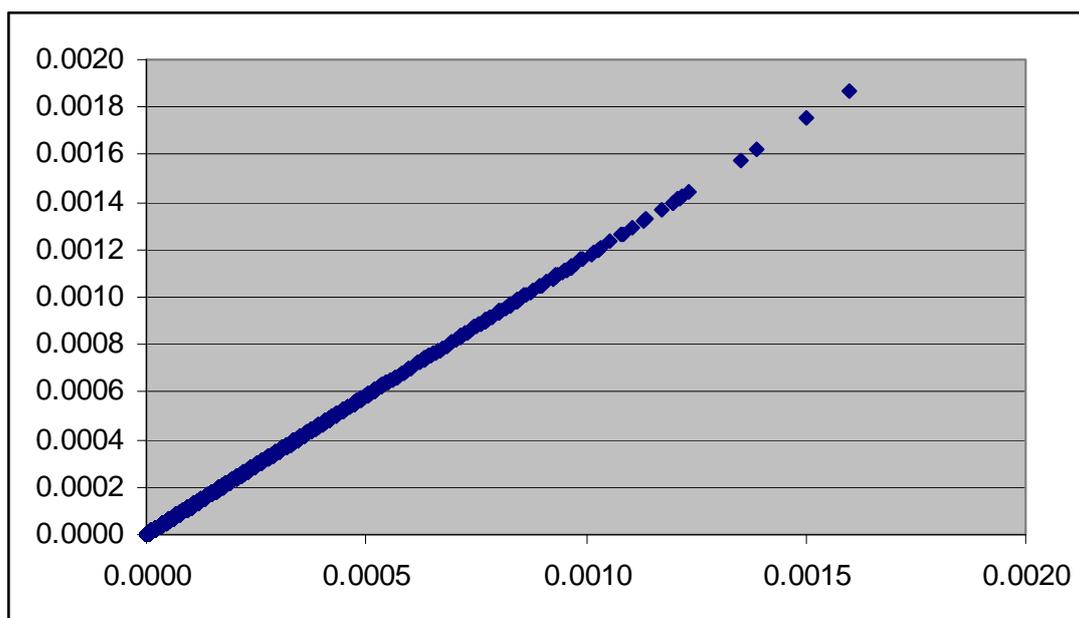

Figure 2 – Crossplot of De and GSe residuals, all secondary schools in England, 2004

When either set of residuals, such as those illustrated in one of the axes in Figures 1 and 2, are added together and divided by two to obtain one of four versions of the indices, all resultant indices also correlate very highly. Across the 10 years in this analysis, Dt and GSt correlate at 1.00, and De and GSe correlate at 0.98. Table 1 presents the correlations, over 10 years, of all four estimates of segregation between school.[v] If we use take-up as our preferred measure of disadvantage, perhaps because it reveals those who are most in need, then it makes no difference at all whether we assess segregation in this era (1996-2005) using D or GS. We have a free choice of index, and the results will be exactly the same. If we prefer using eligibility as our measure of disadvantage, because it is unaffected by the dietary requirements perhaps, then it still makes very little difference in practice whether we assess segregation in this era (1996-2005) using D or GS. We still have a relatively free choice, and the substantive results will be the same – measurement and recording error in the ASC will far outweigh any minor difference obtained using D or GS which correlate at 0.98.

Table 1 – Pearson correlation between four estimates of segregation, all secondary schools in England, 1996-2005

	GS take-up	GS eligibility	D take-up	D eligibility
GS take-up	-	0.97	1.00	0.92
GS eligibility	0.97	-	0.96	0.98
D take-up	1.00	0.96	-	0.89
D eligibility	0.92	0.98	0.89	-

Note: presented to only two decimal places for ease of reading.

If we now turn this analysis around, we can see a further illustration of the reasons advanced in Gorard (2007) why GS might be preferred to D. If we use GS, it does not matter much whether we measure poverty in terms of take-up or eligibility. The results will correlate at 0.97, so that, allowing for the more sizable errors likely in the initial figures, we get the same substantive finding either way. GS is very tolerant of the precise measure of disadvantage. On the other hand, if we select D as our index, our choice of a measure of disadvantage matters more. The results for take-up and

eligibility will correlate at only 0.89, so D is less tolerant than GS of the precise measure of disadvantage. This advantage of GS has been termed 'strict compositional invariance' (Gorard and Taylor 2002), and it is one of the main reasons why GS was used in the first place (Gorard 1997).[vi] Until 1992, the ASC only collected FSM take-up figures, whereas FSM eligibility figures were collected from 1993. Thus, any analysis that includes both of these years either has to use take-up throughout or else use an index that is not much affected by the changeover from take-up to eligibility (i.e. GS).

The trend in school segregation

It has long been established (Gorard and Fitz 1998, 2000, 2006), and now independently confirmed several times (see above), that segregation between secondary schools in terms of FSM pupils fell from the year before the Education Reform Act 1988 was enacted (1989) to a settled lower level in the mid-1990s before rising again from 1997 onwards. What has happened since then? Table 1 shows the four estimates of segregation over the ten years from 1996. There has been a steady growth in between-school segregation by poverty until 2001/2002, whichever estimate is used. In fact the proportion of FSM pupils who would have to exchange schools to achieve an even spread has now risen above 34%, only a little below the estimate for 1989 (near 36%) when segregation began to fall. Since 2002, however, segregation has settled again, just as it did in the mid-1990s. Whatever it was that drove the increase in segregation 1996-2001 appears to have abated. Future papers will address the likely determinants of this increase and the subsequent stasis. The purpose of this paper is to show the overall pattern with the same substantive result appearing whichever of the four estimates of segregation is used, and to illustrate, again, that GS is marginally preferable when analysing the changing *distribution* of an indicator such as FSM which is also liable to change in *prevalence* over time.

Table 2 - Comparison of GS and D for both FSM take-up and eligibility, 1996-2005, all secondary schools in England

Year	GS take-up	GS eligibility	D take-up	D eligibility
1996	.30	.31	.35	.38
1997	.30	.31	.34	.37
1998	.31	.32	.35	.38
1999	.32	.32	.36	.39
2000	.33	.33	.37	.39
2001	.33	.33	.37	.39
2002	.34	.33	.39	.39
2003	.34	.34	.38	.39
2004	.34	.34	.38	.39
2005	.34	.34	.38	.39

Note: presented to only two decimal places for ease of reading.

Figure 3 shows the same results as Table 2, but in graphical form. The figure emphasises how D and GS track each other (like distances in metres and yards), whether it is in terms of eligibility (the top and bottom lines in 2005), or take-up (the middle two lines in 2005). Figure 3 also illustrates that D is always higher than GS for the same years, but that GS for take-up and GS for eligibility (the bottom two lines)

remain very close to one another in a way that D for take-up and eligibility (the top two lines) does not. This is a graphical illustration of the characteristic of strong compositional invariance.

Figure 3 - Comparison of GS and D for both FSM take-up and eligibility, 1996-2005, all secondary schools in England
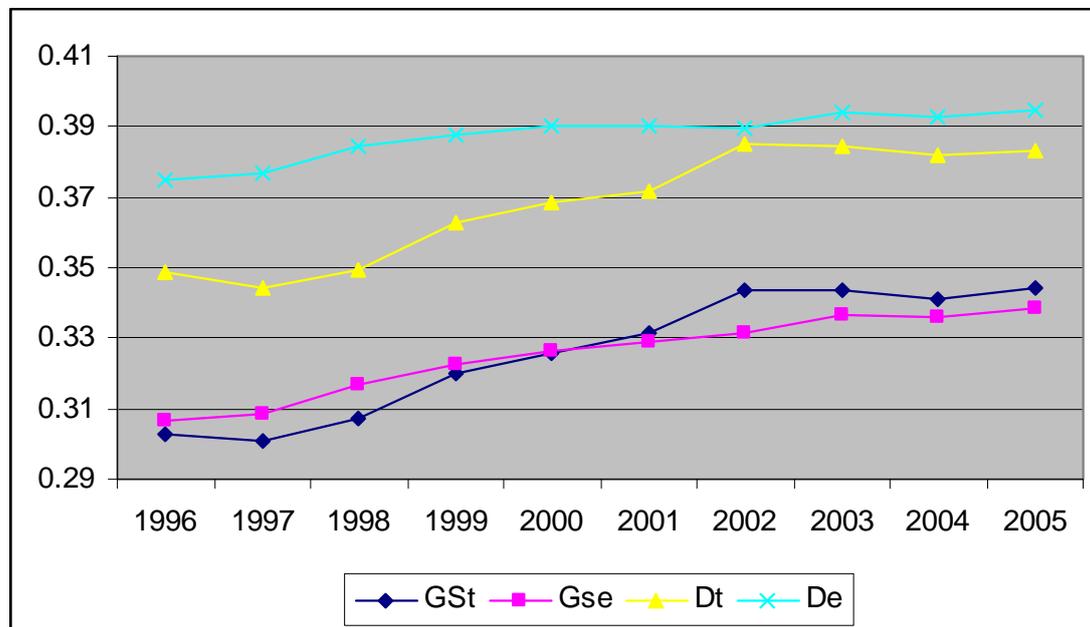
Conclusion

It hoped that the findings presented in this paper will now help to end the debate about how to measure segregation between schools in England. Apart from during the changeover from measuring FSM by take-up in 1992 to measuring by eligibility from 1993, it makes no substantive difference whether GS or D is used. The clear and confirmed pattern is that school segregation by poverty in England decreased from 1990 onwards, after the Education Reform Act 1988, before rising from 1996 until 2001/02. Of course, the precise figure for segregation depends upon the measure and index used. Using GS, the proportion of FSM pupils who would have to exchange schools to ensure no clustering of poor children in particular schools was 36% in 1989, dropping to around 30% in 1995. Segregation then increased from 1996 onwards to around 34% in 2001, following the School Standards and Framework Act 1998. Segregation remained at that level until the most recent figures presented here for 2005/06. In future papers, it will be necessary to consider in more detail why this pattern of recent change occurred.

In the period covered by this paper, 1996 to 2005, there was no change in the measurement of FSM and so no large annual jump in the proportion of FSM pupils. Under these conditions, it makes no substantive difference whether D or GS is used to summarise patterns of segregation. However, even here GS shows its superiority in the sense that it is empirically more tolerant of the precise measure being used. The correlation over time for segregation assessed via GS and using take-up or eligibility for FSM is considerably higher than the equivalent correlation for D. These

11

methodological findings about how to estimate segregation between schools, and compare the results between areas with differing proportions of FSM pupils may be more important now that local authorities are required to monitor their situation in this respect. Without a fuller awareness of the nature of compositional invariance in an index there is a danger that public policy will be misled into reading differences of prevalence in FSM as though they were differences in distribution. Poorer areas are likely to suffer more as a consequence.

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

Burgess, S., Wilson, D. and Lupton, R. (2005) Parallel lives? Ethnic segregation in schools and neighbourhoods, *Urban Studies*, 42, 7, 1027-1056

Casey, L., Davies, P., Kalambouka, A., Nelson,N. and Boyle, B. (2006) The influence of schooling on the aspirations of young people with special educational needs, *British Educational Research Journal*, 32, 2, 273-290

Clotfelter, C. (2001) Are whites still fleeing? Racial patterns and enrolment sifts in urban public schools, *Journal of Policy Analysis and Management*, 20, 2, 199-221

Croxford, L. and Paterson, L. (2006) Trends in social class segregation between schools in England, Wales and Scotland since 1984, *Research Papers in Education*, 21, 4, 381-406

Davies, I. and Evans, M. (2002) Encouraging active citizenship, *Educational Review*, 54, 1, 69-78

DfES (2002) *Citizenship: the National Curriculum for England*, http://www.dfes.gov.uk/citizenship [Accessed August 2003].

Duncan, O., Cuzzort, R. and Duncan, B. (1961) *Statistical geography: problems in analyzing area data*, Glencoe IL: Free Press

Dupriez, V. and Dumray, X. (2006) Inequalities in school systems: effect of school structure or of society structure?, *Comparative Education*, 42. 2. 243-260

European Group for Research on Equity in Educational Systems, (2005) Equity in European Educational Systems: a set of indicators, *European Educational Research Journal*, 4, 2, 1-151

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

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. (2005a) Academies as the 'future of schooling': is this an evidence-based policy?, *Journal of Education Policy*, 20, 3, 369-377

Gorard, S. (2005b) 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) Is there a school mix effect?, *Educational Review*, 58, 1, 87-94

Gorard, S. (2007) What *does* an index of school segregation measure? A commentary on Allen and Vignoles, *Oxford Review of Education*, 33, 5

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

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

Gorard, S. and Fitz, J. (2006) What counts as evidence in the school choice debate?, *British Educational Research Journal*, 32, 6, 797-816

Gorard, S. and Rees, G. (2002) *Creating a learning society?*, Bristol: Policy Press

Gorard, S. and Smith, E. (2004) An international comparison of equity in education systems?, *Comparative Education*, 40, 1, 16-28

Gorard, S. and Smith, E. (2006) Beyond the 'learning society': what have we learnt from widening participation research?, *International Journal of Lifelong Education*, 25, 6, 575-594

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

Gordon, I. and Monastiriotis, V. (2006) Urban size, spatial segregation and inequality in educational outcomes, *Urban Studies*, 43, 1, 213-236

Haahr, J., with Nielsen, T, Hansen, E. and Jakobsen, S. (2005) *Explaining student performance: evidence from the international PISA, TIMSS and PIRLS surveys*, Danish Technological Institute, www.danishtechnology.dk

Hames, M. (2003) *The logic of school choice – a dialogue, Education Forum Briefing Papers*, No. 8, April 2003, www.educationforum.org.nz

Harrison, M. (2004) *Education matters: Government, markets and New Zealand Schools*, Wellington: The Education Forum

Hoover, E. (1941) Interstate redistribution of population 1850-1940, *Journal of Economic History*, 1, 199-205

Hutchens, R. (2004) One measure of segregation, *International Economic Review*, 45, 2, 555-578

Johnston, R., Burgess, S., Harris, R. and Wilson, D. (2006) *'Sleep-walking towards segregation'? The changing ethnic composition of English schools, 1997-2003*, CMPO Working Paper Series 06/155

Massey, D. and Denton, N. (1998) *American Apartheid: Segregation and the making of the underclass*, Harvard: Harvard University Press

Massey, D. and Fischer, M. (2006) The effect of childhood segregation on minority academic performance at selective colleges, *Ethnic and Racial Studies*, 29, 1, 1-26

Noden, P. (2000) Rediscovering the impact of marketisation: dimensions of social segregation in England's secondary schools, 1994-1999, *British Journal of Sociology of Education*, 21, 3, 371-390

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

Print, M. and Coleman, D. (2003) Towards understanding of social capital and citizenship education, *Cambridge Journal of Education,* 33, 1, 123-149

Reardon, S. and Firebaugh, G. (2002a) Measures of multigroup segregation, *Sociological Methodology*, 32, 1, 33-67

Reardon, S. and Firebaugh, G. (2002b) Response: segregation and social distance, *Sociological Methodology*, 32, 1, 85-101

Simpson, L. (2004) Statistics of racial segregation: measures, evidence and policy, *Urban Studies*, 41, 3, 661-681

[i] The Gorard (1997) paper and what followed was something of a breakthrough, both in terms of the methods and data used to examine the impact of school choice and in the results that ran contrary to almost all of the UK work that had gone before it – as any citation search will attest. That is why direct confirmation such as that by Allen and Vignoles (2006) is important. What is peculiar is that these authors, and others, still wish to argue about the relative merits of different indices more than they are willing to state clearly (in their abstract, for example) the significance of their direct replication of my work from a decade earlier. Whichever decent approach is used leads to the same substantive findings as mine. That is crucial in what Allen and Vignoles, and others, have now done. The rest is interesting, but empirically and substantively much less relevant.

[ii] But they also state that the true level of segregation is lower than I have stated. This is because they use a different metric, and so the numbers they generate are smaller (although highly correlated with mine). This unfounded comment by Allen and Vignoles is almost exactly like a claim that someone has overstated a length because they measure it in feet rather than yards. For more on this, see Gorard (2007).

[iii] Although some readers might think that this perfect correlation between the residuals of D and GS would be obvious from the similarity in their calculations, it is worth stressing here because there are some commentators to whom this identity is not obvious, but who are nevertheless taken seriously by others. Allen and Vignoles (2006), for example, try to portray the two indices as fundamentally different. They are not. And, as this paper shows, unless the distribution of the underlying measure changes fundamentally (as it did in 1992) then they give exactly the same substantive results – as should all decent indices of segregation.

[iv] Some commentators, again including Allen and Vignoles (2006), might object that using these real-life data on schools is somehow unfair and that at extremes the values of D and GS would diverge in some way. We have no reason from the formulae, from simulations, or from real-life to expect this perfect correlation to go wrong for any given set of figures. Again, for further discussion of the errors made by Allen and Vignoles (2006), such as those surrounding the boundaries for D and GS, see Gorard (2007).

[v] The use of two decimal places in all tables for this paper must not mislead readers of a less numerate disposition into imagining that the variation of less than 5/1000ths between some measures in any one year in Table 2 mean that Dt and GSt, for example, cannot correlate at a value within less then 5/1000ths of 1.00 in Table 1.

[vi] See Figure 3 for why this difference might matter in an analysis changing indicators between takeup and eligibility. Of course, an analyst might not want such invariance if their object of study was focused on the differences in trends between segregation by takeup and by eligibility. As above, each analysis needs to justify its use of an index. There is no one perfect index for all situations.