This article was downloaded by: [5.67.153.220]
On: 12 June 2014, At: 12:08
Publisher: Routledge
Informa Ltd Registered in England and Wales Registered Number: 1072954 Registered office: Mortimer House, 37-41 Mortimer Street, London W1T 3JH, UK

International Journal of Research & Method in Education
Publication details, including instructions for authors and subscription information:
http://www.tandfonline.com/loi/cwse20

Measuring equity in educational effectiveness research: the properties and possibilities of quantitative indicators

Anthony Kelly[a]
[a] School of Education, University of Southampton, Highfield, Southampton SO17 1BJ, UK
Published online: 19 May 2014.

To cite this article: Anthony Kelly (2014): Measuring equity in educational effectiveness research: the properties and possibilities of quantitative indicators, International Journal of Research & Method in Education, DOI: 10.1080/1743727X.2014.914486

To link to this article: http://dx.doi.org/10.1080/1743727X.2014.914486

PLEASE SCROLL DOWN FOR ARTICLE

Taylor & Francis makes every effort to ensure the accuracy of all the information (the "Content") contained in the publications on our platform. However, Taylor & Francis, our agents, and our licensors make no representations or warranties whatsoever as to the accuracy, completeness, or suitability for any purpose of the Content. Any opinions and views expressed in this publication are the opinions and views of the authors, and are not the views of or endorsed by Taylor & Francis. The accuracy of the Content should not be relied upon and should be independently verified with primary sources of information. Taylor and Francis shall not be liable for any losses, actions, claims, proceedings, demands, costs, expenses, damages, and other liabilities whatsoever or howsoever caused arising directly or indirectly in connection with, in relation to or arising out of the use of the Content.

This article may be used for research, teaching, and private study purposes. Any substantial or systematic reproduction, redistribution, reselling, loan, sub-licensing, systematic supply, or distribution in any form to anyone is expressly forbidden. Terms &

Conditions of access and use can be found at http://www.tandfonline.com/page/terms-and-conditions

Downloaded by [5.67.153.220] at 12:08 12 June 2014

International Journal of Research & Method in Education, 2014
http://dx.doi.org/10.1080/1743727X.2014.914486

Routledge
Taylor & Francis Group

Measuring equity in educational effectiveness research: the properties and possibilities of quantitative indicators

Anthony Kelly*

School of Education, University of Southampton, Highfield, Southampton SO17 1BJ, UK

(*Received 1 July 2013; accepted 4 April 2014*)

There has never been a published discursive review of equity measurement methodology in educational effectiveness research, though the literature on equity is growing. This paper sets out several indices that have the potential to measure it in terms of pupil attainment or in terms of how far a school (or group of schools) is from having a 'fair' proportion of its success attributable to a 'fair' proportion of its student population. The paper explores the principles and properties of three relatively simple metrics (the *Range Ratio* family of measures, the *Coefficient of Variation* and *McLoone's Index*) leading to two complex measures (*Theil's T* and finally to the Gini-based *Attainment Equity Index*). The paper investigates the warrant that the measures have (or do not have) for measuring equity by exploring them in a theoretical way so that their strengths, provenances and presumptions reveal themselves, and concludes with a discussion – the first of its kind in educational effectiveness research – on their desirable characteristics and properties. Worked examples from 2009 to 2011 UK pupil attainment data are presented in the Notes by way of worked illustrations, as are another two indices: the *Variance of the Logs* and the *Atkinson Index*.

Keywords: equity; measuring equity; educational effectiveness research

Downloaded by [5.67.153.220] at 12:08 12 June 2014

Introduction

Education policy-makers have not yet agreed measures to capture their aspiration to deliver equity in schooling outcomes. Whether using threshold pupil attainment indicators or school-level value-added measures that take account of 'context', there is no accepted quantitative metric for measuring the extent to which outcomes are equitable in terms of attainment, though as we shall see in this paper there are several possibilities. As Kelly (2012) has pointed out, many governments and supra-governmental organizations have defined equity explicitly in terms of attainment outcomes:

> Equity is the extent to which individuals can take advantage of education in terms of opportunities . . . *and outcomes*. Equitable systems ensure that the *outcomes* of education are independent of [all] factors that lead to educational disadvantage . . . Inequity in relation to gender, ethnic minority status, disability and regional disparities etc. *is not the prime focus*, [except insofar] as it contributes to overall socio-economic disadvantage. (EU 2006, 2. Emphasis added)

*Email: a.kelly@soton.ac.uk

© 2014 Taylor & Francis

And there is no reluctance, either in the USA under the 'No Child Left Behind' Act (see Owens and Sunderman [2006] and earlier Baker and O'Neil [1994]) or in Europe, to link it to accountability:

> Accountability systems in the form of central exit examinations . . . exist in most European countries. [They] should be designed to ensure a full commitment to equity . . . . (EU 2006, 6)

Traditionally, attempts to address the problem of equity in education have focused on access and on different 'unequal' sub-groups attaining reasonably 'equal' outcomes, so that the interrogation of pupil attainment data has not been to date an important tool in assessing the efficacy of policy in relation to closing attainment gaps and identifying at-risk groups. Educational effectiveness research is underpinned by the tenet that schools can make a small but significant difference to disadvantaged pupils, so the belief is implicit that education should be made more equitable. Most early studies focused on the under-achievement of disadvantaged and ethnic minority children in reading and numeracy (Sammons 2007), but as the field progressed, the idea that the general measurement of attainment was integral to understanding equity became more important, and systemic effectiveness was linked to the distribution of attainment across particular populations, most explicitly through international comparison studies like PISA[1] rather than through 'between-school' or 'within-school' studies where appropriate equity metrics have not yet been developed. Of course, any metric aimed at capturing equity and effectiveness will mask the finer detail of learning and schooling, but summative metrics nevertheless have their uses, not least in enabling comparisons to be made between and within schools over time, the importance of which has been well established through research in many countries (Goldstein et al. 2000; Gray, Goldstein, and Thomas 2001; Kyriakides and Creemers 2008; Scheerens and Bosker 1997; Teddlie, Stringfield, and Reynolds 2000). Equity is a complex issue that can impact in various ways on the experiences and outcomes of students in formal education (and in Higher Education, see Willems [2010]), just as it does in economic (Sen and Foster 1997) and sociological (Allison 1978) spheres. The methodological disadvantage in our historic attempts to define it *in education* is that we have perhaps drawn too heavily on notions of social justice, particularly in relation to equal opportunities and compensatory education for disadvantaged minority groups, without any sustained attempt to develop metrics for it. What few *have* been developed have focused on the distribution of financial resources (Berne and Stiefel 1994) and the delivered curriculum.

The family of range ratios and the median absolute deviation from the median

Any statistical measure of equity is linked intrinsically to measures of central tendency, variance, skew and dispersion. *Range* is the simplest measure of dispersion – the difference between the highest and lowest values of a given variable – but its application is limited by the fact that it uses only two values from what can be a very large set. *Range Ratio* (RR) is a better measure. It is calculated by dividing the value at a certain percentile *above the median*, by the value at a certain percentile *below the median*.

$$\mathrm{RR} = \frac{\text{high value}}{\text{low value}}.$$

Downloaded by [5.67.153.220] at 12:08 12 June 2014

In the USA, a common version of the RR in an education context is the *Federal Range Ratio* (FRR). It is used for gauging inequality in educational expenditure, and divides the difference between expenditure on the pupil at the 95th percentile and the pupil at the 5th percentile, by the expenditure on the pupil at the 95th percentile. The percentiles are used to reduce the influence of 'extreme' values.

$$\mathrm{FRR} = \frac{(\text{Spend at 95th} - \text{Spend at 5th})}{\text{Spend at 95th}}.$$

The National Centre for Education Statistics (NCES) in Washington uses FRR extensively, so for example, in the fiscal year 2009, they revealed that 'school districts had median total revenues per pupil of $11,620' and that 'the federal range ratio was 1.9 which indicated that the magnitude of the difference between total revenues per pupil at the 5th ($8323) and 95th ($23,971) percentiles of districts was approximately 190% of the value at the 5th percentile' (NCES 2012).

A third type of RR used for gauging inequality in educational expenditure is the *Inter-Quartile Range Ratio* (IQRR), got simply by dividing the expenditure on the pupil at the 75th percentile by that on the pupil at the 25th percentile (though in theory any percentiles, and not just the 75th and 25th, could be used).

$$\mathrm{IQRR} = \frac{\text{Spend at 75th}}{\text{Spend at 25th}}.$$

All these measures are clearly adaptable to pupil attainment equity whether the 'population' is a class cohort, a school, a local authority/school district or indeed a nation state, but it does require that examination *grades* be converted into *points* in some agreed way[2] or left as raw percentages. In all three cases, the larger the ratio, the greater the inequity. The lower limit is $+1$, which occurs when the numerator and the denominator are equal, and represents zero disparity between cohorts. The upper limit is $+\infty$. The advantage of these three measures, and of other similar measures like the ratio of range to inter- quartile range (IQR), is that comparisons can be made without taking extreme outliers into account, but the disadvantage remains that they use only two values from the data set and they are still not very robust to outliers, though that would depend on what percentile range is selected. An improvement might be the ratio IQR to Median, which in a crude way is analogous to the coefficient of variation (CoV) (see below). Like the CoV, it only makes sense to use it with *ratio* data – that is to say, when zero means actual zero – but an even better metric (and a better nonparametric analogy with the CoV) would be the *median absolute deviation from the median* (MADM).

$$\mathrm{MADM} = \mathrm{median}_i(|x_i - \mathrm{median}_j(x_j)|).$$

MADM is more resilient to outliers in a data set than range or even standard deviation, where the distances from the mean are squared so that the large deviations of outliers are weighted more heavily. Since it is a more robust estimator of scale than variance or standard deviation, it works better with distributions without a (defined)

Downloaded by [5.67.153.220] at 12:08 12 June 2014

mean or variance.[3] It is related to standard deviation by a scale factor constant, which depends on the distribution.

The coefficient of variation

The CoV measures variability around the mean. It is calculated by dividing the standard deviation by the mean

$$\mathrm{CoV} = \frac{\sigma}{\mu}.$$

In contrast to the various RRs, CoV takes into account *all* regions of a distribution, but like RR, for use in the area of attainment equity, it would require examination grades to be converted to 'points' in an agreed way. Perfect equity is represented by a lower limit of zero; greater inequity/disparity is represented by an upper limit of $+\infty$. In the USA, a CoV of 10 or less is considered to indicate an acceptable level of equity (CPRE 2014).

Graphically, CoV describes the kurtosis of a distribution; i.e. the extent to which the variable tends to 'pile up' around the centre. For a leptokurtic distribution (a small spread closely bunched around the mean), the peak will be high and the CoV small; a platykurtic distribution (more dispersed with less of a peak) will have a higher CoV to represent lower equity.

CoV can be broken into between-school and within-school components (see Allison [1978] and see Theil's T below) – the lower limit being the between-school component – but it can only be used for data measured on a ratio scale.[4] It is similar to the standard deviation itself, but has one major advantage over it; namely, it is dimensionless so unlike the standard deviation, it is suitable for comparison between data sets with different units or with widely differing means. The disadvantages of CoV are that when the mean is close to zero it approaches infinity and is very sensitive to small changes in the mean, and unlike standard deviation cannot be used to construct confidence intervals.

The McLoone Index

The third simple metric is the *McLoone Index* (McL), which unlike RRs and the CoV does not increase as inequity increases but increases as the distribution becomes *more equal*. Like RR and CoV, it has been used almost exclusively to examine fairness of expenditure (Sherman and Poirer 2007) and in that field it is the preferred measure when the lower part of the distribution (the 'have-nots') is of interest. It is calculated by taking the sum of per capita expenditure for each region below the median, and then dividing by the sum that would exist if each region below the median had a per capita expenditure equal to the median. (A worked example is presented in Note 1.)

$$\mathrm{McL} = \sum \frac{(\text{expenditure at or below the median})}{(\text{Number at or below the median}) \times (\text{the median expenditure})}.$$

The McL is part of a suite of school finance equity statistics in the USA,[5] which typically address both 'Fiscal Neutrality' and 'Horizontal Equity'. Fiscal Neutrality refers to the extent to which the resources available to schools vary with local fiscal

Downloaded by [5.67.153.220] at 12:08 12 June 2014

capacity (using indicators like property values and household income) and is usually measured either by the 'Correlation Coefficient (CC)' or by 'Wealth Elasticity (WE)'. The CC measures the strength of the linear relationship between two variables; typically property wealth and per capita student expenditure. The value of CC ranges between -1 and $+1$, with the upper value representing a perfect positive relationship (i.e. the two variables change in the same direction) and a value of zero indicating no relationship between the variables. (For example, as per household disposable income (say) goes down, per capita student subsidies *from central government* might go up, signifying a negative correlation, but overall per capital student expenditure on education might go down, signifying a positive correlation.) WE, on the other hand, measures the *size* of the relationship between the variables discussed above. So while disposable income (say) and overall per capita educational expenditure (say) might be positively correlated, a large increase in the former might result in only a very small increase in the latter, and this obviously would have implications for policy in terms of the amount of 'leverage' that policy-makers can exert. More specifically, WE measures the percentage change in one variable relative to a percentage change in the other variable. It ranges from zero upwards or downwards, with $+1$ indicating that the two variables increase together at the same rate.

Horizontal Equity is about students in similar (demographic, socio-economic and educational) contexts having the same levels of expenditure and teaching. It measures the extent of the inequality that exits in expenditure, and the two common measures are the CoV (discussed above) and the McL. The McL was created to indicate the degree of equality for those schools or school districts below the 50th percentile. It ranges from 0 to 1, with 1 representing perfect equality, and an index of 0.95 is considered desirable (CPRE 2014). It suffers from the disadvantage of not using all the data – data above the median is not used – and for this reason the Index has rarely been used on its own (Sherman and Poirer 2007), but unlike RR at least it takes a relatively large amount of data into account and not just two randomly nominated values. The lower limit of zero occurs when the population below the median receives none of the variable, and $+1$ represents 'perfect' equity when everyone achieves the median. It *increases* as the distribution becomes *more equitable* because the per capita below the 50th percentile is in that case approaching the median. However, when the 'disadvantaged' group is the group *above* the median, which would be the case with (say) 'pupil–teacher' ratios, it is necessary to invert the McL (or alternatively, invert the variable from 'pupil–teacher ratio' to 'teacher–pupil ratio').

Downloaded by [5.67.153.220] at 12:08 12 June 2014

Theil's *T*

The first of the two more complex metrics proposed as possible attainment equity measures is *Theil's* (1967) *T* statistic. Like the indices described above, Theil's *T* has historically been used to measure financial 'fairness', and when individual data are available it is calculated using the following equation:

$$T_{\text{indiv}} = \sum_{i=1}^{n} \left[\left(\frac{1}{n}\right) . \left(\frac{v_i}{\mu}\right) . ln \left(\frac{v_i}{\mu}\right) \right],$$

where n is the number of individuals in the population (so $1/n$ represents every individual's share of the overall T), v_i is the value of the 'achievement'[6] variable (usually income, but it could be pupil attainment) for person i, and μ is the population mean.

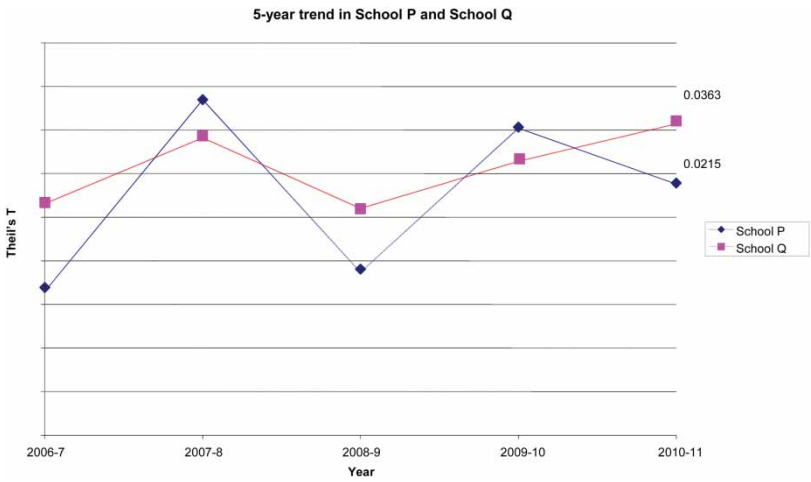

Figure 1. Comparison of Theil's T for School P and School Q over five years.

v_i/μ is the ratio of 'individual to average' and the (natural) log of it determines whether that individual Theil element is positive (when the individual value is greater than the mean), negative (when the individual value is less than the mean) or zero (when the individual value is equal to the mean). In the last case, when every individual has exactly the same (the average) amount of the variable – in other words, when there is perfect equity – T will be at its lower limit of zero; and when one individual has everything – in other words, when there is total inequity – then:

$$T = \{0\} + \{0\} + \cdots + \left\{\left(\frac{1}{n}\right).\left[\frac{v}{(v/n)}\right].ln\left[\frac{v}{(v/n)}\right]\right\}$$
$$= ln(n).$$

So the upper limit of T is $\ln(n)$ and the overall Theil for the population is the sum of the individual T_i. A worked example is presented in Note 2.

The result from two or more schools or school groups *can* be compared, *but with care*. Say School P (cf. the example in Note 2) with 420 pupils has T_P=0.021 and School Q with 680 pupils has T_Q= 0.036. The upper limit for T_P is ln(420)=6.0403, whereas the upper limit for T_Q is ln(680)=6.5221. Since School Q has more pupils than School P, T_Q would *ipso facto* be greater than T_P even if everything else about the two schools were identical, so it is difficult to draw any firm conclusions from the comparison. Theil's T only comes into its own if we are looking at trends over several years, as in Figure 1.

Theil's T can cater for hierarchical data by calculating within- and between-group components.[7] For n schools (rather than n individuals), T as a measure of equity *between schools* is given by

$$T_{\text{bet-sch}} = \sum_{i=1}^{n}\left[p_i\left(\frac{\mu_i}{\mu}\right)\cdot ln\left(\frac{\mu_i}{\mu}\right)\right],$$

Downloaded by [5.67.153.220] at 12:08 12 June 2014

where μ_i is the arithmetic mean of group i, μ is the overall population mean and p_i is the fraction of the population in school i (the equivalent of $1/n$ in the 'individual' T equation above). This is the *between-school* Theil's T, which is equivalent to the T that would be obtained if everyone in each school had that school's average.

The weighted average[8] of the *within-school* Theil's T is given by

$$T_{\text{in-sch}} = \sum_{i=1}^{n} \left[p_i \left(\frac{\mu_i}{\mu} \right) \cdot T_i \right],$$

where T_i is the T of school i, and the *overall* T for the whole distribution is given by these two terms added together. (A worked example is presented in Note 4.)

$$T_{\text{overall}} = \sum_{i=1}^{n} \left[p_i \left(\frac{\mu_i}{\mu} \right) \cdot ln \left(\frac{\mu_i}{\mu} \right) \right] + \sum_{i=1}^{n} \left[\left(\frac{p_i \mu_i}{\mu} \right) \cdot T_i \right].$$

The lower limit of the Theil's T for groups of schools is the between-school component, and sometimes when it is impossible to compute the within-school component, this is the only part of the measure calculable.[9] Theil's T for a population equals the limit of the 'between-school' component as the number of schools approaches the size of the population. The fact that a bigger population (e.g. a local authority with more schools) will have a higher T, all other things being equal, than a smaller population [because T is bounded by $\ln(n)$] is the principle disadvantage of the measure so that we need data from several years to enable strong conclusions to be drawn. And of course, for use as an attainment equity measure, like all the other indices described above, examination grades must first be converted to point scores. Another index, the Atkinson Index, can be computed from a normalized Theil's T, and this is described briefly in Note 5.

Downloaded by [5.67.153.220] at 12:08 12 June 2014

The Attainment Equity Index

Although policy-makers clearly perceive a link between equity and examination success, it is not clear what the targets *should be* or how attainment should be spread across the range of prior attainment, which school effectiveness research tells us is the major determinant of success. For example, is it 'fair' to expect the bottom 40% (say) of a school's population, as measured by their prior attainment at age eleven (say)[10] to achieve 40% of standard[11] public examination pass grades A*–C at age 16? On the one hand, it could be argued that a school should expect a *greater* proportion of pass grades from its 'more able' (and more privileged) students; on the other hand, that a school might reasonably expect a greater proportion of *higher* grades (like A* and A, in the UK) to come from its more able students, but a more or less equal distribution of *standard* grades (A*–C, in the UK) across the whole cohort. Kelly (2012) suggests that the evidence from the literature is that the latter is the official aspiration (e.g. in the UK, DFEE [1997]; in the USA, NCLB [2001] and see Owens and Sunderman [2006]; in Europe, EU [2006]), so that assuming that a given proportion of a non-selective school's examination grades is attributable to an equal proportion of the student population is a warranted starting point for developing a Gini-type measure of equity in respect of attainment, and he has used this as a starting point for developing his *Attainment Equity (Æ) Index*.

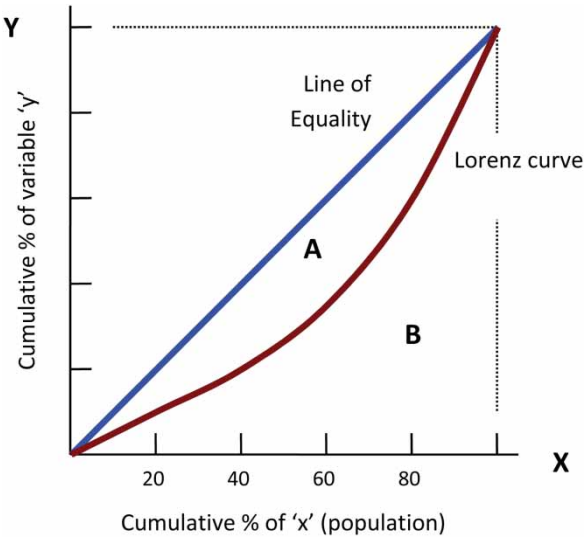

Figure 2. The Lorenz curve and the Gini 'area'.

A Gini Coefficient (Figure 2) is a measure of statistical dispersion. The straight line $y=x$ represents 'perfect' equity in the distribution of variable y over the population x, and the curve represents *actual* distribution. If B is the area under the curve, and A is the area between the line and the curve, the Gini is defined as

$$\frac{A}{(A+B)} = 1 - 2B \text{ for normalized axes.}$$

The Lorenz curve that separates A and B plots the proportion of a variable y that is cumulatively attributable to the population x, and as such it is a distribution function where every point represents a Pareto-type statement.[12] This curve defines all Gini indices, and all scale-invariant measures that exhibit 'transferability' (see below) are related to the Lorenz function.[13] If the Lorenz curve is represented by the function $y = L(x)$, the Gini is given by

$$1 - 2\int_0^1 L(x)dx.$$

The Gini is usually defined in terms of the Lorenz curve, but it can also be defined as the average absolute difference between all pairs of individuals, divided by twice the mean:

$$\frac{1}{n^2}\sum_{i=1}^{n}\sum_{j=1}^{n}\frac{|v_i - v_j|}{2\mu}.$$

And from this we can see that any Gini-based index and the CoV are related to each other as special cases of the same metric. We know (Kendall and Stuart 1977) that

Downloaded by [5.67.153.220] at 12:08 12 June 2014

variance is given by

$$\sigma^2 = \frac{1}{(2n^2)} \sum_{i=1}^{n} \sum_{j=1}^{n} (v_i - v_j)^2$$

so that

$$\text{CoV} = \frac{\sqrt{\left[(1/(2n^2)) \sum_{i=1}^{n} \sum_{i=1}^{n} (v_i - v_j)^2\right]}}{\mu}.$$

And since Gini is given by

$$\frac{1}{n^2} \sum_{i=1}^{n} \sum_{i=1}^{n} \frac{|v_i - v_j|}{2\mu}$$

both Gini and CoV are special cases of an index given by

$$\frac{\left[(1/(2n^2)) \sum_{i=1}^{n} \sum_{i=1}^{n} (v_i - v_j)^r\right]^{1/r}}{\mu},$$

where $r = 1$ for a Gini index and $r = 2$ for CoV (see Allison 1978, 870).

In developing a Gini-type index specifically for equity in pupil attainment, Kelly returns to the Gini's relationship with the Lorenz curve. Even when the Lorenz equation is unknown or undefined, he notes that area B can be approximated to trapezoids (Figure 3) and therefore that an attainment equity index can be given by[14]

$$Æ = 1 - \sum_{k=1}^{n} (X_k - X_{k-1})(Y_k + Y_{k-1}),$$

where (X_k, Y_k) are the points on the Lorenz, with $X_0 = Y_0 = 0$ and $X_n = Y_n = 1$. A worked example of the *Æ* Index is presented in Note 6.

It can be shown that it does not appreciable affect the metric if the cumulative intervals are calculated every 10% instead of every 20%, although like most measurements of this kind, the metric is lowered by lower 'granularity'.

Having developed an equation for *Æ*, the conceptual problem in developing it to take account of school characteristics (or 'context') like socio-economic circumstance is how to incorporate these factors into the Index when the research evidence does not exist to quantify their precise (percentage) impact. We know that socio-economic status has a positive correlation with attainment and that this effect tends to be greater in high-performing schools, which suggests that policies to encourage these schools to expand their provision in order to accommodate a diverse range of students has not translated into a more equitable distribution of learning outcomes across the system. However, according to Rousseau and Tate (2003), *not* linking issues of access, retention and completion to student attainment restricts teachers' conceptualizations of equity – especially as it relates to pedagogic practice with ethnic minority students – to the detriment of students. One way of adjusting the Index would be to make an 'elbow' in the line of intended equity to reduce area A at the low attainment end, but such an approach would unjustifiably produce the same size reduction in the Index for each variable (Kelly 2012), and there is no sensible way of iterating two or more consecutive adjustments without running the risk of reducing the Index to zero for no good reason.

Downloaded by [5.67.153.220] at 12:08 12 June 2014

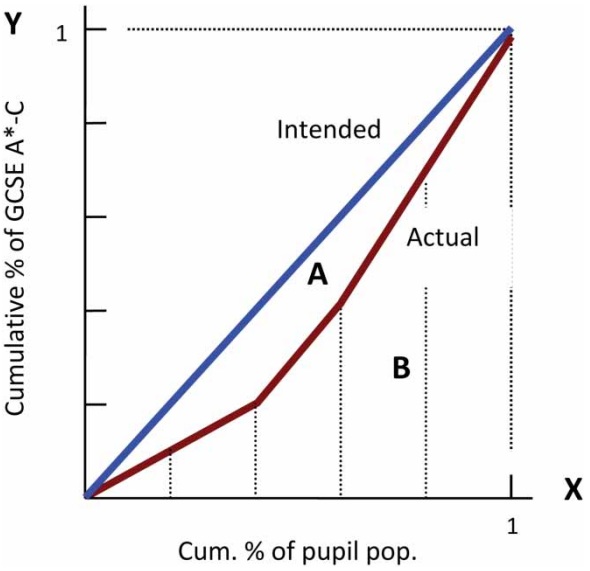

Figure 3. The Lorenz curve approximated.

However, there are some adaptations that can be made to overcome the shortcoming of not being able to take account of context. One suggestion is to combine the *Æ* Index (which is output-focused) with existing value-added measures (which are process-focused) to categorize schools in terms of both attainment equity *and context* (Figure 4), which in many ways can be treated as a proxy for a social welfare function: those with a low *Æ* Index and high contextual value-added (CVA) are adding value across the range of pupil ability and background characteristics, whereas schools with a high *Æ* Index and high CVA are adding value but not across the range of ability and background, and this can be used to refine our definition of 'differentially effective' schools.

Like Theil's *T* (Figure 1), individual school and local authority results can be used to reveal trends in equity, thereby circumventing the need for more and more context variables. For example, Figure 5 shows the calculated historical five-year trends for two statistical-neighbour local (city) authorities in England: Portsmouth and Southampton (Kelly 2012). (The calculations for the 10 *Æ* Indices are shown in Note 7.) Again, the usefulness of the Index lies not in the absolute values it calculates, but in the way it can prompt researchers and policy-makers to ask why the two authorities diverged suddenly in 2008/2009.

Kelly's Attainment Equity (Æ) Index uses 'raw' examination data, albeit adjusted for prior attainment, which in the UK at least is in line with current government thinking. It is true that the Index on its own ignores the variables that effectiveness research seeks to control, but unlike other metrics, the Index does not seek to isolate the 'school effect' beyond catering a priori for its most important factor, prior attainment. It is based on the Gini concept, which is a well-regarded metric in economics, and it has the advantages of incorporating all data and not just the extremes (as with RRs) or half of it (as with McLoone's Index), thereby allowing direct comparison between units with different size populations. It is also compatible with the Lorenz dominance

Downloaded by [5.67.153.220] at 12:08 12 June 2014

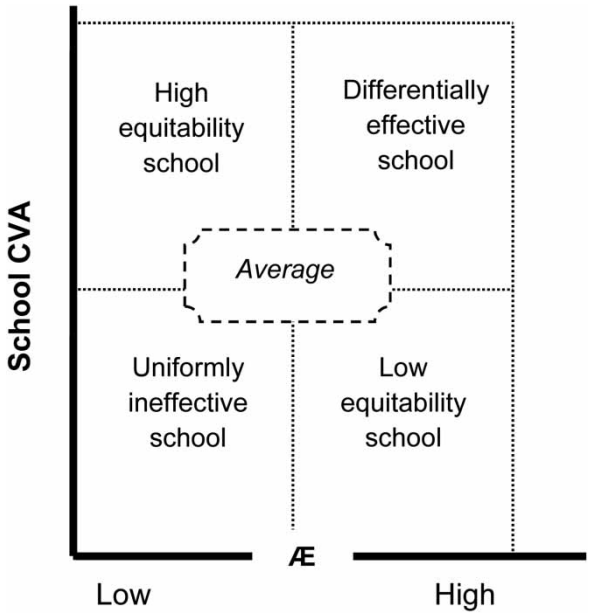

Figure 4. Combining the Attainment Equity Index with CVA.

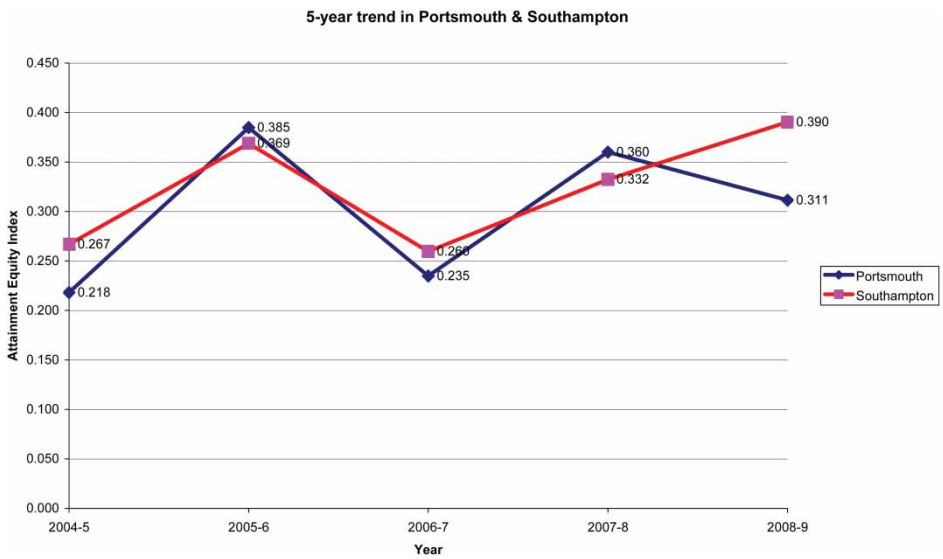

Figure 5. *Æ* Indices for Portsmouth and Southampton over five years.

criterion (cf. endnote 13) and has methodological synergies with sociological theories like intersectionality[15] (Crenshaw 1991): the Index juxtaposed with contextual measures conceptualizes equity as a matrix of multiple interrelated forms of disadvantage, which is similar to the challenges faced by educational effectiveness researchers in coping with factors impacting simultaneously on one another and on outcomes.

Downloaded by [5.67.153.220] at 12:08 12 June 2014

The way the *Æ* Index can be used in combination with measures like CVA is a strength, as is the fact that its ratio analysis allows suggestive (rather than definitive) comparisons to be made in line with other effectiveness measures (Goldstein and Spiegelhalter 1996; Knapp et al. 2006; Shen and Cooley 2008; Wohlstetter, Datnow, and Park 2008). What we can call a 'transfer principle' also applies to the Index – that is to say, if equity is transferred from one pupil or group of pupils with a lot of it, to another pupil or group of pupils with a shortage of it, the resulting distribution becomes more equal – which is not the case with other school effectiveness metrics like CVA. It is easily interpreted, can track changes over time, can act as a prompt for improvement and can give fairly immediate feedback to policy-makers (Sammons et al. 1997; Van Damme et al. 2002; Thomas, Peng, and Gray 2007). This is not to suggest that the Index does not have its disadvantages: it measures equity, but on its own does not measure opportunity, capability or wider aspects of social injustice; like other school effectiveness measures, indices for different sub-populations cannot be averaged to obtain an index for the whole population; and as with all school effectiveness measures, if desirable commodities (like how schools encourage a range of intellectual, sporting and cultural interests among young people, enable friendships and develop the ability to interact socially) are not counted in the input, they cannot be reflected in the output.

Downloaded by [5.67.153.220] at 12:08 12 June 2014

Desirable characteristics and properties of equity metrics

Scaling and zero: Any measure of inequity should be zero when everyone has the same amount of the variable, and it should have some positive value when this is not the case. How large the scalar is depends on the metric, but equity measures cannot be used meaningfully with anything other than ratio data; they cannot be used securely with interval data, for example, even when comparing different interval scales that use the same origin, although one can compare distributions on the same interval scale provided there is an underlying ratio scale behind it. Fortunately, *attainment* in most of its manifestations *has* an absolute zero – that is to say, a zero examination grade-point score *is* an absolute zero – but ratio scaling in equity effectiveness research can be an issue in the wider sense because 'zero *achievement*' from education is non-sensical when every pupil derives *some* benefit from attending school.

Scale invariance: Equity metrics should be scale invariant and multiplying by a constant should leave the dimensionless relative measures unchanged. Simple variance (σ^2), for example, fails this test. And in addition to allowing comparison between different units of measurement, there are other reasons why we should demand scale invariance:

(i) Equity should be invariant to changes of unit because instinctively, changing units should not lessen or increase equity; for example, it should not matter to a school's relative equity whether one is measuring pupil attainment in unadjusted grades or unadjusted point scores.
(ii) Real percentage increases in pupil attainment should leave equity unchanged since high achievers benefit more in absolute terms even though relative difference remains unchanged, and research suggests that self-reported well-being and happiness depends on *relative*, not *absolute*, inequity.
(iii) Scale-invariant measures decline when a positive constant in added to everyone's attainment, which seems intuitively correct as the differences between

pupils or groups becomes less significant as everyone's raw attainment increases.

Fortunately, most measures (CoV, Theil's *T* and the *Æ* Index) can be made scale invariant simply by dividing by the mean.

Transferability and sensitivity: 'Transferable' in the case of attainment means that equity decreases when attainment is transferred from someone with less of it to someone with more of it. It is an important principle and is related to the Lorenz curve that underpins the *Æ* Index. Not all scale-invariant measures obey the principle of transferability – for example, equity transfer between pupils who are on the same side of the mean does not affect equity one way or the other – and even when metrics *do* obey the principle of transferability, they are differentially sensitive to it.

(i) CoV is *equally* sensitive to all transfers, which means the measure is in fact very *in*sensitive.
(ii) The sensitivity of the Attainment Equity Index depends on the transferor's/transferee's rank – specifically on how many people are lower/higher than the transferor/transferee, respectively – rather than on actual attainment, so this Index is most sensitive to transfers around the middle of the distribution.
(iii) Regarding the sensitivity of Theil's *T*, it can be shown (Allison 1978) that as the transfer approaches zero, the effect on Theil's *T* depends only on the mean and on the number of occurrences. Whereas the change in CoV resulting from transfers depends on the difference between attainments, the change in Theil's *T* resulting from transfers depends on the ratio of attainment of the transferee to that of the transferor, so that the lower the level of attainment the more sensitive is Theil's *T*.

Of course, in school effectiveness it is important to question whether there is transfer at all, and if there is, to what extent. 'Goods' like examination results do not at first sight appear to have transferability except perhaps where a normative assessment system is in place – that is to say, where pupils are judged against the performance of others and where there are fixed percentages awarded at each level – or where teachers are triaging pupils who are on the threshold of certain grades, at the expense, one supposes, of other pupils. Measures sensitive in the lower range will tend to show less inequality because developing groups are more likely to have larger homogeneous populations of 'have-nots', and this is unfortunately the case in relation to Theil's *T* in underperforming schools.

Transforming the limits and grouped data: For full populations, CoV and Theil's *T* vary between 0 and $+\infty$, and Kelly's *Æ* Index varies between 0 and $+1$, but there is no great significance in these bounds because simple transformations can easily change them,[16] and for finite populations the measures can be made to vary between 0 and $+1$ simply by dividing each measure by its upper limit.[17] For sample data (ungrouped), the best approach is simply to apply the relevant formula, although this makes it difficult to obtain confidence intervals and standard errors. For grouped data, the mid-points of the intervals can be used.

The marginal utility value of educational attainment: It is not clear, in the context of equity measurement, whether or not attainment has diminishing marginal utility; that is to say, whether the utility gained from an increase (or lost from a decrease) in the 'consumption' of examination success is diminishing, increasing or of no relevance. What

Downloaded by [5.67.153.220] at 12:08 12 June 2014

can be said however is that when the variation in attainment does *not* have diminishing marginal utility or when it has no relevance, CoV (having a flat response to transfer) should be preferred to measures like Theil's *T* and the *Æ* Index; and when achievement has diminishing marginal utility, Theil's *T* is preferable because transfer among low achievers is more important than transfer among high achievers. Since sensitivity to transfers for Gini-based measures depends on the shape of the distribution, and since most distributions in education are normal, the *Æ* Index should (in this author's opinion) be preferred when one is concerned with changes in equity generally or among middle ranks, as will be the case with 'coasting' or comprehensive-intake schools that have survived underperformance . . . most schools in the UK in fact.

Conclusion

The ranking of one distribution as being more equitable than another has both theoretical and methodological implications, and the ordering can differ depending on which metric is chosen. In many ways the choice of metric is really a choice between definitions of equity and involves a normative contested judgement as to which distribution is the preferred/desired one. We should strive to make more explicit these judgements and the social welfare functions that drive them because while policy-makers generally favour a more-or-less equal distribution of wealth and achievement, this need not be the case, and in some circumstances may not even be desirable.

Educational inequity can occur for many reasons: the personal preferences of pupils in selecting the balance between their academic and 'other' activities; innate physical, psychological and intellectual abilities being distributed unequally over the population; socio-economic and familial factors like the pressure to leave school early; the disproportionate risk of poverty in some sections of society; public and government agency on matters like access and funding; the impact of ineffective schooling and substandard teaching.[18] Whatever its cause, measuring inequity can help gauge the effectiveness of policies aimed at reducing it, and can generate the empirical data necessary to use equity as an explanatory variable in policy analysis, particularly in relation to the distribution of 'hard' outcomes like examination attainment. Although it has been suggested (Sen 1973) that our conceptualization of inequality may be too imprecise for conventional measurement, and that the best we can hope for is a partial ordering of preferences and distributions, it seems reasonable to posit that attainment equity, as one aspect of what Sen might regard as inequality, can more easily be achieved by using data to monitor who is achieving what and understanding the gaps. At the teacher/practitioner level, where policies are (or are not) actualized in practice, equity is viewed through the lens of *process*, and differential attainment within schools is not per se the main channel for reflection. Teachers typically do not question patterns of student under-achievement in their classes, and some research (Rousseau and Tate 2003) suggests that students, particularly African-American and Hispanic-American learners, are 'allowed' to be unsuccessful because their under-achievement is not perceived as unacceptable or problematic. So any measure of equity must be useful to, and have face validity with, teachers, which when coupled with the theoretical considerations discussed in this paper, strongly suggests the Attainment Equity and Theil's *T* indices. The fact that the former uses raw grades and does not require a conversion to point scores, and can be combined with established contextual measures may give it an advantage over Theil, although it cannot be disaggregated into between-school and within-school components. Furthermore, the fact that the theoretical

Downloaded by [5.67.153.220] at 12:08 12 June 2014

underpinning for the *Æ* Index (and any Gini-based measure) are inherently strong because of the clarity of the link between the Lorenz, the social welfare function that drives desirable rankings, and transferability,[19] also adds to the appeal of the *Æ* Index, but in applying the measures to education we must appreciate that they all work better for fixed totals like norm-referenced attainment distributions, even if the spread differs between the distributions under consideration.

Notes

1. PISA is the OECD's Programme for International Student Assessment. It does not use classroom observations or specific attainment equity indices.
2. For example, the 'General Certificate in Secondary Education' (GCSE) in England and Wales – the examination taken by (nearly) all 16-year-old pupils after (typically) five years of secondary schooling – is the most widely used 'output measure' of school effectiveness in the UK, and its (full) grades are converted to points as follows: $A^*=58$; $A=52$; $B=46$; $C=40$; $D=34$; $E=28$; $F=22$; $G=16$. There are other 'grade-to-points' scales for 'half' GCSEs, vocational awards and post-16 A-level (AS and A2) examinations.
3. For example, a Cauchy or Lorenz distribution.
4. When only a sample is available, the CoV for a population is estimated using the ratio of the sample standard deviation to the sample mean, but care should be taken as this tends to be biased on the low side.
5. In the USA, expenditure figures are adjusted to reflect regional cost differences, using the Geographic Cost of Education Index (for which the Comparable Wage Index is used) and weighted for student needs (EPE 2014). Students in poverty get 1.2; students in special education get 1.9 (prior to 2001, this was 2.3). From 2005, the US national average is calculated using an average of the States' averages (prior to that, the US average was calculated using a grand mean).
6. We use 'achievement' to cover the acquisition of generic desirable outcomes, and 'attainment' to refer specifically to examination performance.
7. It is not the only index that can do this; see Note 3 on the 'Variance of the Logs'.
8. Allison (1978, 876) regards this as the fraction of the total earned by group *i*.
9. Hale (n.d.) has produced a useful online summary of Theil's T for a University of Texas 'inequality project'.
10. In England and Wales, compulsory schooling is divided into 'Key Stages': KS1 (Years 1 and 2) for ages 5–7; KS2 (Years 3–6) for ages 7–11; KS3 (Years 7–9) for ages 11–14; KS4 (Years 10 and 11) for ages 14–16. The end of Key Stage 2 (KS2) happens at age 11and is the accepted marker for secondary school attainment.
11. In the UK, GCSE grades range from A^* to G, but A^*, A, B and C are the *standard* benchmark pass grades, and the simple percentage of grades $A^* - C$ obtained across all subjects is the metric most used (and valued) by UK policy-makers in measuring rates of progression to post-compulsory and higher education (HEFCE 2005, 2010; Access 2008; Thompson 2010). See also reports from the London Challenge (DfES 2003, 2006a, 2006b; Ofsted 2010), though it is striking that there are no references in any of these reports to 'equity', 'equitability' or 'equality'.
12. The Pareto Principle (also known as 'the 80–20 rule') states that, for many natural situations, roughly 80% of effects come from 20% of causes. There is nothing special mathematically about the '80%'. When something is shared across a large population, there must be some number n between 50 and 100 such that n% is taken by $(100 - n)$% of the population (Kelly 2012).
13. It can be shown that if one Lorenz curve L_1 is everywhere lower than another L_2, then it represents lower equity no matter which index calculates it. This is 'Lorenz dominance'. The problem arises when there is no dominant Lorenz, in which case different equity indices might give rise to a different rank ordering of distributions.
14. Other estimation techniques such as Monte Carlo integration would also work.
15. Particularly 'inter-categorical' complexity which acknowledges the de facto existence of inequality categorizations and uses them to make measurements across multiple dimensions and over time.

Downloaded by [5.67.153.220] at 12:08 12 June 2014

16. For example, to change the upper limit of the CoV from ∞ to 1, take (CoV)/(CoV+1).
17. $1-(1/n)$ for Gini-type; $\sqrt{(n-1)}$ for the CoV; and $\ln(n)$ for Theil's *T*.
18. Frempong, Reddy and Kanjee (2011) in their South Africa study of Grade 6 system-level data, identified three indicators of equity: the socio-economic status of learners' home environments, racial discrimination, and inclusivity with reference to special educational needs.
19. An ordering of Lorenz curves (L) implies an ordering of the social welfare function (S) and this can easily be tied to transferability If one Lorenz curve L_1 is everywhere equal to or lower than another L_2, then $S(L_1) < S(L_2)$, where S is the social welfare function; and if L_1 can be moved up to L_2 by a series of transfers from 'haves' to 'have-nots', then (again) $S(L_1) < S(L_2)$.

References

Access. 2008. *GCSE Equivalence and the Access to Higher Education Diploma*. Gloucester: QAA. Accessed October 29, 2010. http://www.accesstohe.ac.uk/AboutUs/Publications/Documents/GCSE-equivalence-and-AHE.pdf

Allison, P. 1978. "Measures of Inequality." *American Sociological Review* 43: 865–880.

Atkinson, A. 1970. "On the Measurement of Economic Inequality." *Journal of Economic Theory* 2 (3): 244–263.

Baker, E., and H. O'Neil. 1994. "Performance Assessment and Equity: A View from the USA." *Assessment in Education: Principles Policy and Practice* 1 (1): 11–26.

Berne, R., and L. Stiefel. 1994. "Measuring Equity at the School Level: The Finance Perspective." *Educational Evaluation and Policy Analysis* 16 (4): 405–421.

CPRE (Consortium for Policy Research in Education). 2014. University of Wisconsin, Madison. Accessed Feb 21, 2014. http://cpre.wceruw.org/contact/index.php

Crenshaw, K. W. 1991. "Mapping the Margins: Intersectionality, Identity Politics, and Violence Against Women of Color." *Stanford Law Review* 43 (6): 1241–99.

DfEE. 1997. *Excellence in Schools*. London: DfEE.

DfES. 2003. *The London Challenge: Transforming London Secondary Schools*. Nottingham: DfES.

DfES. 2006a. *London Challenge: Families of Schools, May 2006, Secondary Schools*. London: DfES.

DfES. 2006b. *London Challenge: Families of Schools, May 2006, Primary Schools.* 4 volumes. London: DfES.

EPE (Editorial Project in Education). 2014. Accessed February 21, 2014. http://www.edweek.org/rc/2007/06/07/edcounts.html

EU (Commission of the European Communities). 2006. *Communication from the Commission to the Council and to the European Parliament: Efficiency and Equity in European Education and Training Systems. 8.9.2006, COM*. Brussels: EU.

Frempong, G., G. Reddy, and A. Kanjee. 2011. "Exploring Equity and Quality Education in South Africa Using Multilevel Models." *Compare: A Journal of Comparative and International Education* 41 (6): 819–835.

Goldstein, H., P. Huiqi, T. Rath, and N. Hill. 2000. *The Use of Value Added Information in Judging School Performance*. London: Institute of Education.

Goldstein, H., and D. Spiegelhalter. 1996. "League Tables and Their Limitations: Statistical Issues in Comparisons of Institutional Performance." *Journal of the Royal Statistical Society A* 159: 385–443.

Gray, J., H. Goldstein, and S. Thomas. 2001. "Predicting the Future: The Role of Past Performance in Determining Trends in Institutional Effectiveness at a-level." *British Educational Research Journal* 27 (4): 1–15.

Hale, T. n.d. "The Theoretical Basics of Popular Inequality Measures." University of Texas Inequality Project. Accessed February 21, 2014. utip.gov.utexas.edu/tutorials/theo_basic_ineq_measures.doc

HEFCE (Higher Education Funding Council for England). 2005. *Higher Education Admissions: Assessment of Bias. Report 2005/47*. Bristol: HEFCE.

HEFCE (Higher Education Funding Council for England). 2010. *Trends in Young Participation in England. Report 2010/03*. Bristol: HEFCE.

Downloaded by [5.67.153.220] at 12:08 12 June 2014

Kelly, A. 2012. "Measuring Equity and Equitability in School Effectiveness Research." *British Educational Research Journal* 38 (6): 977–1002.
Kendall, M., and A. Stuart. 1977. *Advanced Theory of Statistics*. New York: Macmillan.
Knapp, M. S., J. A. Swinnerton, M. A. Copland, and J. Monpas-Huber. 2006. *Data-informed Leadership in Education*. Seattle, WA: University of Washington Center for the Study of Teaching and Policy.
Kyriakides, L., and B. Creemers. 2008. "A Longitudinal Study on the Stability Over Time of School and Teacher Effects on Student Outcomes." *Oxford Review of Education* 34 (5): 521–45.
NCES. 2012. "Revenues and Expenditures for Public Elementary and Secondary School Districts School Year, 2008–09." Washington DC. Accessed February 21, 2014. https://nces.ed.gov/pubs2012/2012313/findings.asp
NCLB (No Child Left Behind) Act. 2001. "Public Law Print of PL 107–110." Accessed October 20, 2010. http://www2.ed.gov/policy/elsec/leg/esea02/index.html
Ofsted. 2010. *London Challenge – Reference no. 100192* London: Ofsted.
Owens, A., and G. Sunderman. 2006. *School Accountability under NCLB: Aid or Obstacle for Measuring Racial Equity?* Cambridge, MA: Harvard University.
Rousseau, C., and W. F. Tate. 2003. "No Time Like the Present: Reflecting on Equity in School Mathematics." *Theory into Practice* 42 (3): 210–216.
Sammons, P. 2007. *School Effectiveness and Equity: Making Connections*. Reading: CfBT.
Sammons, P., S. Thomas, P. Mortimore, A. Walker, A. Hind, R. Smees, J. Bausor, and R. Cairns. 1997. *Forging Links: Effective Schools and Effective Departments*. London: Paul Chapman.
Scheerens, J., and R. Bosker. 1997. *The Foundations of Educational Effectiveness*. Oxford: Pergamon.
Sen, A. 1973. *On Economic Inequality*. New York: Norton.
Sen, A., and J. Foster. 1997. *On Economic Inequality*. Oxford: Oxford University Press.
Shen, J., and V. Cooley. 2008. "Critical Issues in Using Data for Decision-making." *International Journal of Leadership in Education* 11 (3): 319–29.
Sherman, J. D., and J. M. Poirer. 2007. *Educational Equity and Public Policy: Comparing Results from 16 Countries. UIS Working paper 6*. UNESCO: Institute for Statistics.
Teddlie, C., S. Stringfield, and D. Reynolds. 2000. "Context Issues within School Effectiveness Research." In *The International Handbook of School Effectiveness Research*, edited by C. Teddlie and D. Reynolds, 160–186. London: Falmer Press.
Thomas, S., W. J. Peng, and J. Gray. 2007. "Modelling Patterns of Improvement Over Time: Value Added Trends in English Secondary School Performance Across Ten Cohorts." *Oxford Review of Education* 33 (3): 261–95.
Thompson, J. 2010. *Male and Female Participation and Progression in Higher Education*. Oxford: Higher Education Policy Institute.
Theil, H. 1967. *Economics and Information Theory*. Chicago, IL: Rand McNally.
Van Damme, J., B. De Fraine, G. Van Landeghem, M. Opdenakker, and P. Onghena. 2002. "A New Study on Educational Effectiveness in Secondary Schools in Flanders: An Introduction." *School Effectiveness and School Improvement* 13 (4): 383–97.
Willems, J. 2010. "The Equity Raw-score Matrix: A Multi-dimensional Indicator of Potential Disadvantage in Higher Education." *Higher Education Research and Development* 29 (6): 603–621.
Wohlstetter, P., A. Datnow, and V. Park. 2008. "Creating a System for Data-driven Decision-making: Applying the Principal-agent Framework." *School Effectiveness and School Improvement* 19 (3): 239–59.

Downloaded by [5.67.153.220] at 12:08 12 June 2014

Note 1

In 2011, School P in the UK had 420 pupils doing GCSE and their scores in points, converted in the usual way (see endnote 2), are shown in Table 1. [The range of actual scores has been simplified for the purposes of this example, but attainment ranged from four Es and four Fs (worst, 200 points) to 10 A*s and one A (best, 632 points).] The median pupil is one of the 96 pupils on 440 points, having achieved grades A,BB,CCCC,DDDD (or equivalent).

Table 1. GCSE points for School P in 2011

School P	
GCSE grade points 2011	
GCSE points per pupil	Number of pupils
200	3
242	15
296	35
314	37
374	58
440	96
482	100
500	45
542	20
602	5
614	3
626	2
632	1
	$N = 420$

$$McL = \sum \frac{(\text{attainment} \leq \text{the median})}{[(\text{Number} \leq \text{the median}) \times (\text{median attainment})]}$$
$$= \frac{(600 + 3630 + 10360 + 11618 + 21692 + 27280)}{(210 \times 440)}$$
$$= 0.8136.$$

Note 2

Using the School P data in Table 1, T for the school works out at 0.0215, being the sum of all the individual T_i (Table 2).

Downloaded by [5.67.153.220] at 12:08 12 June 2014

Table 2. The calculation of Theil's T for School P (GCSE 2011).

GCSE points (v_i)	No. pupils (n_i)	Ratio of points to average ($r = v_i/\mu$)	ln(r)	Individual T_i $n_i[(1/n)(r)\ln(r)]$
200	3	0.47081	−0.75330	−0.00253
242	15	0.56968	−0.56268	−0.01145
296	35	0.696798	−0.36126	−0.02098
314	37	0.739171	−0.30223	−0.01968
374	58	0.880414	−0.12736	−0.01548
440	96	1.035782	0.03516	0.00832
482	100	1.134652	0.12633	0.03413
500	45	1.177024	0.16299	0.02056
542	20	1.275895	0.24365	0.01480
602	5	1.417137	0.34864	0.00588
614	3	1.445386	0.36838	0.00380
626	2	1.473635	0.38773	0.00272
632	1	1.487759	0.39727	0.00141
Average (μ) = 424.8	n = 420			T_P = 0.0215

Note: The upper limit for T in this example is ln(420)=6.0403

Downloaded by [5.67.153.220] at 12:08 12 June 2014

Note 3: the Variance of the Logs

Variance of the Logs (VoL) is a scale-invariant measure defined as

$$\frac{1}{n}\sum_{i=1}^{n}[\ln v_i - \text{average}\ln v_i]^2.$$

VoL is not defined when achievements of zero are included, but like Theil's T and CoV, it can be broken into 'between-school' and 'within-school' components (Allison 1978, 876). However, unlike Theil's T and CoV, the lower limit is *not* the 'between-school' component. In fact, the 'between-school' component cannot be calculated because the geometric means are not calculable.

VoL does not obey the principle of transferability at high (defined as greater than e, the base of the natural log, times the mean) levels of achievement (though it does obey the principle at low levels), when equity perversely *increases* with the transfer of achievement from the 'have-nots' to the 'haves'.

Note 4

Local Authority 'L' has five schools (A–E). Each school's GCSE results (in points scores, rather than grades) is given in Table 3; so for example, in School A, 8 pupils scored 296 points exactly, 20 scored 374 exactly, and so on. (Obviously, this is a simplified scenario as in real life the 88 pupils in School A had a much wider range of points scores.) To compute the equity for L, it is necessary to calculate within-school and between-school components. (Individual school data have been calculated and is given in the first column of Table 3, and the calculation of individual school T is done on Table 4).

Table 3. The calculation of Theil's T for local authority L using 2011 GCSE results.

Local authority (L) five schools: A–E	No. pupils (n_i)	GCSE points score (v_i)	$T_{\text{Between-school}}$: $p_i(\mu_{\text{sch}}/\mu_{\text{L}})\ln(\mu_{\text{sch}}/\mu_{\text{L}})$	T_{School}: $1/n_{\text{sch}}(v_i/\mu_{\text{sch}})\ln(v_i/\mu_{\text{sch}})$ (Table 4)	$T_{\text{Within-sch}}$: $p_i(\mu_{\text{sch}}/\mu_{\text{L}})T_{\text{sch}}$
School A	8	296	0.005484	0.01616	0.00304
$n_{\text{A}} = 88$	20	374			
$\mu_{\text{A}} = 455.55$	32	482			
$\mu_{\text{A}}/\mu_{\text{L}} = 1.02960$	20	500			
$n_{\text{A}}/N = P_i = 0.1826$	8	602			
School B	11	296	−0.03445	0.01111	0.00212
$n_{\text{B}} = 110$	29	314			
$\mu_{\text{B}} = 369.25$	43	374			
$\mu_{\text{B}}/\mu_{\text{L}} = 0.83455$	20	440			
$n_{\text{B}}/N = P_i = 0.2282$	7	482			
School C	9	242	0.011854	0.04980	0.01091
$n_{\text{C}} = 100$	20	296			
$\mu_{\text{C}} = 467.06$	32	440			
$\mu_{\text{C}}/\mu_{\text{L}} = 1.05562$	20	626			
$n_{\text{C}}/N = P_i = 0.2075$	19	632			
School D	10	296	0.017919	0.01150	0.00306
$n_{\text{D}} = 120$	35	440			
$\mu_{\text{D}} = 473.25$	43	482			
$\mu_{\text{D}}/\mu_{\text{L}} = 1.06961$	27	542			
$n_{\text{D}}/N = P_i = 0.2490$	5	614			
School E	4	200	0.003529	0.03006	0.00410
$n_{\text{E}} = 64$	10	296			
$\mu_{\text{E}} = 454.06$	32	482			
$\mu_{\text{E}}/\mu_{\text{L}} = 1.02624$	16	542			
$n_{\text{E}}/N_{\text{L}} = P_i = 0.133$	2	602			
LA average, $\mu_{\text{L}} = 442.45$	$N_L = 482$		$T_{\text{Between-school}} = 0.00434$	$T_L = 0.02757$	$T_{\text{Within-school}} = 0.02323$

Table 4. The calculation of individual school Theils for L using 2011 GCSE results.

Calculating T_i for each school (to calculate $T_{\text{Within-school}}$)						
Local authority (L) five schools: A–E	No. pupils (n_i)	GCSE points score (v_i)	$r = v_i/\mu_{\text{Sch}}$	$\ln(r)$	$n_i(1/n_{\text{Sch}})(r)\ln(r)$	T_{Sch} (Σ)
School A	8	296	0.64976	−0.43115	−0.02547	
$n_{\text{School}} = 88$	20	374	0.82099	−0.19725	−0.03680	
$\mu_{\text{A}} = 455.55$	32	482	1.05806	0.05644	0.02171	0.01616
	20	500	1.09757	0.09310	0.02322	
	8	602	1.32148	0.27875	0.03349	
School B	11	296	0.80162	−0.221114	−0.01772	
$n_{\text{School}} = 110$	29	314	0.85037	−0.162081	−0.03634	
$\mu_{\text{B}} = 369.25$	43	374	1.01287	0.0127819	0.00506	0.01111
	20	440	1.19160	0.1753008	0.03798	
	7	482	1.30535	0.2664702	0.02214	

(*Continued*)

Downloaded by [5.67.153.220] at 12:08 12 June 2014

Table 4. (*Continued*).

Calculating T_i for each school (to calculate $T_{\text{Within-school}}$)						
Local authority (L) five schools: A–E	No. pupils (n_i)	GCSE points score (v_i)	$r = v_i/\mu_{\text{Sch}}$	ln(r)	$n_i(1/n_{\text{Sch}})(r)\ln(r)$	T_{Sch} (Σ)
School C	9	242	0.51814	−0.65752	−0.030662	
$n_{\text{School}} = 100$	20	296	0.63375	−0.45610	−0.057811	
$\mu_{\text{C}} = 467.06$	32	440	0.94206	−0.05968	−0.017992	0.04980
	20	626	1.34030	0.29289	0.0785127	
	19	632	1.35315	0.30243	0.0777545	
School D	10	296	0.62546	−0.46926	−0.02446	
$n_{\text{School}} = 120$	35	440	0.92974	−0.07285	−0.01976	
$\mu_{\text{D}} = 473.25$	43	482	1.01849	0.01832	0.00669	0.01150
	27	542	1.14527	0.13564	0.03495	
	5	614	1.29741	0.26037	0.01408	
School E	4	200	0.44047	−0.81991	−0.02257	
$n_{\text{School}} = 64$	10	296	0.65190	−0.42787	−0.04358	
$\mu_{\text{E}} = 454.06$	32	482	1.06153	0.05972	0.03169	0.03006
	16	542	1.19368	0.17704	0.05283	
	2	602	1.32582	0.28203	0.011685	

$T_{\text{Between-school}}$ has been calculated (see column 4 of Table 3), and T_i for each school has then been calculated using the data in Table 4 (reported in column 5 of Table 3). Using this, $T_{\text{Within-school}}$ has been calculated (see column 6 of Table 3). The result is $T_{\text{Between-school}} = 0.00434$, $T_{\text{Within-school}} = 0.02323$, and T_{overall} for the local authority is 0.02757, from which we can conclude that the difference in attainment within schools causes most (84%) of the inequity and the difference in attainment between schools is relatively insignificant.

Downloaded by [5.67.153.220] at 12:08 12 June 2014

Note 5: Atkinson's Index

Atkinson's Index is useful in determining which end of a distribution contributes most to inequity (Atkinson, 1970). Greater weight is placed on changes in a given part of the distribution by choosing an appropriate positive coefficient (ε, the 'inequality aversion coefficient') between 0 and 1: the Index can be made more sensitive to changes at the low/'have-not' end of the distribution by making ε high; conversely, as ε approaches zero, the Index becomes more sensitive to changes in the upper end of the distribution. The Atkinson Index is defined in two parts (to avoid a zero denominator) as follows:

When the coefficient $\varepsilon \neq 1$:

$$\text{AI} = 1 - \left(\frac{1}{\mu}\right)\left[\left(\frac{1}{n}\right)\sum_{i=1}^{n}(v_i)^{1-\varepsilon}\right]^{1/1-\varepsilon}.$$

When the coefficient $\varepsilon = 1$:

$$\text{AI} = 1 - \left(\frac{1}{\mu}\right)\prod_{i=1}^{n}(v_i)^{1/n},$$

where v_i is again the variable in question, n is the number of individuals in the population and μ is the mean.

A form of the Atkinson Index can also be computed from a normalized Theil Index (T), where $\varepsilon=1$, using the formula $\mathrm{AI}=1-\mathrm{e}^{-T}$.

Note 6

School R has the following GCSE A*–C distribution for 2009 (using prior attainment at KS2 as benchmark): the bottom 20% (at KS2) obtained 8% GCSE grades A*–C; the bottom 40% obtained 19% GCSE grades A*–C; the bottom 60% obtained 40% GCSE grades A*–C; the bottom 80% obtained 68% GCSE grades A*–C.

Clearly $X_k - X_{k-1} = 0.2$, so the Attainment Equity Index for School R is

$$\begin{aligned} Æ_{\mathrm{R}} &= 1 - 0.2\sum_{k=1}^{n}(Y_k + Y_{k-1}) \\ &= 1 - 0.2[(0.08+0)+(0.19+0.08)+\cdots+(1+0.68)] \\ &= 0.260 \end{aligned}$$

Note 7

Table 5 below calculates the *Æ* Indices over a five-year period for Portsmouth and Southampton local authorities in 2009.

Table 5. *Æ* Index 5-year trend for Portsmouth and Southampton in 2009 [N_{p} = number of pupils (in all schools) with matching GCSE and KS2 data].

Cum % by KS2 prior attainment			0	20	40	60	80	100	Æ Index
	Year	N_p							
Portsmouth LA 851	2004–2005	1932	0	8.74	27.96	40.87	67.90	100	0.218
	2005–2006	1990	0	2.43	10.37	32.07	58.92	100	0.385
	2006–2007	1853	0	8.03	22.75	42.7	67.83	100	0.235
	2007–2008	1856	0	3.53	8.62	34.66	63.19	100	0.360
	2008–2009	2215	0	5.30	16.7	35.97	64.19	100	0.311
Southampton LA 852	2004–2005	2319	0	6.64	23.31	37.97	65.35	100	0.267
	2005–2006	2243	0	3.48	11.94	31.66	60.71	100	0.369
	2006–2007	1848	0	7.35	21.25	40.16	66.34	100	0.260
	2007–2008	1946	0	5.11	14.25	34.97	62.56	100	0.332
	2008–2009	2366	0	2.86	11.08	29.01	59.45	100	0.390

Downloaded by [5.67.153.220] at 12:08 12 June 2014