RESEARCH ARTICLE

The probity of free school meals as a proxy measure for disadvantage

Abstract

The use of free school meal (FSM) data is widely prevalent in official estimates of educational disadvantage as well as in educational research reports in Britain. However, while there has been some concern expressed about the measure, there has, to our knowledge, been no systematic test of its appropriateness. In this paper we test for its appropriateness as a measure, taking into account the dynamics of poverty and the error that can be associated with its application in judging school performance. We find that it is a coarse and unreliable indicator by which school performance is judged and leads to biased estimates of the effect of poverty on pupils' academic progress. These findings raise important policy questions about the quality of indicators used in judging school performance.

Keywords: Free School Meals Eligibility; flexible labour markets; measurement error; reliability; bias; value added analysis; progress in mathematics at KS1.

RESEARCH ARTICLE

The probity of free school meals as a proxy measure for disadvantage

DAPHNE KOUNALI[a]*, TONY ROBINSON[b] , HARVEY GOLDSTEIN[c] AND HUGH LAUDER[d]

Abstract

The use of free school meal (FSM) data is widely prevalent in official estimates of educational disadvantage as well as in educational research reports in Britain. However, while there has been some concern expressed about the measure, there has, to our knowledge, been no systematic test of its appropriateness. In this paper we test for its appropriateness as a measure, taking into account the dynamics of poverty and the error that can be associated with its application in judging school performance. We find that it is a coarse and unreliable indicator by which school performance is judged and leads to biased estimates of the effect of poverty on pupils' academic progress. These findings raise important policy questions about the quality of indicators used in judging school performance.

Keywords: Free School Meals Eligibility; flexible labour markets; measurement error; reliability; bias; value added analysis; progress in mathematics at KS1.

---

[a]* Centre for Multilevel Modelling, Graduate School of Education, University of Bristol, 2 Priory Road, Bristol, BS8 1TX.
Email: masdkk@bath.ac.uk
[b] Department of Mathematical Sciences, University of Bath, Bath BA2 7AY
[c] Graduate School of Education, University of Bristol, 35 Berkeley Square, Bristol BS8 1JA
[d] Department of Education, University of Bath, Bath, BA2 7AY

Introduction

The use of free school meal (FSM) data is widely prevalent in official estimates of educational disadvantage as well as in educational research reports in Britain. However, while there has been some concern expressed about the measure, there has, to our knowledge, been no systematic test of its appropriateness. The appropriateness of FSM eligibility as a proxy measure of deprivation is an issue worthy of debate in the context of the questions that educational research on pupil's academic progress seeks to address. In particular, judgements about school performance that use this measure may be criticised on two grounds. The first and more widely acknowledged, concerns the approximate nature of FSM eligibility as a measure of social deprivation. The second concerns the possibility that the population of those on FSM is highly unstable and that any calculation or judgement is likely, therefore, to be an underestimate of the real disadvantage that a school or student confronts. It is with this latter possibility that this paper is concerned. For example, if the pool of students who take up FSM over a given period is larger than those counted at a specific time then the proportion of disadvantaged in a school will be higher than acknowledged. If the stability varies across schools this in turn will raise equity questions.

The paper provides some background in terms of the nature of the flexible labour market in Britain and the distribution of welfare benefits related to it, since both impact on the nature of child poverty of which FSM is assumed to be a reliable indicator.

Economic Deprivation and the Nature of the Flexible Labour Market

Britain has one of the highest levels of child poverty as measured by the OECD (Bradbury et al., 2001). There are at least two related reasons for this. Firstly, many

children in poverty are in single parent families (Gregg & Wadsworth, 2003). Secondly, the nature of the labour market is such that mothers are deterred from entering it and when they do, they may find paid work unstable[1]. The British labour market can be described as flexible, that is, hiring and firing is much easier in this country than in many European countries (Brown et al., 2001). It can be hypothesized that this has led to a degree of instability in careers, especially of the low skilled who move between low wage employment and state benefits. At the same time, provision for child care is not well developed. In contrast, in the Nordic countries the state provides both jobs and childcare for women workers (Esping-Andersen, 2006). The consequence has been a far lower incidence of child poverty (Bradbury et al., 2001). As a result, in Britain, low wage workers and especially lone parents may have children who are eligible for FSM but this eligibility may be unstable, either because they re-partner and their economic fortunes rise or because they find temporary employment. Underlying these observations is the point that the economic circumstances of welfare populations may be unstable, as families change in composition with children arriving and leaving and labour market fortunes change. If FSM is to stand as a proxy indicator of disadvantage, then in the light of the above its reliability may be in question.

FSM and how it is used

The eligibility for FSM is frequently used as a factor representing economic disadvantage in investigations of educational attainment including valued-added analyses, and truancy (Goldstein, 1997, Plewis & Goldstein, 1997, Sammons et al., 1997, Yang et al., 1999), studies of school composition (Hutchison, 2003, Schagen &

[1] There has been an increase in employment for lone mothers by 11% between 1993-2002 but it is still low by Nordic standards. A range of other policies have also been implemented to help support solo parents.

Schagen, 2005, Strand, 1997) and research on socially-segregated schooling (Allen & Vignoles, 2006, Goldstein & Noden, 2003) and school choice (Gorard et al., 2003). More directly, Local Education Authorities incorporate FSM figures in their calculations of extra provision for Special Educational Needs and Additional Educational Needs. The Department of Education and Skills includes FSM in the publication of school league tables (DFES, 2003, DFES, 2005a, DFES, 2005b) while in Scottish schools is also used for target setting purposes (Croxford, 2000). Over recent years the eligibility criteria have changed consequent to changes in benefits. This can lead to additional problems in using FSM data when investigating economic deprivation over a prolonged timescale. The current eligibility criteria are that parents do not have to pay for school meals if they receive any of the following:

- Income Support
- Income-based Jobseeker's Allowance
- Support under Part VI of the Immigration and Asylum Act 1999
- Child Tax Credit, provided they are not entitled to Working Tax Credit and have an annual income (as assessed by HM Revenue & Customs) that does not exceed £13,480
- The Guarantee element of State Pension Credit. Children who receive Income Support or income-based Job Seeker's Allowance in their own right qualify as well.

The popularity of FSM is based mainly upon its availability. There is no other measure reflecting individual economic disadvantage that is universally or even widely available[2].

[2] One other measure that is becoming popular in research is the Index of Multiple Deprivation (IMD) Noble, M., Wright, G., Dibben, C., Smith, G., McLennan, D., Anttila, C., Barnes, H., Mokhtar, C., Noble, S., Avenell, D., Gardner, J., Govizzi, I. & Lloyd, M. (2004) Indices of Deprivation 2004(London, Deputy Prime Minister Office: Neighbourhood Renewal Unit).. However this does not relate directly to individuals but to the small geographical area in which they live, known as a low level Super Output Area (SOA) containing on average about 1500 people. IMD is a composite index based on indices grouped within seven domains: Income, Employment, Health, Deprivation and disability, Education, skills and training, Barriers to housing and services, Living environment, Crime.

The study – HARPS project

This paper examines the proportion of children who were eligible for free school meals over time using data from Hampshire schools. The probity of FSM as an indicator of economic disadvantage is also compared to other measures of disadvantage, namely home ownership and receipt of working tax credit for a sub sample of Hampshire schools where more detailed information was collected. In order to understand what FSM is measuring, we also examine the association between family SES and economic status as measured by these three indicators.

Study design

The HARPS project is an acronym for 'Hampshire Research with Primary Schools' and looks at the impact of school composition upon student academic progress. The main aim of the study is to estimate and better understand *compositional effects* at the primary school level. Compositional effects are the peer group effects on pupils' achievement, over and above those of an individual's own characteristics. The research design is both quantitative and qualitative. The project has 3 nested parts;

- A large scale analysis of over 300 primary schools
- A study of a sub sample of 46 schools in the Greenwood (pseudonym) area.
- More detailed case studies of 12 schools.

A further remit of the study is to evaluate the probity of FSM[3] as a proxy for the economic element of SES. To this end we have collected and analysed detailed family background information from the year 3 children in the 46 Greenwood schools. These schools have been selected so as to compare the findings from these schools with those of the overall Hampshire cohort. This will test the probity FSM as a measure of

[3] Individual FSM eligibility for these pupils is available from Hampshire LEA over the current and a number of previous years.

disadvantage for the Hampshire dataset and will throw light on its usefulness in school effectiveness studies in general.

Figure 1: The deprivation geography of Hampshire based on the Multiple Deprivation Index.

The Greenwood sub sample contains family background data on 1653 year 3 pupils from a total of 1942 students attending 46 out of all 50 schools in the Greenwood area during the second semester of the academic year 2004 - 2005. Relevant to economic status these data include: occupational group (Goldthorpe & Hope, 1974), working status; home ownership, whether in receipt of Working Tax Credit, whether in receipt of FSM, level of education of the parent and house movements during the child's lifetime. It should also be noted that this sub sample is not a random sample of the Hampshire County as can be seen in Figure 1. The deprivation geography[4] of Hampshire according to the multiple deprivation index suggests that the selected Greenwood schools cover most of the deprivation range in Hampshire.

Previous research suggests that higher SES people are more likely to respond in surveys than their lower strata counterparts (Goyder et al., 2002). Our sample does not seem to be exempt from these difficulties. Despite the high questionnaire return rates (85%), we have found that non-responders were twice more likely to be FSM eligible according to their LEA records. This should be borne in mind when interpreting our findings.

Background data on the Greenwood sub sample

Female responders accounted for 90% of the returned questionnaires. This is also a sample that is predominantly white with 92.7% of the responders being white-British

[4] Figure 1 depicts the distribution of rankings for the multiple deprivation index where higher values represent lower levels of deprivation.

or Irish, another 3.4% being white-mixed and another 3.3% all other ethnic or racial backgrounds.

The proportion of the responders who were single parents was 17.2% (n=284) and another 2.5% (n=41) did not live with the child. The majority (57%) of these children were living in small families of two children or less (9.6% were single-child families) and less than 18% were families with more than 4 children.

Only 32.6% of these children had lived in the same house for most of their lives (7-8 years) and 15.6% changed houses more than 3 times. The proportion of single parents among children who changed house three times or more was was (30.6%) much higher when compared with that among children who lived in the same house for most of their lives (12.5%). Home changes are also related to home ownership**.** In this sample 71.5% of parents owned their homes. Only 15.4% of the children who had always lived in the same house were in rented accomodation. The proportion of rented housing among children who changed home once or twice was 26.1% and this rose to 49.2% for children with more home changes.

The characterization of family SES status was based on the occupation of the male carer whenever such response was available (73%). This was not always available, in particular, in the case of lone mothers. Lone parenthood was predominantly (64%) the reason for the unavailability of occupational information for the male carer, followed by lack of employment.

When there were two carers**,** the occupations of both carers as well as their employment history were also recorded. Combining occupational information from both partners by considering the highest ranked occupation reported by the couple including past occupations is commonly used to characterize family SES (Daly et al., 2006). We found that such characterizations of family SES led to inconsistencies with

local and national statistics and grossly underestimated family SES in this population (Hampshire, 2006). Further analysis revealed that this bias could be related to gender differences in employment and low-paid work. We found that 45% of the highest ranked present or previous occupations in a family were those of male responders. The majority of the highest ranked occupations refer to the past (64.4%). Moreover, among the occupation records referring to the present, the majority (55.7%) refer to a male carer with the highest ranked occupation in the family while the majority of past ones (61.1%) refer to a female carer from low occupational categories. This suggests that the major income earner is male.

Table 1 gives a summary of the distribution of socioeconomic status as recorded from the occupational data with the level of education attained by the responder.

Table 1: Parental education attainment and family SES

In Table 2 we summarize the distribution of the working mode of the couple according to SES. We also present the incidence of single parenthood and family size according to working mode. We first note that non-response to questions on occupation is predominantly due to unemployment since 93.4% of such non-responders were found not to be working currently. The remaining couples with current non-working status mainly do not report occupation.

Table 2: The association of family SES with employment, lone parenthood and family size

It is also worth noting that amongst non-working families the majority are single parents with three children or more. The most prevalent working mode of a couple (38%) consists of one partner working full time and the other part-time (n=592) or both being self-employed (usually running a family business (n=33). The next most

prevalent working mode is that of a couple where only one is full-time (26.5%) followed by couples where both are working full-time (12.3%). The proportion of families consisting of single part-time bread winners is 5.5%. In another 4.6% of families, one is working full-time and the other is self-employed and 4.7% consists of families where there is a single self-employed bread-winner (n=77) or both are part-timers (n=8).

How closely does FSM eligibility reflect deprivation?

Out of the 1653 families, 124 (7.5%) reported that they were in receipt of FSM. Of these families, 13 also claimed to receive Working Tax Credit which implies that their income was less than £13,230, according to the 2004 eligibility criteria. Single parents appear to account for the majority (68%) of FSM recipients. None of the carers were working in the majority (83%) of families claiming FSM while 86% of these families were renting their home. So, it is not surprising that only 9.5% of these children changed house at most once in their lifetimes compared with 71.5% of their peers from non-claimant families. The majority (56%) of FSM claimants were parents whose highest level of education was that of secondary school 16-19 years. The children of FSM claimants were also in families with three (28%) or more (34%) children.

These claimant data are consistent with the FSM eligibility criteria of non-working or very low income families with limited capital assets. Moreover, we also find that FSM-claimant families are not small and the level of education for the carer in these families is mostly low.

The Nature of Economic Deprivation among Low Income Families

In Table 3 we present the incidence of three other economic indicators in this sample, namely: history of FSM eligibility, receipt of working tax-credit and home ownership.

It should be noted that for a significant proportion subjects (13.2%, n=218) in this sample the history of FSM eligibility was unknown from their Hampshire LEA records.

Table 3: Incidence of economic deprivation measured by four economic Indicators

Nearly half (46%, n=763) of the families in this sample were in receipt of either FSM and/or working tax-credit or did not own their home. Among these, the most prevalent (19%, n=312) pattern of economic disadvantage as recorded from the three indices (i.e. FSM, working tax-credit and home renting) consists of people who only receive working tax-credit but don't claim FSM and own their home[6] (Table 4). The majority of these families, however, did not own their home and were either in low paid jobs or FSM eligible.

Table 4: Distribution of the most prevalent patterns of economic disadvantage according to three economic indicators: FSM eligibility, home renting and receipt of working tax credit.

Among those who were currently FSM eligible, 44% (n=55) had been also FSM eligible in the past three years and another 27% (n=33) were in and out of FSM eligibility in the previous three years. In other words the majority of those currently in FSM have also a history of FSM eligibility in the recent past.

[6] Renting on its own is not necessarily a measure of economic deprivation but it does imply a lack of wealth accumulated through home ownership. In this sample, as the tables above show, renting is most likely to be an indicator of disadvantage when linked to other indicators such as FSM or working tax credit.

It is also worth noting that among non-working or part-time working lone-parent families (n=167, 10.1%) with no capital assets i.e. renting their home, a significant proportion 35.3% were not observed to be FSM-eligible over the previous four-year period. In other words FSM eligibility data did not identify a significant proportion of low income families. This group of families is part of a larger pool (20.5%) of low-ranking families in terms of occupation among which 31% are not in receipt of benefits such as working tax-credit or FSM. The rest will be either in low income (30% in receipt of working tax-credit) or FSM eligible (39%). Thus, it seems that FSM-eligibility is a very coarse index of economic disadvantage with a moderate share of 39% in the population of low income families with low capital assets in the Greenwood area. Moreover, among families with no history of FSM eligibility, a significant proportion (10%), relative to FSM-eligibility prevalence, are low income families with no capital assets.

In summary, the error associated with measures of economic disadvantage when using FSM-eligibility data can be large. This is because such data measure extreme economic disadvantage and therefore can be unstable. The processes giving rise to these errors can be complex. These could be due to social processes underlying child poverty in flexible labour markets combined with data collection procedures. For example, the Pupil Annual School Census data records (PLASC) FSM eligibility as registered by the parent. Parents might not know about their entitlement or might not be willing to register it for a variety of reasons including shame or concerns related to the nutritional quality of the meal (Storey & Chamberlin, 2001).

Moreover, at the time of the study, the PLASC administrative records were updated at fixed times yearly and thus might not capture a family's change of FSM status. We also recorded the FSM take-up of the family as reported by the caregiver(s) and when

compared with the PLASC data we found discrepancies. The PLASC data suggested that there were 114 FSM-eligible families while there were 124 parents reporting FSM take-up. Twenty eight (23%) out of these 124 FSM claimants were not identified as eligible from the PLASC data which could be due to administrative error associated with the frequency of PLASC data updating. Half of these pupils had either changed schools or/and had a history of FSM entitlement. Thus, is also reasonable to assume that such errors are more likely to happen to those pupils changing schools and more importantly among pupils with a history of FSM eligibility. In fact we found that as our Hampshire-wide cohort moved from reception to KS1 tests, the proportion of pupils moving schools among non-FSM claimants was 6% compared to 15% among consistent FSM claimants and 16% - 17% among the new FSM and withdrawn claimants, respectively. This confirms previous findings on the association of child poverty and turbulent circumstances (Strand & Demie, 2006). It also suggests that the size of misclassification error could depend on the level of deprivation.

There were also 18 (1.2%) pupils identified as FSM-eligible according to PLASC but were not reported by the parent, which could be due to a parental decision not to take-up their eligibility. This seems to be an underestimate for errors associated with take-up of true eligibility according to other research where 13% of pupils qualified for FSM but only 3% were actually registered (Storey & Chamberlin, 2001). A reason why our sample estimate of misclassification error might underestimate the true extent of such errors could be associated with poverty related non-response. If we consider that the non-responders in our Greenwood sample are twice more likely to be FSM eligible then our sample estimates of FSM, take-up would be underestimated.

The amount of time spent under FSM-eligibility is associated with a future of low income. The proportion of the currently low income families with no capital assets

who had spent the previous one, two and three years under FSM-eligibility was 51%, 73% and 74%, respectively. We now consider how to identify those who could also be considered deprived but do not consistently fall below the set thresholds for benefits by looking at the longitudinal patterns of FSM eligibility using all the data in Hampshire concerning this cohort.

The stability of FSM eligibility

Given the nature of the labour market in Britain and especially with respect to single mothers, the question is to what extent we can consider the vagaries of the labour market to be such that the fortunes of low income and especially one parent families may change from one year to another, either because of family break up and/or because in a flexible labour market (Brown et al., 2001) it is easy to hire and fire workers, especially the low skilled. From the available figures (Table 5) it would appear that the percentage of pupils eligible for FSM remains relatively stable over time but assuming that this is a consequence of individuals' FSM status remaining stable can lead to erroneous analyses and gross underestimation of the extent of deprivation in the population. Table 5 shows the FSM status over the years of the cohort of all Hampshire children from which our Greenwood sample is drawn. These data were extracted from the PLASC data base. According to PLASC 2001/2, the size of the Hampshire-wide cohort of pupils at Reception year at 2001/2 is 14329. According to PLASC 2003/4, the size of Hampshire-wide cohort of pupils at year 2 at 2003/4 is 14308. However, we have test results and complete follow-up from 2002 - 2005 for 85% of this cohort. Further data inconsistencies related to correct identification of pupils and excluding schools which merged or closed reduces our Hampshire sample to 11702 pupils.

Table 5: FSM eligibility over time for the HARPS cohort (N=11,702)

Examinations of follow-up losses of previous cohorts of children progressing from baseline tests at reception year to KS1 tests from 1996-2004 and according to Hampshire maintained data bases suggested that the losses observed for our cohort are quite typical for the Hampshire County. These losses are consistent with the proportion of internal migrants in Hampshire (11%) according to the 2001 census (Hampshire, 2006).

Examination of FSM eligibility over time (Table 5) suggests that there is a substantial change in individual FSM eligibility over this 5-year period. Although the yearly average remains relatively constant at about 9%, almost 15% was actually FSM eligible over this period. There is an implication that there is a large pool of disadvantaged families whose children move in and out of FSM eligibility, as the labour market for single parents might suggest**.** This would also appear to signal no great change in real economic fortunes, since low income parents move in and out of poverty within a low income tunnel and therefore the level of disadvantage with respect to FSM over this period may be underestimated by as much as 70%. When we look closer at Table 6 we see that once eligible for FSM the most prevalent pattern of longitudinal eligibility consists of families who were eligible all the time. In other words once eligible for FSM it is possible to move out but this is not typically the case as is shown in Table 6**.**

Every year from 2003 – 2005 there are around 2% of new cases of FSM claimants (cells shaded grey: column totals for grey shaded cells range from 2.3% in 2003 to 1.6% in 2005). These new cases represent 6% of the Hamsphire-wide 2002 cohort of 5-6 year olds over this four year period. Among these families, the majority (61%) will remain FSM eligible until the end of that period (blue-shaded rows added

together represent 3.6% of the total). In other words, the next most prevalent pattern concerns families entering FSM eligibility status.

Furthermore, FSM eligible families at the beginning of this period who subsequently are observed to withdraw from receiving FSM comprise 4.6% of the cohort. Among these families only 37% (light-red shadowed cell, totalling 1.66% of the cohort) seem to have taken up FSM and then withdraw. If we look at the ratios of the proportion of FSM claimants over non-claimants calculated annually from the pool of vulnerable families who could have been claiming FSM at any point over the four year period we find that the probability that a family is eligible if it has been identified as eligible at any point during this four year period ranges between 54% to 66%.

Table 6: Observed distinct pattern of FSM eligibility over time for the HARPS cohort (N=11,702)

It is also worth noting that every year 2% - 3% change their FSM eligibility status from eligible to non-eligible. At the same time the proportion of those exiting FSM eligibility status consistently (1.66%) is similar to those entering FSM eligibility status (1.57%) after being non-eligible for the rest of the period. Moreover, if we look at the ratios of the proportion of FSM claimants over non-claimants calculated annually from the pool of vulnerable families who could have been claiming FSM at any point over the four year period and excluding those who are consistently eligible, we find that for half of the time – the first two years - they are less likely to be eligible for a single year with a probability of approximately 45% and for the other half more likely with a probability of approximately 55%. These observations again suggest that the population of FSM claimants is volatile and oscillates between poverty states which could equally include or exclude the income thresholds defined by the FSM-eligibility criteria.

In summary, these patterns suggest that this 15% pool of families who have become FSM eligible at any-point during this four-year period consists of children vulnerable to long-term poverty and a large proportion of these may not be identified from single year estimates.

The consequences of these longitudinal patterns in the take up of FSM are significant in judgements made about schools either through official data on school performance or through studies of school effectiveness, because what this analysis suggests is that the levels of economic deprivation in a school could be significantly underestimated.

However, there are other concerns regarding FSM as a measure of economic disadvantage. In the next section, we examine the consequences of underestimating the true extent of deprivation in the context of value-added (VA) analysis of school performance.

The consequences of measurement error in value-added analysis

We now investigate the effect of measurement error on a value-added analysis. The analytic methods used for measurement error adjustment in the sensitivity analysis that follows are based on the measurement error models developed by Goldstein et al. (2007). Each of these analyses is a repeated analysis of the same basic regression model (Table 7), allowing each time for different levels of reliability for FSM eligibility. In this sensitivity analysis we also included scenarios that allowed for measurement error in the tests results (Table 8).

The value-added (VA) model for mathematics attainment at KS1

This basic analysis that ignores measurement error, models the effects on test performance at KS1 for mathematics, of a number of factors most of which but not all, are typically used. These include gender, tests in mathematics and literacy at the beginning of reception year and special education needs (SEN) at KS1. Test scores

scales at both KS1 and baseline were normalized. We also take into account FSM eligibility status at both baseline and at KS1, by including appropriate regression terms. These terms quantify the effects of FSM-eligibility at baseline and separate gradients for FSM-eligibility at the KS1-year for the groups which were and were not FSM eligible at reception. Finally, our predictor's list includes a categorical variable representing low-income groups based on data on occupation rankings, receipt of working tax-credit, renting and family employment (Table 7).

It should be noted that the FSM eligible children had significantly lower baseline scores in mathematics, with mean difference adjusted for sex and special education needs of 0.6 standard deviations (95% Confidence Interval = [0.4 0.8]). The results of this VA analysis for the Greenwood sub-sample revealed some surprising results. Our analysis of their progress in mathematics, suggests that conditional on these baseline scores these FSM eligible children make significantly more progress compared with their peers who were not FSM eligible at baseline. These positive effects are additional to those of low income status at KS1. The poorest progress was made by children who were new FSM eligibility cases or whose families were in low incomes and were renting their homes or were not in full employment. It should be noted, that for the purposes of the current exposition we limited the list of predictors to the most important ones. More extensive analysis revealed that there were also significant interactions of gender and FSM eligibility status at baseline with subsequent SEN. These suggest that both gender and baseline FSM entitlement differences in KS1 progress in mathematics are reduced according to the degree of special education needs. The Hampshire-wide data suggest a strong relationship between SEN status and poverty as well as between SEN status and gender. The prevalence of SEN among those without FSM entitlement at baseline was 18% whereas among those

with FSM entitlement this rises to 40%. The prevalence of SEN among boys was 26% whereas among girls was 13%. There are measurement error issues surrounding the register of SEN in schools which is judged by teachers with reference to achievement levels in their schools (Croll, 2002). The extent of these errors was not possible to assess with the data at hand. Assessment of the reliability of the SEN register is further complicated with changes on the coding schemes of the degree and type of such needs. This is the reason why these are not taken into account in this analysis and we choose to present a sensitivity analysis under a number of conservative "what if" scenarios.

Variation between schools accounts for 14% of the total variability in the KS1 test scores in mathematics in this sample.

Sensitivity Analysis

For the sensitivity analysis, we initially made a rough estimate of misclassification probabilities from examination of the patterns on FSM status over time (Table 6). We first excluded those patterns which could be defended as consistent for example always eligible and always ineligible. Those with three years of eligibility were considered to contribute to the false positives i.e. unidentified FSM-eligibility cases. Those with only a year of eligibility were considered to contribute to the false negatives i.e. observed FSM eligible who are truly withdrawn. Those with two years of eligibility were considered to contribute equally to both the false positives and the false negatives and these proportions were weighted by a half. Then the estimates of Probability(classed as eligible when actually not eligible) and Probability(classed as ineligible when truly eligible) were calculated by weighting the observed proportions by the proportion of time over the four years spent in eligibility or ineligibility. These were then summed and divided by the total considered misclassified plus those always

eligible or always non-eligible to give the Probability(classed as eligible when actually not eligible) and Probability(classed as ineligible when truly eligible) respectively. This resulted in values for the Probability(FSM when truly ineligible)= 0.02 and Probability(not FSM when truly eligible)= 0.26. In order to examine the sensitivity to these choices we chose alternative values of 0.6 and 0.02 respectively and further allowed for measurement error in the baseline measurements.

The value of 0.6 is a maximum value based on the comparison of approximately 8% FSM in any one year as against 15% in FSM at some time over the four year period as mentioned previously. This estimate of the error arising from failing to identify FSM-eligible cases is consistent with yearly estimated false negative rates. These rates were produced when we tested how well future FSM-status can be predicted from the current one. The false negative rates ranged between 22% – 27% while the corresponding false positive rates ranged between 1.7% - 3.1%.

The later misclassification probabilities, i.e. Probability(of observing a pupil as being FSM eligible when not actually eligible) were also consistent with the estimate derived from the Greenwood sub sample. Among the responders in the Greenwood area who took-up FSM, 1.2% were not eligible according to PLASC. It is reasonable to assume that the observed discrepancies giving rise to this latter type of misclassification error are due to the time-lag between administrative record-updating and study data collection and are not likely to vary much in areas of low pupil mobility (Goldstein et al., 2006). What is also worth considering is that these types of errors are more likely to appear among those who have been FSM-eligible in the past.

Results of sensitivity analysis

The results in Table 8 suggest that ignoring the measurement error in the predictors could significantly change the associated effect estimates and their standard errors and

thus our inferences. The effect estimates of the covariates with no measurement error are not affected. Adjusting for measurement error in the continuous predictors, also leads to reductions in the level 1 variance and between school variation increases by up to 3% of the total variability in test scores.

In this analysis we start by introducing measurement error in the FSM eligibility at KS1 for the group of children who represent new FSM-eligibility cases for two reasons. The first is concerned with the findings from the standard VA analysis which suggested that this is an important predictor. The second is related to the finding that the unreliability of FSM status is likely to be higher in this group rather than the group of those who consistently meet the eligibility criteria.

We find that increase in the proportion of unidentified FSM eligibility cases weakens the associated effect. However, the changes induced by this type of error alone, are small. The latter is not surprising since the counts affected by such an error would be low as a result of the low prevalence of FSM eligibility. If combined with measurement errors in the tests, however, large increases of this type of error can lead to significant reductions (25% change) in the effect estimates.

On the other hand, the effect estimate associated with FSM is found to be sensitive to even small increases in the proportion of false eligible cases. Adjustment of this type of error results in significant increases up to 59% change in the resulting estimates among the scenarios considered as well as striking increases in the associated standard errors. These changes could be attenuated when measurement error in the tests is also accounted for. These latter changes tend to be smaller when we relax assumptions of independence of the measurement errors of the test scores. Another consequence of introducing measurement error in the test scores and the baseline tests especially,

relates to further increases in the standard errors associated the effect estimate of FSM entitlement. Allowing for measurement error in the response leads to similar changes.

In this analysis we have assumed that the measurement error in baseline tests is independent of misclassification in FSM since these data are assessed by different agents, i.e. the teachers and the Local Education Authorities, respectively.

Table 7: The effect of measurement error on effect estimates on Value-Added analysis of performance at KS1 tests on Maths.

Table 8: The effect of measurement error on effect estimates on Value-Added analysis of performance at KS1 Maths tests.

Discussion

The research reported in this paper examines the longitudinal patterns of FSM eligibility over time for the cohort of all Year 3 primary school pupils at 2004/2005 in Hampshire. We observed high levels of individual fluctuation in FSM status over time which renders FSM an unreliable index of economic deprivation. Closer examination of such volatility using other indices of SES collected from the Greenwood area revealed associations with low income and education level, and turbulent family circumstances as reflected by family structure and home and school changes.

We used this data to assess the magnitude of error that can be introduced in estimates of the prevalence of economic disadvantage in this population when FSM is used to index it. We found that FSM is both a coarse and unstable instrument and that single year snapshots of FSM eligibility underestimates the pool of disadvantaged by 70%.

Entitlement to FSM is a crude measure of socio-economic circumstances. We saw that the income cut-off imposed will characterise a significant proportion (61%) of low-income families with low-capital assets as “non-disadvantaged”. The “non-disadvantaged” families which are close to the threshold will then be averaged with those from more privileged backgrounds, driving the mean of the truly non-

disadvantaged towards lower values. The resulting comparisons between the groups formed in this way will lead to estimates of difference which are smaller. In fact, our VA analysis (Table 7) suggests that this low-income group is very similar in terms of progress in mathematics, to those eligible for FSM. There is a need for more fine-grained measures for economic circumstances in order to explain differences in attainment more accurately. This finding has profound implications for policy because it suggests that children from low income families, regardless of whether they are eligible for FSM, under perform at school. Given the government's emphasis on taking children out of poverty through mechanisms such at the WTC this finding casts doubt on the implications of such a policy for educational achievement. Indeed, it suggests a broader strategy which is much better resourced such as in the Nordic countries may be required (Esping-Andersen, 2006)

In order to understand the direction of bias that could be expected according to increases of the two different types of misclassification errors, consider the following. Intuitively, correction for increases in the misclassification probability associated with the unidentified FSM cases is equivalent to moving the associated income eligibility cut-off towards lower values. This will in turn weaken the effect of FSM entitlement.

The changes that Welfare policies usually impose however, involve moving these cut-offs towards higher incomes. So, in practice this error is mostly due to low FSM take-up rates and these have been discussed elsewhere (Storey & Chamberlin, 2001). We found that adjusting for this type of error leads to the expected decrease in the effect estimate of FSM entitlement. This type of error can be large. In fact, if we also allow for high levels of this type of error in the estimation of the effect of FSM eligibility at baseline, it no longer appears to have an impact on pupils' progress. In other words, ignoring this type of error could lead to overestimating the progress of pupils with

very poor backgrounds early in life. However, the size of the bias introduced, is fairly insensitive to large increases of its value. Thus, ignoring the measurement error due to low take-up rates will not grossly over-estimate the effect of FSM entitlement when balanced with its lack of sensitivity to capture disadvantaged pupils who are close but do not meet the eligibility criteria. This finding simply confirms that if we adjust for this type of error, FSM eligibility proves to be a very poor indicator of poverty. If FSM eligibility continues to be used as a proxy then efforts needs to be made to ascertain the take-up rates in schools and action needs to be taken to improve take-up rates in schools.

The second type of error which is associated with the proportion of false FSM eligibility seems to be more important. Adjustment to even small increases in this type of error, increase the effect of FSM entitlement considerably. Another striking effect of adjusting for this type of error, relates to the two-fold increase in the standard errors associated the effect estimate of FSM entitlement. It is important to consider the process which could give rise to this error. Inconsistent updating of the administrative records of FSM entitlement can be difficult when pupils move schools or when schools respond poorly with their PLASC returns. It can most likely occur to those pupils who have a history of FSM entitlement and remain in the poverty tunnel surrounding the income eligibility cut-offs.

This second type of error is important for another reason. In the VA analysis we saw that indicators of poverty are strong predictors of poor progress in mathematics test scores at KS1. Change in FSM entitlement is not usually considered in most VA analysis and we saw that this is important when the reliability of a blunt indicator of poverty such as FSM entitlement is also poor. We found that children with poor backgrounds i.e. FSM eligible at the beginning of this period, have lower baseline

scores, but progress significantly better. They can catch-up. These effects however are cancelled by subsequent poverty. The level of poverty during the KS1 year, seem to be important in explaining differences in attainment. In this comparison children from low income families with low capital assets who do not meet below FSM eligibility criteria do not seem to fare better in their progress in mathematics at KS1 when compared with new FSM eligible cases.

We also examined how these changes in the effect estimates could be affected by likely errors in the baseline test scores. Even under a conservative scenario where moderate levels of measurement errors of both types errors is considered along with relatively high levels of unreliability for the baseline test scores, there will be a 38% underestimation of the effect of FSM entitlement. When measurement error in the KS1 test scores is also assumed there is still 13% underestimation.

Methods for correcting for the bias introduced by measures of SES with poor validity i.e. poor indicators of what is of real interest that is poverty exist e.g. instrumental variable or latent variable approaches, but depend on the use of richer sources of SES data which are not usually available. The construction of indices of poverty using such sources along with sophisticated statistical methodology is widely considered to be beneficial and convenient as it combines information from a multitude of sources in a single measure. However, the fact that the resulting index, is an estimate and thus subject to uncertainty and error is rarely taken into account in subsequent analysis. Accounting for measurement error in a VA added analysis is not a trivial task especially when this error is differential e.g. depends on the true levels of the measure it approximates. Development of such indices usually results in values which are strongly dependent on their associated standard errors. Failing to account for the

differential nature of measurement error in subsequent VA will result in equally poor inferences and more importantly predictions.

In this work, we do not address issues associated with the validity of FSM as a measure, directly. Instead we address the bias by studying its reliability. It is the income interval around the FSM entitlement cut-off where the bias is likely to be meaningfully and more precisely maximised. This is the same interval where FSM would be mostly unreliable. We argue that the unreliability of FSM is intrinsic to the extremity of the economic disadvantage it measures and thus bound to be unstable in the light of a flexible British Labour Market.

In summary, we found that ignoring the uncertainty associated with FSM eligibility can lead to biased inferences on the effect of FSM on pupil's academic progress and inflated optimism for the associated standard error estimates which in turn can lead to incorrect inferences. If FSM entitlement continues to be used in VA analysis, it is important to also account for the change in FSM eligibility status. Adjustment for the measurement error associated with FSM eligibility counts in a value-added analysis also seem to be important, although the size of likely errors is difficult to ascertain. Further work needs to be done on the elicitation of these errors on a large scale in order to understand the processes causing them. These independent error estimates could then be used to correct FSM-entitlement estimates in VA analysis.

Our error estimates were based solely on assessments of the reliability of FSM from the small number of repeated measurements covering the period between reception and KS1 tests. The instability of FSM eligibility over time, however, only reflects one aspect of deprivation predominantly related to family unemployment and lone-parenthood. In fact, Vignoles (2006) reports that these latter components of deprivation account for only 18% of the FSM-gap in KS1 attainment in mathematics,

using longitudinal data from the ALSPAC study and other factors such as family income and maternal education level account are far more informative. More fine grained indicators of poverty which combine FSM eligibility with other indicators such as working tax credit are needed in order to more reliably assess the effect of socio-economic circumstances on pupil's academic progress, especially during early phases of schooling.

In conclusion, FSM eligibility is not just a coarse indicator of socio-economic disadvantage but is also unreliable. As a result, it will underestimate the pool of disadvantaged considerably. This in turn can also bias the effect of SES in standard value-added analyses. It underestimates the effect of poverty on the progress in mathematics of children in families living below extremely low income thresholds during the year of their KS1 tests. Moreover, this progress for children from very poor backgrounds early in life could also be overestimated in schools with low FSM take up rates.

Finally, and most importantly these findings raise important questions about the way progress in schools is 'officially' measured and raises doubts about the trust that is invested in FSM as a reliable indicator of deprivation. It also raises questions about the estimates of school effects based on models where FSM entitlement is used as a measure of disadvantage.

Acknowledgments:

This work was supported by the ESRC grant Reference No.: RES-000-23-0784. Special thanks are due to the Hampshire Children's services and especially Nigel Hill and Paula Guy for all their support. We are deeply indebted to all the Year 3 school teachers and school heads in Greenwood and Eddie Izzard in particular. Their support and collaboration made this project possible. Special thanks are also due to Ceri

Brown and Martin Thrupp who designed and organized the collection of the original Greenwood data. We would also like to thank Ruth Lupton for her perceptive comments on a draft of this paper.

References

Allen, R. & Vignoles, A. (2006) What should an index of school segregation measure? *Centre for Economics of Education*.

Bradbury, B., Jenkins, P. & Micklewright, J. (2001) *The Dynamics of Child Poverty in Industrialised Countries* (Cambridge, Cambridge University Press).

Brown, P., Green, A. & Lauder, H. (2001) *High skills: Globalization, Competitiveness and Skill Formation* (Oxford, Oxford University Press).

Croll, P. (2002) Social deprivation, school level achievement and special education needs., *Educational Research*, 44(1), pp. 43-53.

Croxford, L. (2000) Is free school entitlement a valid measure of school intake characteristics, *Educational Research and evaluation*, 6(4), pp. 317-335.

Daly, A., McNamara, J., Tanton, R., Harding, A. & Yap, M. (2006) Indicators of social exclusion for Australia's children: An analysis by state and age group.*Social Reseach Centre Conference* (Queensland University,

DFES (2003) Pupil progress by pupil characteristics: 2002, *Statistics of Education, London, Department of Education and Skills*.

DFES (2005a) Youth Cohort study: The activities and experiences of 16 Years Olds: England and Wales 2004, *Statistical First Release. London, Department of Education and Skills*.

DFES (2005b) National Curriculum Assessment, GCSE and Equivalent Attainment and Post-16 Attainment by Pupil characteristics in England 2004, *Statistical First Release. London, Department of Education and Skills*.

Esping-Andersen, G. (2006) Social Inheritance and Equality of Opportunity Policies, in: H. Lauder, P. Brown, J. Dillabough & A.H. Halsey (Eds) *Education, Globalization and Social Change.* (Oxford, Oxford University Press).

Goldstein, H. (1997) Methods in school effectiveness research, *School effectiveness and School Improvement*, 8, pp. 369-395.

Goldstein, H., Burgess, S. & McConnell, B. (2006) Modelling the impact of pupil mobility on school differences in educational achievement*CPMO Working Paper No. 06/156*

Goldstein, H., Kounali, D. & Robinson, A. (2007) Modelling Measurement errors, *SMIJ, in press*.

Goldstein, H. & Noden, P. (2003) Modelling social segregation, *Oxford Review of Education*, 29, pp. 225-237.

Goldthorpe, J.H. & Hope, K. (1974) *The social grading of occupations: A new approach and scale* (Oxford, Clarendon Press - Oxford).

Gorard, J., Taylor, C. & Fitz, J. (2003) *School, Markets and Choice Policies* (London, Routledge Falmer).

Goyder, J., Warriner, K. & Miller, S. (2002) Evaluating socio-economic status (SES) bias in survey nonresponse, *Journal of Official Statistics, Statistics Sweden*, 18(1), pp. 1-11

Gregg, P. & Wadsworth, J. (2003) Workless Households and Recovery, in: R. Dickens, P. Gregg & J. Wadsworth (Eds) *The Labour Market Under New Labour* (Basingstoke, Palgrave).

Hampshire, C.C. (2006) Census 2001 results for the Hampshire area in: Environment (Ed)

Hutchison, D. (2003) The effect of group-level influences on pupils' progress in reading, *British Educational Research Journal*, 29(1), pp. 25-40.

Noble, M., Wright, G., Dibben, C., Smith, G., McLennan, D., Anttila, C., Barnes, H., Mokhtar, C., Noble, S., Avenell, D., Gardner, J., Govizzi, I. & Lloyd, M. (2004) Indices of Deprivation 2004(London, Deputy Prime Minister Office: Neighbourhood Renewal Unit).

Plewis, I. & Goldstein, H. (1997) The 1997 Education White Paper - A failure of standards, *British Journal of Curriculum and Assessment*, 8(1), pp. 17-20.

Sammons, P., West, A. & Hind, A. (1997) Accounting for variations in pupil attainment at the end of Key Stage 1, *British Educational Research Journal*, 23(4), pp. 489-511.

Schagen, I. & Schagen, D. (2005) Combining multilevel analysis with national value-added data sets - a case study to explore the effects of school diversity, *British Educational Research Journal*, 31(3), pp. 309-328.

Storey, P. & Chamberlin, R. (2001) Improving the take-up of free school mealsInstitute of Education).

Strand, S. (1997) Pupil progress during Key Stage 1: a value-added analysis of school effects, *British Educational Research Journal*, 23(4), pp. 471-487.

Strand, S. & Demie, F. (2006) Pupil mobility, attainment and progress in primary school, *British Educational Research Journal*, 32(4), pp. 552-568.

Yang, M., Goldstein, H., Rath, T. & Hill, N. (1999) Using value-added analysis for school improvement purposes, *Oxford Review of Education*, 25, pp. 469-183.

List of Tables:

Table 1: Parental education attainment and family SES

SES (⅄) Counts	Missing	Secondary <16 years	Secondary 16 – 19 years	Further and Vocational qualifications	University graduates and postgraduates	Total
High	1	26	17	73	130	247 (14.94)
Middle	2	110	95	247	139	593 (35.87)
Low	8	236	90	253	37	624 (37.75)
Unknown	12	85	28	54	10	189 (11.43)
Total	23	457	230	627	316	1653
(%)	(1.39)	(27.65)	(13.91)	(37.93)	(19.12)	

⅄: This is the occupation of the male carer when this is present. The Goldthorpe scale was used to rank occupational categories

Table 2: The association of family SES with employment, lone parenthood and family size.

Family employment status SES (*) Counts	None work	Only one Part-time	At least one full-time	Both full-time	Lone parenthood	No. of families with >=3 children
High	-	7	201	39	20	97
Middle	3	17	486	87	46	233
Low	7	54	490	73	110	275
Unknown	131	13	41	4	108	106
Total	141	91	1218	203	284	711
	(8.53%)	(5.51%)	(73.68%)	(12.28%)	(17.18%)	(43.01%)

Table 3: Incidence of economic deprivation measured by four economic Indicators

	N	%
Known History of FSM eligibility	158	9.5
Renting	434	26.3
Receipt of working tax-credit	483	29.2

Table 4: Distribution of the most prevalent patterns of economic disadvantage according to three economic indicators: FSM eligibility, home renting and receipt of working tax credit (*).

	FSM eligibility	Home Rent	Receipt of Working Tax-credit	N	%
Pattern 1			X	312	40.9
Pattern 2		X		169	22.1
Pattern 3		X	X	158	20.7
Pattern 4	X	X		92	12.5
Total				763	

(*): X denotes the presence of the attribute

Table 5: FSM eligibility over time for the HARPS cohort (N=11,702).

FSM	2002	2003	2004	2005
Eligible	1,022	1,009	1,135	1,042
%	(8.73)	(8.62)	(9.70)	(8.90)

Number of years eligible	N	%
0	9,987	85.34
1	541	4.62
2	339	2.90
3	351	3.00
4	484	4.14

Table 6: Observed distinct pattern of FSM eligibility over time for the HARPS cohort (N=11,702).

		Pattern (*)			
Count	%	2002	2003	2004	2005
484	4.14	X	X	X	X
141	1.20	X	X	X	-
21	0.18	X	X	-	X
91	0.78	X	X	-	-
51	0.44	X	-	X	X
20	0.17	X	-	X	-
20	0.17	X	-	-	X
194	1.66	X	-	-	-
138	1.18	-	X	X	X
64	0.55	-	X	X	-
11	0.09	-	X	-	X
59	0.50	-	X	-	-
133	1.14	-	-	X	X
104	0.89	-	-	X	-
184	1.57	-	-	-	X
9,987	85.34	-	-	-	-

(*) X represents FSM eligibility and (-) FSM non-eligibility

Table 7: Test performance in mathematics at KS1 – Value Added Analysis

Predictors	Mean	Standard Error	95% CI (*)	
Baseline mathematics	0.46	0.03	[0.40	0.52]
Sex - male	0.23	0.04	[0.15	0.30]
Baseline literacy	0.14	0.03	[0.08	0.20]
SEN status at KS1 ⊥				
Mild	-0.46	0.07	[-0.59	-0.34]
Severe	-0.70	0.10	[-0.90	-0.50]
Income group ‡				
Group 2	-0.16	0.04	[-0.24	-0.08]
Group 3	-0.25	0.08	[-0.41	-0.10]
FSM eligibility at baseline	0.24	0.12	[0.02	0.47]
FSM eligibility at KS1 for the group NOT FSM eligible at baseline	-0.32	0.12	[-0.56	-0.10]
FSM eligible at baseline	-0.17	0.15	[-0.45	0.12]

95% CI (*) : 95% Confidence Interval

⊥ : categorical variable with reference category the group with No SEN

‡ : categorical variable with reference category those who are not low income as judged by either Receipt of working tax-credit or low ranking occupations. Group 3 represents those in low income who were also burdened by rent or were not in full time employment.

Table 8: The effect of measurement error on effect estimates on Value-Added analysis of performance at KS1 Maths tests

Measurement error Scenario	FSM at KS1	Baseline Maths	Baseline Literacy	Level 2	Level 1 Variance
(†) $P(0\mid 1)$ (‡) $P(1\mid 0)$	Mean (SE)	Mean (SE)	Mean (SE)	Mean (SE)	Mean (SE)
(*) R, (**) ρ, (°) R_y					
$P(0\mid 1)$=0%, $P(1\mid 0)$=0% R=1, ρ=0, R_y=1	-0.32 (0.12)	0.46 (0.03)	0.14 (0.03)	0.08 (0.02)	0.49 (0.02)
$P(0\mid 1)$=26%, $P(1\mid 0)$=0% R=1, ρ=0, R_y=1	-0.31 (0.13)			0.08 (0.02)	0.49 (0.02)
$P(0\mid 1)$=60%, $P(1\mid 0)$=0% R=1, ρ=0, R_y=1	-0.30 (0.12)			0.08 (0.02)	0.49 (0.02)
$P(0\mid 1)$=80%, $P(1\mid 0)$=0% R=1, ρ=0, R_y=1	-0.28 (0.12)			0.08 (0.02)	0.49 (0.02)
$P(0\mid 1)$=26%, $P(1\mid 0)$=1% R=1, ρ=0, R_y=1	-0.42 (0.18)			0.08 (0.02)	0.49 (0.02)
$P(0\mid 1)$=60%, $P(1\mid 0)$=1% R=1, ρ=0, R_y=1	-0.39 (0.17)			0.08 (0.02)	0.49 (0.02)
$P(0\mid 1)$=26%, $P(1\mid 0)$=2% R=1, ρ=0, R_y=1	-0.49 (0.24)			0.08 (0.02)	0.49 (0.02)
$P(0\mid 1)$=60%, $P(1\mid 0)$=2% R=1, ρ=0, R_y=1	-0.45 (0.24)			0.08 (0.02)	0.49 (0.02)
$P(0\mid 1)$=26%, $P(1\mid 0)$=2% R=0.9, ρ=0, R_y=1	-0.51 (0.30)	0.61 (0.05)	0.04 (0.05)	0.08 (0.02)	0.46 (0.02)
$P(0\mid 1)$=26%, $P(1\mid 0)$=2% R=0.9, ρ=0.5, R_y=1	-0.41 (0.31)	0.54 (0.04)	0.13 (0.04)	0.08 (0.02)	0.46 (0.02)
$P(0\mid 1)$=26%, $P(1\mid 0)$=2% R=0.8, ρ=0.5, R_y=1	-0.44 (0.23)	0.65 (0.05)	0.08 (0.06)	0.08 (0.02)	0.42 (0.02)
$P(0\mid 1)$=26%, $P(1\mid 0)$=2% R=0.8, ρ=0.5, R_y=0.9	-0.36 (0.23)	0.60 (0.05)	0.10 (0.05)	0.07 (0.02)	0.39 (0.02)
$P(0\mid 1)$=60%, $P(1\mid 0)$=0% R=0.8, ρ=0.5, R_y=0.9 (∇)	-0.24 (0.12)	0.65 (0.06)	0.08 (0.06)	0.08 (0.02)	0.42 (0.02)
$P(0\mid 1)$=60%, $P(1\mid 0)$=2% R=0.8, ρ=0.5, R_y=0.9	-0.34 (0.21)	0.65 (0.06)	0.08 (0.06)	0.08 (0.02)	0.42 (0.02)

†: $P(0\mid 1)$ denotes the misclassification Probability of observing a pupil as not being FSM eligible when he is actually eligible

‡ $P(1\mid 0)$ denotes the misclassification Probability of observing a pupil as being FSM eligible when he is not actually eligible.

(*) R=1 denotes the Reliability of the baseline tests; the reliability is assumed to be the same for both tests

(**) ρ=0 denotes the correlation between the measurements errors for the baseline tests

(°) R_y=1 denotes the reliability of the outcome i.e. KS1 test scores in mathematics

(∇) Introducing $P(0\mid 1)$=60% and $P(1\mid 0)$=0% for both FSM at baseline and KS1 modifies the mean (SE) of the corresponding effect estimates to -0.09 (0.08) and -0.20 (0.11) respectively.

List of Figures

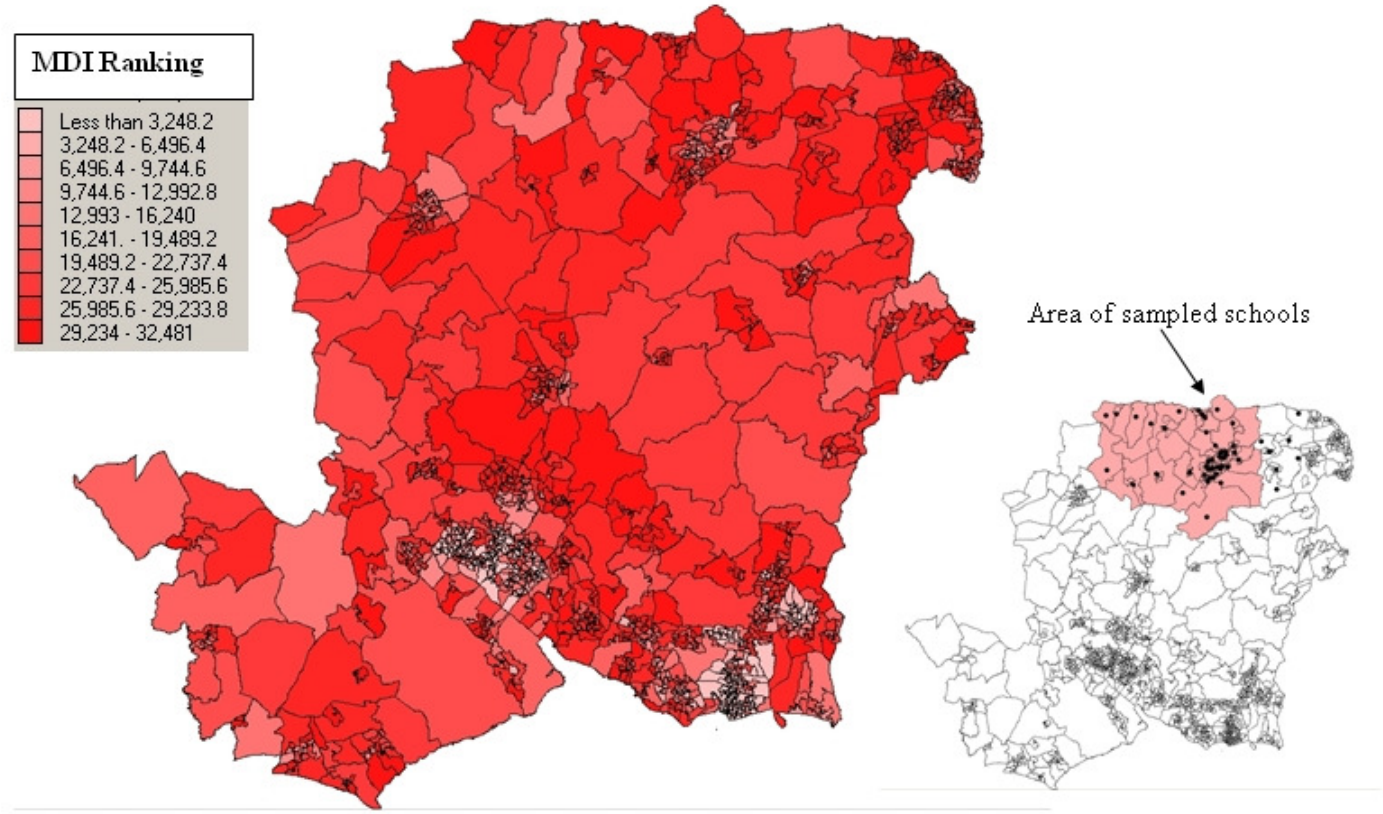

Figure 1: The deprivation geography of Hampshire based on the Multiple Deprivation Index.