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 labor 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, Bristol, UK. Email: Daphne.Kounali@bristol.ac.uk
[b] Department of Mathematical Sciences, University of Bath, Bath, UK
[c] Centre for Multilevel Modelling, Graduate School of Education, University of Bristol, Bristol, UK
[d] Department of Education, University of Bath, Bath, UK

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 labor markets; measurement error; reliability; bias; value added analysis; progress in mathematics at KS1.

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. Moreover, among policy makers it is seen as referring to a stable population of disadvantaged pupils who, in effect are depicted as a sub-set of the working class or as part of an underclass. For example, Ruth Kelly, when Minister for Education said of FSM:

> "We have no data on the social class of the parents of children in school at age 11, so we proxy social class by whether or not the pupil is in receipt of FSM. Importantly, in the absence of administrative data on the FSM status of KS2 pupils in 1998, we assume that their FSM status is the same as it was at age 16 in 2003. This is an approximation, but as FSM status is relatively stable through time it should not be too unrealistic as a means of eliciting the key trends", Rt Hon Ruth Kelly, Secretary of State for Education and Skills 'Education and Social Progress' briefing Note, 26 July, 2005. (DFES, 2005,2)

In this paper we argue that FSMs as a measure of deprivation are not only coarse but also unreliable. We provide empirical evidence that does not support the assumptions of stability of FSM eligibility status over time. Such assumptions form the basis of official statistics to support policy makers and it is clearly expressed in the statement above. The data we present here suggest that it is not clear what the group of those identified as eligible for FSMs represents in terms of disadvantage. We find that those identified as eligible for FSMs from administrative data bases at any single year are only a small section of a much larger group of disadvantaged pupils and their families. This implies that the proportion of disadvantaged in a school is higher than acknowledged. It also suggests that the population of those on FSMs is not stable and any calculation or judgement is likely to be an underestimate of the real disadvantage that a school or student confronts.

There are good theoretical reasons for believing that the circumstances of disadvantaged families may change making them eligible for FSMs at one point and not another.

While there has been some concern expressed about the measure, there has, to our knowledge, been no systematic test of its appropriateness. We find that the quality and use of official data records for education policy does not allow for adequate assessment of the nature and extent of socio-economic disadvantage. We show that the statistics currently used are a gross under-estimate of socio-economic disadvantage and that such bias also leads to under-estimation of education disadvantage.

The Structure of the Paper

The paper starts by presenting some background in the nature of flexible labour market in Britain and the distribution of welfare benefits related to it. Both impact on the nature of child poverty of which FSM is assumed to be a reliable indicator. We then provide details on what FSMs intend to measure and what is actually recorded in official databases.

We proceed by outlining the data set used to conduct our analysis. This is followed by a discussion of the methodology employed. We then analyse the sources of error in the use of FSM as a proxy for disadvantage.

We first look at the nature of the FSM population relative to a sub-section of the year 3 population of a county in England for whom detailed socio-economic data were collected.

We then examine the degree of change in FSM eligibility over time for this cohort in the whole county in order to characterize the reliability of FSM eligibility status in official records. This is followed by a discussion on the magnitude and the processes that could give rise to errors of misclassification of those who are FSM eligible and is interesting that one of these errors may well be related to pupil turbulence. This in turn occasions a discussion of the assumption that FSM eligible pupils constitute the stable core of the most deprived.

Having discussed these sources of error we examine the consequences of measurement error with respect to FSMs on a value added analysis of the effects of deprivation on numeracy at Key Stage 1 which compares the FSM measure to other variables such as occupation, receipt of working tax credit, renting and family employment in explaining KS1 outcomes for disadvantaged families in our sample.

Economic Deprivation and the Nature of the Flexible Labor 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. If FSM is to stand as a proxy indicator of disadvantage, then in the light of the above its sensitivity may be in question.

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.

The Use of FSM

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, 1997,369-395, Plewis & Goldstein, 1997,17-20, Sammons *et al.*, 1997,489-511, Yang *et al.*, 1999,469-183), studies of school composition (Hutchison, 2003,25-40, Schagen & Schagen, 2005,309-328, Strand, 1997,471-487) and research on socially-segregated schooling (Allen & Vignoles, 2007,643-668, Goldstein & Noden, 2003,225-237) and school choice (Gorard *et al.*, 2003). More directly, Local Authorities incorporate FSM figures in their calculations of extra provision for Special Educational Needs and Additional Educational Needs. The Department for children schools and families (DCSF, former Department of Education and Skills (DFES)) 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, 2000,317-335).

Eligibility Criteria

Over recent years the eligibility criteria have changed as a result of changes in benefits. This can lead to additional problems in using FSM data when investigating economic deprivation over a prolonged time period. At the time of the study the (2004) eligibility criteria were 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 as an indicator of disadvantage is based mainly upon its availability. There is no other measure reflecting individual economic disadvantage that is universally or even widely available[2].

In this paper we are primarily concerned with FSM eligibility as recorded by the Pupil Annual School Census (PLASC) and maintained by the DCSF. It is worth noting that these records do not strictly represent FSM eligibility since its recording depends on both the school and the carer's decision to claim. PLASC is statutory for all maintained, special and non-maintained special schools in England, city academies and city technology colleges (Section 537A of the Education Act 1996) and schools have to maintain and prepare their PLASC returns through their school information systems. School Information systems are not centrally controlled and vary across schools. There is no study on the quality of information maintained by the schools or the accuracy of their PLASC returns. However, recent reports by the PLUG (Pupil Annual Census/National Pupil Database of test records User Group) suggest problems in the quality and variation in the quality of PLASC returns across schools (Rosina & Downs, 2007).

Moreover, the DFES (currently DCSF) guidelines to schools on how to complete their PLASC returns on FSM eligibility status state:

> "Pupils should only be recorded as eligible if they have claimed FSMs and (1) the relevant authority has confirmed their eligibility or (2) final confirmation of eligibility is still awaited but the school has seen

---

[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.* Report for Deputy Prime Minister Office: Neighbourhood Renewal Unit (London).. 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.

> documents that strongly indicate eligibility (e.g. an Income Support order book) and on the basis of those who has commenced provision of free school meals."

So, there are issues relating to parental take up as well as how schools support their pupils' families in this process.

Methodology

In this analysis we use three data bases: the National Pupil database (NPD), PLASC and the data collected under the Hampshire Research with Primary Schools (HARPS - ESRC funded project). The NPD is a pupil level database which matches pupil and school characteristic data to pupil level attainment. PLASC is the key source of data for individual pupil characteristics which include ethnicity, FSM, information on Special Education Needs (SEN), and a history of schools attended.

The HARPS project

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.

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. Data collected included: 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. The deprivation geography of Hampshire according to the multiple deprivation index suggests that the children attending the selected Greenwood schools live in areas covering the deprivation spectrum, including pockets of particularly deprived.

Data collected on measures of disadvantage

In this paper we include three proxies for income: FSM, Working Tax Credits and Home Ownership and a measure of socio-economic status (SES) based on occupational categories ranked according to the Goldthorpe scale. Details of the SES characterization and coding from the collected data are presented in Appendix A. Families eligible for FSMs, as we have seen, do not have paid work; Working tax credits are given to families where one adult is in low paid work. In 2004, when the data on our families were collected, a couple or lone parent with one dependent child under 11 and a gross annual income of up to about £13,500 would have been eligible for WTC, although those with higher incomes would also be eligible if they were paying for childcare, or were disabled, or working more than 30 hours per week, or if they had more children. Home ownership can be seen as a form of wealth, whereas it will be seen later that renting is strongly associated with low income.

Statistical Methodology

Assessment of measurement error in FSM eligibility recorded in PLASC

Our purpose is to estimate the underlying but unobserved threshold of poverty as measured by FSM eligibility and also to estimate the dynamics of moving above and below this threshold. We use a Bayesian hierarchical hidden Markov model which specifies that changes in individual eligibility depend only on the previous eligibility status and that there are time independent probabilities for each of the four possibilities resulting from the combinations of remaining in the same eligibility status or of changing status. The probability of an FSM claim then depends only on the underlying eligibility status at the appropriate time.

Specifically, the random variable e_{it} is the hidden eligibility state at time t for individual i (e_{it} = 1, 0 denote eligible and not eligible respectively). The random variable c_{it} is the observation for individual i at time t, (c_{it} = 1, 0 denote claim and no claim respectively)

The probabilities corresponding to the four possible transitions are:

Probability(now eligible given previously eligible) = Probability($e_{it} = 1 \mid e_{it-1} = 0$)

Probability(now eligible given previously ineligible) = Probability($e_{it} = 1 \mid e_{it-1} = 1$)

Probability(now ineligible given previously eligible) = Probability($e_{it} = 0 \mid e_{it-1} = 1$)

Probability(now ineligible given previously ineligible) = Probability($e_{it} = 0 \mid e_{it-1} = 0$),

and so the second and third of these correspond to a change of status.

Then S= Probability($c_{it} = 1 \mid e_{it} = 1$) is the sensitivity or detectability of FSM claims to identify those eligible. We also assume that FSM claims as a test for FSM eligibility have perfect specificity, i.e.

Probability($c_{it} = 1 \mid e_{it} = 0$)= 0

The proposed model allows the estimation of the transition probabilities of the hidden states as well as the sensitivity of official records to detect those below the intended income thresholds.

This Hidden Markov Model (HMM) in which the observed process is the presence of an FSM claim (Figure 1) below shows the general architecture of an instantiated HMM. The arrows in the diagram denote conditional dependencies

Then Probability(c_{it} = 1 | e_{it}) = S e_{it} where S is the specificity as defined above (Kounali *et al.*, 2008). We fitted the model above using the freely-available software WinBugs (Spiegelhalter *et al.*, 2003)

<Insert Figure 1 about here >

Value-added analysis

Value-added analysis on the KS1 performance on mathematics in the Greenwood sample was performed using multilevel modelling. We fitted a variance component model using MLWin (Rasbash *et al.*, 2005).

The basic analysis models the effects on test performance at KS1 for mathematics, of a number of factors. These include gender, prior attainment 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 reported FSM eligibility status at both baseline and at KS1. These terms allow quantification of the separate effects of FSM-eligibility at baseline and those newly eligible at KS1. Our predictor list also includes a categorical variable representing low-income groups based on data on occupation rankings, receipt of working tax-credit, renting and family employment.

Accounting for measurement error in VA analysis

The effect of measurement error on the basic value-added model was investigated through sensitivity analysis. New analytic methods and software were developed to adjust for misclassification error on binary predictors and unreliability of continuous predictors. The technical details of the measurement error model are described elsewhere in detail (Goldstein *et al.*, 2008). The statistical software implementing these techniques is freely available and can be downloaded from the web-site of the Centre of Multilevel Modelling (http://www.cmm.bristol.ac.uk/research/Realcom/)

Results

The Greenwood subsample - background data

Female responders accounted for 90% of the returned questionnaires. This is a sample that is predominantly white with 92.7% of the responders being white-British or Irish, another 3.4% being white-mixed and another 3.3% all other ethnic or racial backgrounds. Table 1 depicts the distribution of FSM eligibility status according socioeconomic status and working mode as well as lone parenthood and home ownership.

[Insert Table 1 about here]

In Table 2 we summarize the distribution of FSM eligibility status according to SES and level of parental education attained.

[Insert Table 2 about here]

Out of the 1653 families, 124 (7.5%) reported that they were in receipt of FSM. We 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 overwhelming majority of those found to respond as eligible for FSMs are families where none of the carers is working (78%) and are renting their homes (86%), (Table 1). A significant proportion (73%) of these FSM eligible families consists of solo parents (Table 1). Secondary education below 16 years was the highest level of education for 53% of these families (Table 2).

Here, we need to distinguish between the parental response on FSM take-up recorded by this study and the official records of FSM-eligibility. We have already discussed the reasons why these official records can be misleading by being labelled as eligibility and have noted the close resemblance in FSM claims as reported by the parent and as recorded by PLASC (Table 3).

[Insert Table 3 about here]

These claimant data are consistent with the FSM eligibility criteria of non-working or very low income families with limited capital assets.

The Nature of Economic Deprivation among Low Income Families

In Table 3 we present three socio-economic indicators in this sample, namely: FSM eligibility (based on PLASC records), receipt of working tax-credit and home ownership. 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. Moreover, the children of those renting suffer a penalty in QCA3 test performance (Lauder *et al.*, 2008). For this reason we have included those renting as a measure of disadvantage. In particular our interest is represented by those who are either FSM or WTC eligible and are renting (patterns 3 and 4 in Table 4).

[Insert Table 4 about here]

We found that among non-working or part-time working families with no capital assets i.e. renting their home (n=167, 10.1%), a significant proportion 32.6% were not observed to be FSM-eligible according to PLASC over the previous four-year period. In other words FSM eligibility data did not identify a significant proportion of very low income families. There were 350 families who were renting their homes and the carers were either in part-time employment or working in occupations ranked among the lowest. Among these families 39% (n=137) were in receipt of WTC and 32% (n=113) were claiming FSMs. Thus, it seems that

FSMs claims is a very coarse index of economic disadvantage with a moderate share of 32% in the population of low income families with low capital assets in the Greenwood area.

Measurement error in the PLASC records of FSM-eligibility

So far we have examined the relationship between those defined as disadvantaged and their relationship to FSM eligibility. If we conceive of those eligible for FSM as part of a wider pool then we might expect a degree of mobility in and out of FSM eligibility. In the following analysis we use data extracted from the PLASC data base. According to PLASC 2001/2, the size of the Hampshire-wide cohort of pupils in Reception in 2001/2 is 14329. According to PLASC 2003/4, the size of Hampshire-wide cohort of pupils at year 2 in 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 exclusions of schools which merged or closed reduces our Hampshire sample to 11702 pupils.

[Insert Table 5 about here]

Examination of FSM eligibility recorded in PLASC over time (Table 5) suggests that there is a substantial change in individual FSM status over this 4-year period. Although the yearly average remains relatively constant at about 9%, almost 15% was actually FSM eligible at some time during this period. This suggests that the pool of disadvantage is underestimated by single year snapshots.

In Table 6 we present estimates of key probabilities characterizing the dynamics of poverty defined at the income thresholds implied by FSM-eligibility. We compare the associated estimates under two scenarios. The first scenario assumes that PLASC records of FSM-claims are an accurate representation of FSM-eligibility. The second scenario makes the more realistic assumption that PLASC records of FSM claims are perfectly specific (i.e. non-eligible pupils do not claim FSMs) but FSM claims do not perfectly identify all those eligible. Estimates of the ability of FSM claimant records to detect those FSM-eligible are also presented.

<Insert Table 6 about here>

The estimate of the detectability parameter (Table 6) implies that the error associated with the official FSM-eligibility data is relative large with an average of 9% FSM-eligible not identified by the claimant records (95% credible interval is [8% 11%]). We also find that ignoring the measurement error associated with claimant records will significantly over-estimate the probabilities of transition into and out of the income thresholds implied by the FSM-eligibility criteria. As a result, the pool of the most disadvantaged pupils i.e. those who consistently remain below such income thresholds is under-estimated by 50% on average (Table 6).

The estimates of the presented transition probabilities – after accounting for measurement error - also imply large reductions in the transition probabilities of new FSM-eligibility cases between 2002-2004 which is the period between baseline and KS1 testing for the children of our cohort. This measurement error (ME) analysis reveals that the expected proportion of families representing new FSM-eligibility cases during the year of KS1-testing as compared to the beginning of schooling is 1.03% on average and could range between 0.5% - 2%. These estimates in turn imply a misclassification probability of entering FSM status of 60% on average and which could range between 24% - 80%.

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 for the Greenwood sub-sample.

Value Added Analysis and Measurement Error

In Table 7 we present the results of a value-added analysis on the performance in mathematics at KS1 for the pupils in Greenwood sub-sample. Assessment of the effect of poverty indicators with such a small prevalence such as FSM in a small sample such as Greenwood subsample on mathematics tests is a rather conservative example for testing the effects of measurement error. This is because of both statistical (power) considerations as well as substantial ones such as the nature of the subject tested. However, interest also lies in comparing the effects of poverty indicators such as FSM-claims with other more sensitive indicators of SES.

<Insert Table 7 about here>

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 subsample revealed some surprising results. The analysis suggests that conditional on these baseline scores these FSM eligible children make significantly more progress in mathematics 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 least progress was made by children who were newly FSM eligible 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, 2002,43-53). 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.

In Table 8 we investigate the consequences of ignoring the measurement error in FSM-claimant data on the resulting estimates for the effect of FSM claims via sensitivity analysis. In this sensitivity analysis we compare the estimates of the effect of FSM under different assumptions on the size of misclassification probabilities for FSM eligibility as assessed by the previous Hampshire-wide analysis on the poverty dynamics. We also included scenarios that allowed for measurement error in the tests results (Table 8).

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 for very small misclassification probabilites. 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. However, for average levels of error (60%) they become more substantial (25% change) in the estimated effect size. Moreover, if combined with measurement errors in the baseline tests, the misclassification error leads to further reductions (33% change) in the effect estimates.

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. Also, 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.

<Insert Table 8 about here >

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. This discounts assumptions of stability of FSM status over time and invalidates the statistics based on such assumptions.

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. A failure to correctly identify eligibility could occur due to social processes underlying child poverty in flexible labour markets combined with the data collection procedures. For example, home changes were found to be related to home ownership. In our Greenwood 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 accommodation. 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. This raises the question of whether such turbulent children are tracked through school changes.

Over and above this there are questions about how accurately the data are reported. For example, the Pupil Annual School Census data records FSM eligibility if claimed 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, not all schools send home forms for parents to fill in, rather as our study of Greenwood revealed some schools estimate the proportion of those eligible for FSM. As one principal explained:

> 'We have tried sending out FSM forms for parents to complete, but with limited success (we do include legal stuff but less than 50% return) so we use our local knowledge.'

We used county-wide data to assess the magnitude of error that can be introduced in estimates of the prevalence of economic disadvantage in this population when FSM official records are used to measure it. We found that FSM is both a coarse and error-prone instrument. The associated error was found to be large (10%). It was also found to lead to underestimation of the proportion of children who consistently remain below the income thresholds implied by the FSM-eligibility criteria, by 50%.

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 test performance 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 imperfect sensitivity of FSM-claimant records to identify those truly eligible (misclassification error) 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.

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. Further analysis is currently being undertaken to assess the size of poverty related educational disadvantage while adjusting for the error in the FSM poverty indicator using test results in literacy from large scale samples (Kounali *et al.*, 2008).

Our findings suggest that ignoring the error in the official FSM claimant records will underestimate the associated educational disadvantage. 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.

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 seems to be important in explaining differences in attainment. In this comparison children from low income families with low capital assets who do not meet the 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 misclassification are considered along with relatively high levels of unreliability for the baseline and KS1 test scores, there will be a 33% underestimation of the effect of FSM entitlement.

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 misclassification error associated with FSM eligibility counts in a value-added analysis also seems to be important, although the size of the resulting bias is difficult to ascertain in small samples. Further work needs to be done on a larger scale VA analysis whilst accounting for measurement error in predictors such as FSM-eligibility.

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 variability of official FSM eligibility records over time, however, only reflects one aspect of deprivation predominantly related to family unemployment and lone-parenthood. In fact, (Hobbs & Vignoles, 2007) report 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 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.

This work questions the architecture of accountability which drives the state theory of learning in England (Lauder *et al.*, 2006). Our findings suggest that many schools will confront far greater levels of disadvantage than what is currently measured by FSMs. In this context Ball's discussion of performativity may be apt when he notes (Ball, 2003,215-228):

> "Truthfulness is not the point – the point is their effectiveness in the market or for inspection, as well as the work they do 'on' and 'in' the organisation – their transformational impact"

It is important not to see the problem of quantifying the poverty related educational disadvantage as just confined to measures such as FSMs (Miles & Evans, 1979). Rather, it can be argued that disadvantaged populations will always be difficult to 'capture' through single catch-all measurements from routinely collected administrative data such as FSMs. To address many of the fundamental questions raised here, there is a need for better documentation of the data already collected. Documentation needs to include concurrent collection procedures and uses as well as more research validating the quality and scope of use.

References

Allen, R. & Vignoles, A. (2007) What should an index of school segregation measure? *Oxford Review of Education,* 33(5)**,** 643-668.

Ball, S. (2003) The teacher's soul and the terrors of performativity. *Journal of Education Policy,* 18(2)**,** 215-228.

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)**,** 43-53.

Croxford, L. (2000) Is free school entitlement a valid measure of school intake characteristics. *Educational Research and evaluation,* 6(4)**,** 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 (2005) Has the social class gap narrowed in primary schools? *A background note to accompany the talk by Rt Hon Ruth Kelly MP, Secretary of State for Education and Skills: Education and social progress.*

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(369-395.

Goldstein, H., Kounali, D. & Robinson, A. (2008) Modelling measurement errors. *Statistical Modelling International Journal (in press).*

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

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, County & Council (2006) Census 2001 results for the Hampshire area. IN Environment (Ed.).

Hobbs, G. & Vignoles, A. (2007) Is free school meals a valid proxy of socio-economic disadvantage (in school research)? *CEE Discussion paper, Paper No: CEEDP0084.*

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

Kounali, D., Robinson, A. & Goldstein, H. (2008) The probibity of free school meals in official records as a measure of educational disadvantage. (*CMM preprints),* Centre of Multilevel Modelling, University of Bristol.

Lauder, H., Brown, P., Dillabough, J. & Halsey, A. H. (2006) *Education, globalization and social change.* (Oxford, OUP).

Lauder, H., Kounali, D., Robinson, A., Goldstein, H. & Thrupp, M. (2008) Social class, pupil composition, pupil progress and school performance: An analysis of primary schools. University of Bath (working paper).

Miles, I. & Evans, J. (1979) *Demystifying social statistics.* (London, Pluto Press).

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.* Report for Deputy Prime Minister Office: Neighbourhood Renewal Unit (London).

Plewis, I. & Goldstein, H. (1997) The 1997 education white paper - a failure of standards. *British Journal of Curriculum and Assessment,* 8(1)**,** 17-20.

Rasbash, J., Steele, F., Browne, W. & Prosser, B. U. O. Bristol (Ed.) (2005) *A user's guide to mlwin version 2.0*. Available online at: http://www.cmm.bristol.ac.uk/MLwiN/index.shtml.

Rosina, H. & Downs, M. (2007) *Collecting data for a school*. Available online at: http://www.bris.ac.uk/Depts/CMPO/PLUG/events/170107/main.htm.

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)**,** 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)**,** 309-328.

Spiegelhalter, D., Thomas, A., Best, N. & Lunn, D. (Ed.) (2003) *Winbugs user manual.* Available online at: http://www.mrc-bsu.cam.ac.uk/bugs/winbugs.

Storey, P. & Chamberlin, R. (2001) *Improving the take-up of free school meals.* Report for Institute of Education.

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

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

List of Tables:

List of Figures:

Table 1: Counts of FSM eligible (*) pupils according to family SES, employment status and lone parenthood.

Family SES class (**)	Family employment status				Lone parenthood	Renting
Counts FSM eligible / Cell count	None work	Only one Part-time	At least one full-time	Both full-time		
High	0 / 0	1 / 7	2 / 201	0 / 39	2 / 20	2 / 17
Middle	1 / 3	2 / 17	5 / 486	0 / 87	4 / 46	4 / 79
Low	6 / 7	5 / 54	4 / 490	0 / 73	11 / 110	12 / 205
Unknown	96 / 131	1 / 13	1 / 41	0 / 4	73 / 108	89 / 133
Total (Column %)	103 / 141 (8.53%)	9 / 91 (5.51%)	12 / 1218 (73.68%)	0 / 203 (12.28%)	90 / 284 (17.18%)	107 / 434 (26.26%)

(*) : FSM eligibility as recorded by the parent / carer
(**):The assessment of Socio-Economic Status is based on parental occupation (see Appendix A)

Table 2: Counts of FSM eligible (*) pupils according to parental education attainment and family SES

SES (⅄) Count of FSM eligible / Cell count	Missing	Secondary <16 years	Secondary 16 – 19 years	Further and Vocational qualifications	University graduates and postgraduates	Total Total FSM eligible / Row count (SES class %)
High	0 / 1	1 / 26	0 / 17	1 / 73	1 / 130	3 / 247 (14.94)
Middle	0 / 2	2 / 110	0 / 95	5 / 247	1 / 139	8 / 593 (35.87)
Low	0 / 8	9 / 236	0 / 90	6 / 253	0 / 37	15 / 624 (37.75)
Unknown	3 / 12	54 / 85	15 / 28	24 / 54	2 / 10	98 / 189 (11.43)
Total (Employment group %)	3 / 23 (1.39)	66 / 457 (27.65)	15 / 230 (13.91)	36 / 627 (37.93)	4 / 316 (19.12)	124 / 1653

⅄: The assessment of Socio-Economic Status is based on the occupation of the male carer (Details Appendix A). The Goldthorpe scale was used to rank occupational categories
(*) : FSM eligibility as recorded by the parent / carer

Table 3: Counts of FSM eligible pupils according to administrative records and parental response

	Parent reports non-eligibility			Parent reports eligibility		
	Administrative Records (Update January 2005)					
Administrative Records (Update: January 2004)	Non-eligible	Eligible	Unknown	Non-eligible	Eligible	Unknown
Non-eligible	1435	3	0	8	18	0
Eligible	20	15	0	3	88	0
Unknown	43	0	13	0	5	0
Total	1498	18	13	11	111	2
According to parent:	Total FSM non-eligibility counts: 1529			Total FSM eligibility counts: 124		

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 (PLASC)	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
Column Total (Sample %)	124 (7.5%)	434 (26.3%)	483 (29.2%)		
Total				763	

(*): X denotes the presence of the attribute

Table 5: Observed distinct patterns of FSM eligibility/claims over time for the HARPS cohort as identified by the PLASC records (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	-	-	-	-
Yearly Total (%)		1022 (8.73)	1009 (8.62)	1135 (9.70)	1042 (8.90)
Observed Total Number of pupils entering FSM-eligibility (%)			272 (2.55)	308 (2.88)	236 (2.23)
Observed Total Number of pupils recovering from poverty thresholds as measured by FSM-eligibility (%)			285 (27.89)	182 (18.04)	329 (28.99)

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

Table 6: Estimated measurement error (ME) and the effect of ME on the estimates of the poverty dynamics associated with the income thresholds implied by FSM-eligibility

Estimates	Ignoring ME		Accounting for ME	
	Mean (%)	95% CI*	Mean (%)	95% CI*
Estimated Transition probability into poverty	2.6	[1.5 1.9]	2.1	[1.9 2.2]
Estimated Transition probability of recovery from poverty	25	[24 27]	17	[16 19]
Estimated probability remaining in poverty for the whole period	4.1	[3.6 4.6]	6.1	[5.2 7.0]
Estimated detectability of poverty thresholds associated with FSM-eligibility by of FSM-claim records			91	[89 92]

(*): 95% Credible Intervals

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 (†) P(0│1) (*) R, (**) ρ, (°) Ry	FSM at KS1 Mean (SE)	Baseline Maths Mean (SE)	Baseline Literacy Mean (SE)	Level 2 Mean (SE)	Level 1 Variance Mean (SE)
P(0│1)=0%, R=1, ρ=0, Ry=1	-0.32 (0.12)	0.46 (0.03)	0.14 (0.03)	0.08 (0.02)	0.49 (0.02)
P(0│1)=26%, R=1, ρ=0, Ry=1	-0.31 (0.13)			0.08 (0.02)	0.49 (0.02)
P(0│1)=60%, R=1, ρ=0, Ry=1	-0.30 (0.12)			0.08 (0.02)	0.49 (0.02)
P(0│1)=80%, R=1, ρ=0, Ry=1	-0.28 (0.12)			0.08 (0.02)	0.49 (0.02)
P(0│1)=60%, R=0.8, ρ=0.5, Ry=0.9 (∇)	-0.24 (0.12)	0.65 (0.06)	0.08 (0.06)	0.08 (0.02)	0.42 (0.02)

†: P(0│1) denotes the misclassification Probability of observing a pupil as not being FSM eligible when he is actually eligible
(*) R denotes the Reliability of the baseline tests; the reliability is assumed to be the same for both tests
(**) ρ denotes the correlation between the measurements errors for the baseline tests
(°) Ry denotes the reliability of the outcome i.e. KS1 test scores in mathematics
(∇) Introducing P(0│1)=60% and P(1│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.

Appendix A: Occupation coding scheme

In this section, we provide some details on the classification system used to characterize social class, having recorded occupation categories using the Goldthorpe occupation-scale (Goldthorpe and Hope 1974).

SES class	Qccupation category used in the questionnaire
High	Professionals
Middle	Managers/Administrators; Associate Professionals
Low	Skilled Craftsmen; Clerical/secretarial; Sales; Machine Operatives; Personal and protective services
Not working	Employement data recording lack of work at both for both of the carers.

The occupation of both carers at present and in the past was recorded and used for assessing SES as follows:

The family SES is the current occupation of the male carer and the current occupation of the female carer in the absence of response from the male carer. We compared different methods of combining current and historical occupational information from both carers. 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 *et al.*, 2006). Based on this analysis, we outline below the factors which were found to be associated to such biases i.e. when the highest occupational class is used among carers at present or historically.

Adopting the widely used strategy of considering the highest occupational class between carers resulted in exaggerated representation of the professional and managerial occupational groups when compared with data with the Hampshire and national statistics on occupation – with the associated proportions almost twice as high as those reported in the county-wide national statistics.

Also we found that almost 45% of the occupation codes determining the family's SES (as the highest occupation in the couple) were those of the male responders or partners. It is also interesting to note, that in the occupational classes associated with the highest and middle SES (as defined in the Table above) the proportion of male-determined codes were close to the average while the lowest and missing or unemployed classes were predominantly determined by females. In those later low SES classes a significant proportion (45% of clerical/secretarial; 49% of Sales / Machine Operatives / Personal & Protective Services) and 67% of the non-responders and unemployed) were single parents. It is clear that family structure (i.e. single parenthood) is associated with SES where the proportion of single parents in the higher SES occupations is 7%, compared to 11.3% and 26% in the middle and low SES occupations, respectively.

Also, we found that the majority of responses on the highest occupational category refer to the past (64.4%). We also see that the majority of the current ones (55.7%) refer to the occupation of the male bread-winner from high occupational categories and the majority of past ones (61.1%) refer to female bread-winner from low occupational categories. This suggests that the bread-winner has a male gender. If we look closer at the change of occupational status for the major bread winner we find that those with higher SES occupations suffer less in the job market (job-stability/ insecurity). A total of 365 families (22.1%) experienced a worsening of their occupational status. Among these families, 81% corresponds to female bread-winners. Among higher SES occupations 20.7% experienced a worsening of their occupational status compared with 23.7% and 24.3% for the middle and low SES occupations. The gender of the bread-winner modifies this relationship and suggests that working mothers might experience a tougher deal in the job market. More specifically, we find that if we control for the gender of the major bread-winner then among females with occupations associated with high SES 27.4% experience worsening of their occupational status. This worsening of occupational status is 36.9% and 39.2% among women with middle and low SES occupations, respectively.

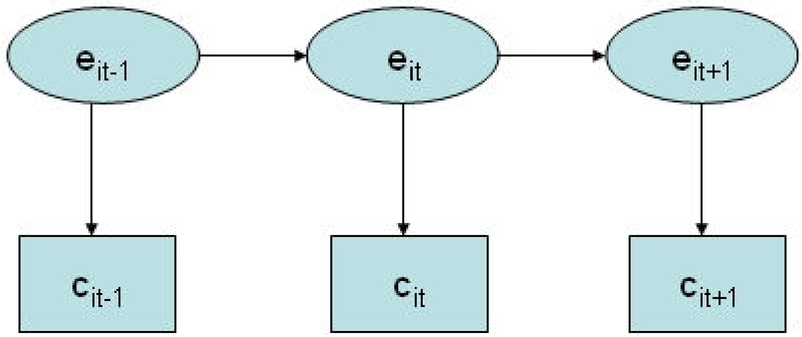

Figure 1: Temporal evolution of a Hidden Markov Model

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.