Can school league tables help parents choose schools?

Rebecca Allen
Institute of Education, University of London

Simon Burgess
CMPO, University of Bristol

November 2010

Abstract
Information on schools' performance is central to the school choice process in England. This paper evaluates school performance tables against criteria of functionality, relevance, and comprehensibility. We estimate a model to judge functionality by setting up the following question: "In which feasible school will my child achieve the highest exam score?" We show that neither of the current leading performance measures score very well. We propose an alternative measure that performs better on these criteria and demonstrate how the measure could be best delivered to parents. We also describe the trade-off between the criteria of functionality and relevance.

Keywords: school choice, performance tables

JEL Category: I20

Thanks to the DfE for access to the NPD data, and to the ESRC for funding Burgess through CMPO. Thanks to Paul Clarke, Lorraine Dearden, Harvey Goldstein, Carol Propper, and Anna Vignoles for comments on a related paper. None of these are responsible for the views expressed here.

1. Introduction

For over two decades English school admissions have allowed parents to express preferences for schools and these are used alongside school oversubscription criteria to allocate pupils to schools. This school choice system requires information to enable parents to identify high quality schools and school performance tables are the most important indicators of academic quality of school. This paper investigates the role of school performance tables as a component of a school choice system, arguing that we should judge their success based on their functionality, relevance, and comprehensibility.[1] We judge a measure as functional for parents if it is a useful predictor of their own child's likely exam performance; this is the core of our analysis. It is judged as relevant if it informs parents about the performance of children very similar to their own in ability and social characteristics. It is comprehensible if it is given to them in a metric that they can meaningfully interpret. We show that neither of the current leading performance measures score very well across our three criteria. We propose an alternative measure that performs better on these criteria and demonstrate how the measure could be best delivered to parents. No measure can be perfect because there are important trade-offs between relevance, functionality and comprehensibility which we set out below and return to in our concluding comments.

Parents consistently rank academic standards as being one of the most important criteria for choosing a school, both in surveys and as revealed in their stated preferences on application forms (e.g. Coldron et al., 2008; Burgess et al., 2010). This school attribute is balanced against other considerations such as proximity to home, the child's own wishes, school ethos, behaviour, and so on. Recent surveys conducted by Coldron et al. show that parents use a wide variety of formal and informal sources of information to choose schools, including Ofsted reports, school visits and word of mouth. Just below half of parents say that they make use of school performance tables, published annually on the Department for Education website; a higher fraction may be aware of the rankings in general terms from the major media coverage of the tables. A large proportion of parents appear to use and understand the statistic that reports the proportion of pupils who successfully achieve five or more GCSEs at grades A* to C, including maths and English; this indicator is seen as critical because it is a threshold needed to proceed to post-16 academic study. A smaller proportion (36 percent) report that they use the contextual value added (CVA) score that is intended to reflect the 'effectiveness' of practices at the schools, controlling for the school's social context.

The information that a government chooses to provide about school performance is critical because it has the capacity to influence the schools that parents choose. A US experiment that provided lower-income families with direct information on school test scores in a public school choice plan showed that receiving the information significantly increased the fraction of parents choosing higher-performing schools (Hastings and Weinstein, 2007). Moreover, where children from lower-income families actually attended these higher-performing schools it had a positive impact on their

[1] We also discuss the extent to which they minimise gaming of examinations, though this is not a major focus here.

final academic achievement. Critics of league tables argue that they simply exacerbate over-subscription at popular schools. However, more intuitive information does have the potential to reduce the social gradient in how parents choose schools and so may achieved more balanced intakes at popular schools if a greater proportion of disadvantaged families apply.

The next section gives a brief overview of secondary schools admissions in England and we then describe the data that is used in the remainder of the paper. Next, we propose and implement a method for testing whether alternative measures of judging the performance of schools are functional, i.e. whether they are good predictors of a child's future academic performance. We then discuss the criteria of comprehensibility and relevance. Finally, we propose an alternative approach to offering parents school performance information. The final section offers some conclusions.

2. School choice in England

In this study we analyse the outcomes of secondary school choices made by parents. Most pupils in England transfer from primary to secondary school at age 11, although there are a few areas where this transfer is slightly different due to the presence of middle schools. Secondary school allocation takes place via a system of constrained choice whereby parents are able to express ordered preferences for three to six schools anywhere in England and these preferences are reconciled locally with published admissions criteria that must adhere to a national Admissions Code. A school-deferring Gale-Shapley two-sided matching algorithm that considers all of a parent's preferences simultaneously is used for the reconciliation (Gale and Shapley, 1962). First priority is usually given to pupils with a sibling already at the school, pupils with statements of special educational needs and to children in public care. However, the largest proportion of places is allocated giving priority to children living within a designated area or on the basis of proximity to school. This means that house location continues to be critical to securing a school place in England, despite the choice system that exists. There are a significant number of schools which do not give priority to local communities: at voluntary-aided religious schools (17 percent of secondary pupils), priority is usually given on the basis of religious affiliation or adherence; there are also state schools that offer a proportion of places on the basis of ability or aptitude for a particular subject, and this includes 164 entirely selective grammars schools.

3. Data

In this analysis we draw pupil-level data from all eight years (2002 to 2009) of the National Pupil Database (NPD) to measure school performance in a variety of ways, described below. NPD is an administrative dataset of all pupils in the state-maintained system, providing an annual census of pupils taken each year in January, from 2002 onwards (with termly collections since 2006). This census of personal characteristics can be linked to each pupil's test score history. We use a single cohort to analyse the potential consequences of the secondary school choices made by over 500,000 pupils who transferred to secondary school in September 2004, completing compulsory education in 2009. These pupils are located in 3143 secondary schools; we exclude non-standard schools such as special schools or those with fewer than 30 pupils in a cohort from the analysis. We drop a small number of pupils from our analysis because they appear to be in the incorrect year group for their age or they have a non-standard schooling career history.

NPD provides data on gender, within-year age, ethnicity, an indicator of whether English is spoken at home (EAL), and three indicators of Special Educational Needs (SEN, measuring learning or behavioural difficulties at a high, medium and low level, respectively). It also provides us with two measures of the socio-economic background of the child. Free School Meals (FSM) eligibility is an indicator of family poverty that is dependent on receipt of state welfare benefits (such as Income Support or Unemployment Benefit). Our FSM variable is a very good measure of the FSM status of the 12 per cent of our cohort who have it, but it has been shown by Hobbs and Vignoles (2010) to be a crude measure of household income or poverty status. We also use the Income Deprivation Affecting Children Index (IDACI), an indicator for the level of deprivation of the household's postcode.[2]

Data on individual pupil characteristics are linked to educational attainment at the ages of 7 (Key Stage 1 – KS1), 11 (KS2) and 16 (GCSE or equivalent examinations). The linked test score data that measures the academic attainment of children in KS2 tests at the end of primary school serves as a useful proxy for academic success to date. We use an overall score (KS2) that aggregates across all tests in English, maths and science, as well as the individual subject scores in our analysis.

4. Functionality of league tables

Performance tables are only helpful to parents in choosing schools if their use improves the decisions made. The question that parents need an answer to is: "In which feasible school will my child achieve the highest exam score?" Our argument is that the best content for school performance tables is the statistic that best answers this question. If no performance measure can provide better guidance than choosing a school at random, then we would conclude that performance tables in this context are valueless.

This is what we test here, and set it up as follows. We identify a feasible choice set of schools for each student, and define a set of school choice decision rules based on different school performance tables. These allow us to identify the school that each pupil would have chosen under each decision rule. We then estimate the counterfactuals: how that particular pupil would have scored if they had attended that school for the following five years. Finally, a comparison of that outcome with a choice at random from the feasible choice set – that is, a choice uninformed by performance tables – tells us whether using that decision rule was successful for that student. Further details on these steps can be found in Allen and Burgess (2010).

(a) Choice sets and Decision Rules

It is impossible for us to know which schools any particular parent is actively considering for their child because this will be a function of their own preferences and constraints; there are no legal restrictions on choices. Instead, we define a choice set for every pupil by including a school in

[2] For more information see http://www.communities.gov.uk/documents/communities/pdf/131206.pdf (accessed 17/05/10).

her/his choice set if another similar pupil from the same neighbourhood attended the school during the eight year period of 2002 to 2009 for which we can observe secondary school destinations. The pupil's neighbourhood is defined as a lower layer super output area (SOA), a geographical unit that is designed to include an approximately equal population size across the country.[3] In our data, on average 123 pupils live within an SOA, adding across all eight cohorts. Our first stage of defining the pupil's choice set is to calculate an SOA destination matrix for all 32,481 SOAs. We also need to exclude schools from the choice set if very few 'similar' pupils attended the school in the main 2009 cohort because we are unable to estimate likely pupil outcomes if there are no similar pupils in the school. Therefore, a school is excluded from a pupil's choice set if fewer than 1% of that school's 2009 cohort are of the same sex, EAL, SEN, ethnicity (white British, Asian, black, other) or KS2 group (indicating low, middle or high ability). The result of all these restrictions is that pupil choice sets are slightly smaller than SOA destination lists: pupils have between one and 18 schools in their choice set (mean 5.07; SD 2.35).

To model the decision rules, we use information that would be available to parents whose children start secondary school in September 2004. Pupils typically take nationally set, high stakes, GCSE or equivalent examinations in 8 to 10 subjects at the age of 16 and these are measured on an eight-point pass scale from grade A*, A, B, ... to F, G. In this paper we focus on:

1. Proportion of pupils achieving 5 or more GCSEs at grades A*-C, including at least a grade C in both English and maths (%5A*-C DR). This rather crude threshold metric has been used to measure school performance since 1992 (with the inclusion of English and maths restrictions from 2006 onwards).
2. The contextual value added (CVA DR) score for the school, similar to that published for all secondary schools from 2006. This is essentially a school residual extracted from a multi-level regression that conditions on the pupil and peer characteristics available in NPD (see Ray, 2006). We calculate our own school CVA-type scores because it was not published by government in 2003.

We estimate the expected GCSE outcome for each school using the 2009 outcome data, which is the cohort making their school choices in 2003, starting secondary school in 2004. We estimate a very flexible model, allowing for maximum school-level heterogeneity by estimating a separate regression for each of the 3143 schools. We include as predictors six measures of the pupil's prior attainment, personal characteristics such as gender, poverty status, ethnicity, language, special educational needs, and neighbourhood characteristics. In addition we include interactions of all these characteristics. The results of these regressions are summarised in Appendix 2 of Allen and Burgess (2010).

An important issue is the potential effects of school selection bias. Students are not randomly assigned to schools and there are unobserved student characteristics that influence both the probability of assignment and subsequent exam performance. We cannot model the assignment process explicitly and so, absent any instrument for school assignment (such as that used by Sacedote, 2010), we will have biased estimates of school effects. Essentially, we will overestimate the quality of schools with unobservably good pupils. This means that we will impute higher scores

[3] A SOA is a small geographical unit, containing a minimum population of 1000 and a mean of 1500.

to the counterfactual pupils not at that school than they would truly have achieved had they attended. This is a well-known problem and it faces all attempts to estimate true school effects and to interpret school performance data; it is not an additional problem for our approach.

We take two practical steps to minimise the bias. First, we use as many observable student characteristics as possible in the school level regressions, including measures of student progress between ages 7 and 11 in some specifications to capture differences in progress from age 11 to 16. All of these are interacted with other individual characteristics. Second, we only consider counterfactual pupils for plausible local schools, and do not use predictions for schools with no similar students to the focus student. Beyond this, we can make statements about the nature of the bias if we explicitly parameterise the assignment process. We show in Allen and Burgess (2010) that in some important cases the impact of selection bias is to change the estimated effectiveness of schools without changing the ranking.

(b) Functionality Results

The results in Table 1 show the chances that each decision rule identifies a school that turns out to have been a good choice. We benchmark each decision rule against an uninformed choice and compute the odds ratio of making a good choice against a bad choice. We compare the outcome of the decision rules with the expected value of a choice at random, namely the mean outcome for each student over all schools in her/his choice set.[4] We consider how often choosing the best school according to the threshold DR is at least as good as a random choice, how often choosing a good school from the tables is at least as good as random, and whether the school identified by the tables as the worst choice turns out to be worse than random. We define a good school as one chosen at random from the top half of the performance table for the pupil's choice set.

Using the %5A*-C decision rule correctly identifies a school where the child should outperform the average across their choice set 1.92 times more frequently than it identifies one where the child performs worse. Clearly this means that a substantial fraction of students would turn out to be badly advised by the performance tables; but the number for whom they proved useful is almost twice as large. The remainder of the table disaggregates the performance of the decision rule. Picking the best school according to this decision rule turns out to be better than random with odds of 2.92 for the top third of KS2 students, compared to the just 1.37 for the bottom third of KS2 students. The %5A*-C at GCSE decision rule also performs better when the variation between schools (on the 2003 decision rule measure) is greater. This intuitively makes sense because where there are greater differences between schools in 2003, there should be a greater chance that the rank ordering is maintained over time.

This %5A*-C has considerably better predictive power than CVA, which delivers good choices only 1.33 times more frequently than bad choices. CVA was introduced to English league tables to capture the underlying effectiveness of the school, controlling for all measured pupil and peer characteristics. However, the poor performance of CVA suggests that 2003 underlying effectiveness

[4] In principle we would model an uninformed choice as a choice at random. However, many students face choice sets with just 2 or 3 schools in and in this case, a literal random choice will produce a very high percentage of ties.

is not a particularly strong predictor of a child's likely 2009 GCSE attainment. It scores particularly poorly relative to the %5A*-C rule in areas with a lot of variation in school effectiveness.

To understand the relative performance, we can distinguish two reasons for a 2003 decision rule performing poorly in explaining 2009 expected outcomes for a child. First, the ranking of local school exam performances may not be particularly stable. Second, the value of the metric for even the contemporaneous cohort may be only weakly related to our estimate of any one specific pupil's estimated exam performance at that school. A decomposition into these components shows that the %5A*-C rule is extremely stable over time, but that the CVA rule is not. This relatively low local stability of CVA is consistent with the low national stability reported by Leckie and Goldstein (2009). However, CVA scores better on the issue of contemporaneous fit: measures that more closely identify a school's effectiveness in 2009 are indeed useful in predicting a child's own likely exam performance. However, this superior predictive power in the contemporaneous cohort is not sufficient to offset the high instability in the CVA rule over time. If parents only had to predict the best school for their child one year ahead, then metrics getting closer to effectiveness do well; over longer time horizons this is outweighed by the slightly lower predictive power but greater stability of the unconditional measures.

We use a bootstrap procedure to express parameter uncertainty in this model. Because our outcome variable is complex, based on a nonlinear function of the predictions of a number of separate regressions, we apply a non-parametric bootstrap to our entire estimation procedure, including the individual regressions for each school. For computational reasons, we restrict attention to pupils in London (approximately 13% of our total sample). These results in Table 2 show that the lower bound on the confidence interval is well above unity for all our decision rules; in fact it is more than five standard errors above one for all except the CVA rule, which clearly performs worse than the others; its confidence interval is non-overlapping with the %5A*-C rule, so the latter is significantly and unambiguously superior.

We postpone discussion of the other two columns of results to section 6.

5. Comprehensibility and relevance of league tables

Having established that school league tables are functional in helping parents make school choices, we now consider the criteria of comprehensibility and relevance. These issues are not of second order importance. Hastings and Weinstein (2007) show that providing simpler school choice information has a significant impact on both choices made and on subsequent pupil test score outcomes (check). This is in line with other recent studies showing that the simplification of college finance applications have significant impacts on decisions (see for example, Bettinger et al, 2009).

The core of the argument about comprehensibility is that the school league tables should be easy to understand correctly. First, they should be in a natural metric. The %5A*-C threshold measure reports the fraction of pupils achieving good passes. This is straightforward and clear, and it is also clear how the measure was derived. However, one straightforward interpretation of this measure is not necessarily correct. The implication that a parent might draw from this is that this is the probability that their own child would have of securing five good passes if they attended that school.

The fact of non-random allocation of pupils to schools means that this will not necessarily be the case in general. The CVA measure is reported as a score centred on 1000. This has no natural interpretation or translation into any metric that the parent is interested in, namely GCSE grades. The derivation of the measure is also – for all but statisticians – obscure and mysterious. CVA measures were created for the purposes of school accountability and so they are deliberately and necessarily complex in both their calculation and their interpretation (Ray, 2006). Gorard (2010, page 757) describes the CVA scores as being '*magic figures emerging from a long-winded and quasi-rational calculation*'. The complexity of CVA's calculation is not a problem for accountability measures since they are to be used by experts, but parental choice measures should be straightforward to interpret.

To be relevant to a particular family, school league tables need to relate directly to that family's child and location. This means that the set of schools displayed need to be ones that the family could feasibly attend given the location of their home. This illustrates the important distinction about who the league tables are for. For accountability purposes, league tables can be complex and should use "statistical neighbours"; for school choice purposes, league tables need to be simple and should provide "physical neighbours". There is a divergence in terms of delivery media here. Printed media typically provide a league table of schools only at an LA level, but web-based delivery enables a much more bespoke format. Figure 1 illustrates the UK Department for Education's school performance website which allows users to enter their postcode and browse the performance tables of local schools. One issue is whether there are socio-economic gradients in the use of web-based delivery as opposed to printed and broadcast media.

The second component of relevance is the characteristics of the family's child, including her/his ability but also other factors such as special needs. It is clear that children arriving at a school with different levels of ability will, on average, continue to perform differentially in the final exams. But different schools may produce different outcomes for children at different parts of the ability distribution and this differential effectiveness is lost in the averages. None of the available measures provide any disaggregation by ability, and this seems to us to be the single most important failing of the current system. The CVA approach attempts to control for individual ability to provide a measure of overall school effectiveness, but again this is assumed to be the same for all students in the school (Jesson and Gray, 1991; Sammons et al., 1993; Thomas et al., 1997; Wilson and Piebalga 2008).

6. An alternative proposal

We have argued above for the importance of functionality, comprehensibility and relevance for league tables. In this section we propose a way of representing school attainment information that is more functional and comprehensible than CVA data and more relevant than the %5A*-C data. Our measure gives parents the expected GCSE performance for a child of similar ability to theirs for all schools in their local choice set in a straightforward outcome metric. It is relatively simple for parents to interpret and we test the functionality of the proposed measure below.

Figure 2 shows the format we propose: GCSE attainment information for six schools in close spatial proximity to each other. The chart shows the average GCSE attainment (in their best eight subjects) for pupils at three different points in the ability distribution: children who scored at the 25th, 50th and 75th percentile on their Key Stage two tests at the end of primary school.[5] It therefore allows schools to be differentially effective and also highlights the importance of prior attainment in determining future exam success.

In the diagram it is possible for the parent to observe both the ordering and the spacing of schools for the ability of their child.[6] This allows them to see how substantively important choice of school can be. For example, this diagram illustrates the larger differences in GCSE outcomes across schools for the lower ability pupils whereas choice of school appears to be less critical to exam success for higher ability pupils. This should give parents a realistic perspective of the extent to which school choice matters for attainment, thereby empowering them to balance this judgement against other school dimensions such as atmosphere, ethos, and so on.

The outcome measure we report on this diagram is the capped GCSE scores, recalibrated into average grade over best eight subjects. We have re-calibrated the measure because we believe that average GCSE grade is a measure that most parents can understand and interpret. However, it is possible to use any outcome measure on these types of diagrams. There are no confidence intervals around the estimates: it would be perfectly possible to generate them using a similar approach to the current CVA measure, but we are concerned that they are not good measures of true uncertainty of performance since they do not incorporate the uncertainty on estimates that result from measurement error on variables such as test scores and they also introduce considerable complexity into the diagram.

By giving measures of school performance based on a capped GCSE outcome variable that differ by prior attainment of the pupil, the measure we propose here bears similarities to that proposed by Dearden et al. (2010). However, there are four differences in our approach. First, we report local rather than national comparisons of schools. Second, we disaggregate information into only three ability groups rather than 8. Third, we do not report p-values or other measures of statistical significance of differences between schools in the performance tables. Finally, we report in a metric that is more straightforward for parents to understand (average GCSE grade rather than GCSE points scores and value-added measures).

We noted above a trade-off between functionality and relevance; there is also a trade-off between comprehensibility and relevance. Our proposed approach clearly produces imperfect measures at the expense of reporting a straightforward and easily interpreted outcome metric. For example, we do not propose removing the national mean average 'effect' of other pupil social characteristics such as sex, ethnicity and poverty from our estimates to maintain the simplicity of interpretation. This will be more of a problem in areas where local schools are very different to each other in characteristics that explain large differences in pupil progress, such as pupil sex and ethnicity.

[5] The attainment at the 25th percentile is calculated as the average GCSE performance for pupils scoring between the 20th and the 30th percentile at GCSE, and is only reported provided that at least 4 pupils can be used in the calculation. There are many alternative approaches to calculating this average performance, including regression based parametric approaches to estimating the school's relationship between prior attainment and GCSE performance.

[6] Primary schools might be able to give expected levels, but general guidance would be needed to help parents.

Other problems are a product of the way we currently measure exam performance and there is nothing that ex-post analysis can do to compensate for this. For example, there is significant censoring at the top end of the distribution because pupils are able to score no better than an A* in any particular subject and cannot contribute their full potential share to these value added type scores. This may partly explain why the performance of very affluent schools tends to be more similar. The censoring would not be a problem if all parents care about is GCSE exam results, but if they are concerned with broader skill and knowledge acquisition of their child, GCSE exams imperfectly measure this for high attaining pupils.

This approach to displaying school performance data is more comprehensible and relevant than the existing measures reported in school league tables. We return to Tables 1 and 2 to assess the functionality of the measure, where we present the results for our proposed measure in the final column, and, for comparison, a simple average GCSE grades measure in column 3. The top row shows that the odds ratio of successful decisions to unsuccessful decisions is 1.69 for the proposed measure, relative to 1.92 for the %5A*-C measure and 1.33 for the CVA measure.

It is initially surprising that the performance of the differential measure is no better than that of the simple average measure. Intuition suggests that the provision of more information should do better; that having information on different parts of the distribution is more useful than just the average. There are several reasons why this might be the case. It may be because schools are not differentially effective in a stable manner over time. Also, differential effectiveness measures will not be more informative than raw effectiveness if only the size, and not the ranking, of school effects varies within a choice set at different parts of the ability distribution. Within our choice set, schools do indeed have greater variability on the differential measure at the low ability point than the high ability point. However, the Spearman's rank correlation within a choice set using our simple average measure rather than the differential measure at the three ability points is high at an average of around 0.7 for each pairwise comparison. This observation that slopes of differential effectiveness as a function of ability often do not cross has been reported in other papers (e.g. Thomas et al., 1997). A final advantage of the simple average score measure is that it incorporates information about school composition, where scores at different points of the distribution do not.

7. Conclusions

School performance tables remain a controversial part of education policy, and their usefulness in informing school choice is strongly debated. The format of the “league tables” is being reviewed again in England, school performance tables are likely to be introduced in Australia, and they are part of the school accountability system in some states in the US. In this paper, we argue that the criteria for evaluating school performance tables are functionality, comprehensibility and relevance. We have shown that the different items of performance information currently emphasised in England are deficient in one or more aspects. The %5A*-C score for a school is functional because it usefully identifies a good school for potential pupils to choose; it is reasonably comprehensible because parents in England are familiar with the meaning of a ‘good’ GCSE; but it has low relevance because it is not a broad indicator of pupil performance at a school overall. The CVA measure is not functional, not comprehensible, but is more relevant.

In designing a better basis for school performance metrics there are important trade-offs between relevance and functionality, and between comprehensibility and relevance. The more disaggregate the form in which performance tables are provided (increased relevance), the less precision they will have (decreased functionality) because of smaller cell sizes. The more factors are taken into account in describing school performance for one specific child (increased relevance), the more complex the reported measure will be (decreased comprehensibility). Any choice on the content of league table information has to make decisions on these trade-offs. Our proposed approach offers a good performance on functionality, though not the highest, offers one dimension of disaggregation, and is simple to understand.

It is clear from Hastings and Weinstein (2007) that information on school performance is important and will change the decisions and outcomes of families, particularly low income families. Given this, and building on the analysis presented here of functionality, comprehensibility and relevance, it would be most valuable to trial different contents and formats of school performance information in field experiments.

References

Allen, R. and Burgess, S. (2010) Evaluating the Provision of School Performance Information for *School Choice. CMPO WP 10/241, CMPO Bristol.*

Eric P. Bettinger, E. P., Long, B. T., Oreopoulos, P. and Sanbonmatsu, L. (2009) The Role of Information and Simplification in College Decisions: Results from the H&R Block FAFSA Experiment. *mimeo*.

Burgess, S., Greaves, E., Vignoles, A. and Wilson, D. (2009) What Parents Want: School preferences and school choice. WP 10/222, CMPO Bristol.

Dearden, L., Micklewright, J. and Vignoles, A. (2010) Information for parents on the effectiveness of English secondary schools for different types of pupil. Mimeo, IOE, London.

Coldron, J., Tanner, E. Finch, S. Shipton, L. Wolstenholme, C. Willis, B. And Stiell, B. (2008) *Secondary School Admissions,* DCSF Research Report RR020.

Gale, D. and Shapley, L. S. (1962). College admissions and the stability of marriage. *The American Mathematical Monthly*, 69(1):9–15.

Gorard, S. (2010) Serious doubts about school effectiveness, *British Educational Research Journal,* 36(5) 745-766.

Hastings, J.S. and Weinstein, J.M. (2007) *Information, School Choice, and Academic Achievement: Evidence from Two Experiments*, Quarterly Journal of Economics, vol. 123, pp. 1373 – 1414.

Hobbs, G. and Vignoles, A. (2010) Is children's free school meal 'eligibility' a good proxy for family income? *British Educational Research Journal*, 36(4) 673-690.

Jesson, D. and Gray, J. (1991) Slants on Slopes: Using Multi-level Models to Investigate Differential School Effectiveness and its Impact on Pupils' Examination Results*. School Effectiveness and School Improvement*, 2(3) pp.230-247.

Leckie, G. and Goldstein, H. (2009) The limitations of using school league tables to inform school choice, *Journal of the Royal Statistical Society: Series A*, 127(4) pp.835-52.

Ray, A. (2006) *School value added measures in England: A paper for the OECD project on the development of value-added models in education system.* London: DfES.

Sacerdote, B. (2010) *When the Saints go Marching Out: Can Hurricances Katrina and Rita teach us about Closing Unsuccessful Schools?* mimeo, Dartmouth College.

Sammons, P., Nuttall, D. and Cuttance, P. (1993) Differential School Effectiveness: results from a reanalysis of the Inner London Education Authority's Junior School Project Data, *British Educational Research Journal*, 19 (4) pp.381-405.

Thomas, S., Sammons, P., Mortimore, P. and Smees, R. (1997) Differential secondary school effectiveness : examining the size, extent and consistency of school and departmental effects on GCSE outcomes for different groups of students over three years, *British Educational Research Journal*, 23 (4) pp.451-469.

Wilson, D. and Piebalga, A. (2008) Performance measures, ranking and parental choice: an analysis of the English school league tables, *International Public Management Journal*, 11 pp.233-66.

Table 1: Decision rule performance (best 2003 school versus mean outcome)

The number reported is the ratio of (pupils who make a choice better than mean) to (pupils who make a choice worse than mean)

	%5A*-C	CVA	Average GCSE	Differential
Overall (choice set>1)	**1.92**	**1.33**	**2.04**	**1.69**
Size of choice set: 2	1.38	1.20	1.43	1.34
3	1.54	1.32	1.75	1.53
4 or 5	1.95	1.36	2.01	1.70
6 to 9	2.21	1.32	2.36	1.89
10 or more	2.39	1.46	2.70	1.80
Lowest ability group	1.37	1.22	1.48	1.25
Middle ability group	1.82	1.35	1.93	1.60
Highest ability group	2.92	1.43	3.10	2.47
Low variation in choice set	1.31	1.11	1.48	1.49
High variation in choice set	2.92	1.61	2.92	1.92
Spearman's rank correlation	0.20	0.11	0.22	0.17

Table 2: Bootstrapped standard errors for London

The number reported is the ratio of (pupils who make a choice better than mean) to (pupils who make a choice worse than mean)

	%5A*-C	CVA	Average GCSE	Differential
England odds ratio	1.92	1.33	2.04	1.69
London odds ratio	1.90	1.69	2.15	1.85
Standard error	0.12	0.11	0.11	0.08
Normal-based 95% confidence intervals	1.66 - 2.14	1.47 - 1.91	1.93 - 2.37	1.70 - 2.00

Notes: choice set>1; Number of observations for London = 131358; Number of replications = 100 (100% sample with replacement).

Figure 1: League tables and local searches

http://www.education.gov.uk/performancetables/schools_09.shtml (accessed 9/11/2010)

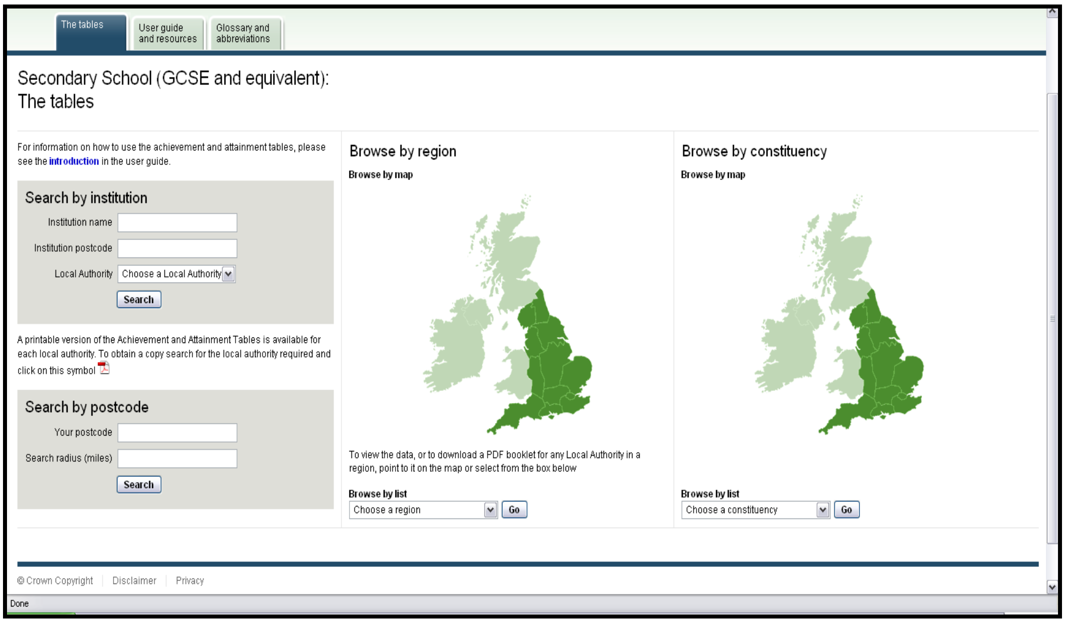

Figure 2: School Performance Information