THE CENTRE FOR MARKET AND PUBLIC ORGANISATION

The Centre for Market and Public Organisation, a Research Centre based at the University of Bristol, was established in 1998. The principal aim of the CMPO is to develop understanding of the design of activities within the public sector, on the boundary of the state and within recently privatised entities with the objective of developing research in, and assessing and informing policy toward, these activities.

Centre for Market and Public Organisation
University of Bristol
Department of Economics
Mary Paley Building
12 Priory Road
Bristol BS8 1TN

Tel: (0117) 954 6943
Fax: (0117) 954 6997
E-mail: cmpo-office@bristol.ac.uk

Who wins and who loses from school accountability? The distribution of educational gain in English secondary schools

Simon Burgess, Carol Propper, Helen Slater and Deborah Wilson

July 2005

Working Paper No. 05/128

ISSN 1473-625X

CMPO Working Paper Series No. 05/128

Who wins and who loses from school accountability? The distribution of educational gain in English secondary schools

Simon Burgess
Carol Propper
Helen Slater
and
Deborah Wilson

CMPO, The University of Bristol

July 2005

Abstract

In 1988 the UK government introduced greater accountability into the English state school sector. But the information that schools are required to make public on their pupil achievement is only partial. The paper examines whether accountability measures based on a partial summary of student achievement influence the distribution of student achievement. Since school ratings only incorporate test results via pass rates, schools have incentives to improve the performance of students who are on the margin of meeting these standards, to the detriment of very low achieving or high achieving pupils. Using pupil level data for a cohort of all students in secondary public sector schools in England, we find that this policy reduces the educational gains and exam performance in high stakes exams of very low ability students.

Keywords: school accountability, high stakes exams, educational value added

JEL Classification: I200, I280, D230

Acknowledgements

We wish to thank Paul Gregg, Paul Grout, Jon Temple and seminar participants at the University of Bristol and the PEUK May 2005 Conference for their helpful comments. We would like to thank the Leverhulme Trust and the ESRC for research funding. All errors remain our own.

Address for Correspondence
Department of Economics
University of Bristol
12 Priory Road
Bristol
BS8 1TN
Carol.Propper@bristol.ac.uk
www.bris.ac.uk/Depts/CMPO/

CMPO is jointly funded by the Leverhulme Trust and the ESRC

1. Introduction

In 1988 the UK government introduced greater accountability into the English state school sector. Schools have to publish information on their performance. But the information that schools are required to make public on pupil achievement is only partial. For secondary schools (those that educate children between the ages of 11 to 18), a key piece of information that is published is the proportion of pupils who achieve above a certain level in national tests taken at age 16. These data are published by government and are used by the media to rank schools in nationally published school ‘league tables’. They are also used by central government to sanction poorly performing schools and by parents in choosing a school for their child. Poorly performing schools may lose pupils, so losing resources, and/or their management may be replaced, while highly performing schools will gain pupils if they are not already at capacity.

The focus of this paper is to examine whether accountability based on a partial summary of student achievement influences the distribution of student achievement. Since school ratings only incorporate student performance via a target indicator based on pass rates, schools have incentives to improve the performance of students who are on the margin of meeting this target. Schools might therefore focus their efforts on the marginal pupils, to the detriment of very low achieving or high achieving pupils, as the former group are a long way from meeting the target and the latter group will meet the target with less input from the school. In the UK education market, a variant on this basic idea is that while the test scores improvements of high achieving pupils do not contribute to the published pass rates, attracting such pupils improves the overall quality of the school body if peer groups matter, either in the production of school outcomes or for parents in their choice of school. This might limit the extent to which schools can focus on marginal pupils at the expense of high ability pupils. The incentives engendered by the focus on pass rates may be thus to divert resources away from very low achieving pupils.

To investigate the impact of this accountability standard on the distribution of pupil achievement, we analyse individual pupil value added as a function of the proportion of pupils in each school who are predicted to be marginal in terms of their contribution to their school's published performance on tests at age 16. We use a census of all children in the English state school system, which contains data on each child's performance in national tests taken at age 11, age 14 and age 16. These data enable us to control for prior ability, other characteristics of the child that may affect performance, and school peer group. We also examine school level data, comparing the performance of schools as a function of the number of students predicted to be marginal in terms of their contribution to the high stakes results.

We find evidence that the primary losers are the low and very low achieving students. As the number of marginal students in their school increases, these students lose both in terms of value added and their performance on the measure used to construct the school league tables. However, the gain of marginal students is small. Where the incentives for schools are sharpened by the potential for competition between schools in a local area, the patterns of gains and losses shift: the gain of the marginal pupils is larger and the loss of the very low ability is smaller. While the magnitude of these losses and gains is not that large, in terms of policy, we find that schools do respond to the short run incentives created by the measure used to assess school accountability, and that this response lowers the educational gain of the lowest ability students. This is not necessarily the effect desired by a government wishing to raise overall educational outcomes.

2. Related literature

2.1 The US findings

Most of the literature on school accountability is from the US, where the cornerstone of Federal educational policy has been expansion of school accountability based on measured student test performance. Some states link accountability to school performance, sanctioning poorly performing schools, whilst others sanction both poorly performing schools and poorly performing students[1]. School ratings can lead to

[1] An example of the former is the accountability system in place in Texas since 1994, the Texas Assessment of Academic Skills (TAAS). Under this system, schools are classified into four categories. Which category a school falls into depends on the fraction of students who pass achievement exams in reading, writing and maths. All students and separate student sub-groups must demonstrate pass rates that exceed year-specific standards for each category (Reback 2004). In addition, schools must have

organisational interventions, changes in school prestige, students transferring to other public schools, changes in local property values and financial rewards to schools and teachers (Reback 2004).

Assessments of the overall impact of accountability have been hampered by the fact that policies are introduced state wide, making comparison within states impossible, while cross-state comparisons are made difficult by the fact that the exact nature of policies differ considerably across states and are confounded by non policy differences that affect educational production. Hanushek and Raymond (2004) use state level data on maths and reading performances from the National Assessment of Educational Progress (NAEP) to assess the impact of accountability on performance. They examine growth in performance between 4^{th} and 8^{th} grade to control for fixed state differences in policies and circumstances, controlling for parental education, school spending, racial exposure in schools and state fixed effects. They find that the introduction of an accountability system leads to larger achievement growth than would have occurred without accountability. However, they also conclude that just reporting results has minimal impact on student performance and the force of accountability comes from attaching consequences to school performance.

A larger body of work has concentrated upon whether or not accountability has led to gaming and subsequent unintended consequences. Several studies have investigated whether schools react to accountability by increasing exclusions, identifying the effects of exclusions by comparing the introduction of accountability with the pre-accountability period within state. Jacob (2002) considers the introduction of test based accountability for Chicago public schools. He finds that the large increases in test scores after accountability went into effect were also accompanied by increases in special educational placement and by increased grade retentions. Deere and Strayer (2001) and Cullen and Reback (2002) find increases in special educational placements with the introduction of accountability in Texas. Figlio and Getzer (2002) focus on special education placement after the introduction of a state accountability system in

maintained drop out rates less than a certain level and attendance rates above a certain level. An example of the latter is the Chicago system in which both pupils and schools face sanctions if they do not reach certain standards. Pupils in 3^{rd}, sixth and eight grades are required to meet certain standards in order to advance to the next grade. Schools in which fewer than 15 percent of students scored at or above national norms are placed on probabation (Jacob 2004).

Florida. They find that placement rates increased relatively over time in grades that enter the accountability system as opposed to grades that do not. In contrast, Hanushek and Raymond (2004) analyse a state level panel. Once they also allow for the nationwide increase in placements over time they find no significant effect of having an accountability system. Other studies have looked at teacher cheating. Jacob and Levitt (2003), for example, look at unexpected test score fluctuations and unusual patterns of answers for students within a classroom as signs of cheating in the Chicago public school system. They find evidence that around 5% of classes in their sample cheated each year. Further, in classrooms with lots of low achieving students, cheating increased, but it did not in classrooms with average or high ability students. Researchers have also examined whether schools focus on teaching of subjects that count towards school performance. Deere and Strayer (2001) find that the passing rate on tests in Texas included in the school rating system (8^{th} grade tests in reading, writing and maths) increased at a higher rate that of Texas tests not used to determine school ratings (8^{th} grade tests in social studies and science). Jacobs (2004) finds in Chicago that maths and reading scores increased sharply on high stakes exams compared to the not high stakes exams after the introduction of the accountability policy.

In general, there has been less direct focus on the distributional effects amongst students of accountability policies (other than the examination of retentions and exclusions). Jacobs (2004) finds, where students as well as schools can be sanctioned, the impact on marginal students was less than on schools: students at risk of grade repetition only made higher gains in reading and not maths and students in the grades where forced retention may occur did not outperform students in other grades. Deere and Strayer (2001) find that students previously scoring near or below the passing score average a larger gain in scores than students previously scoring above the passing score.

The most relevant study for the present paper is Reback (2004). Reback argues that the Texas programme (TAAS) gives schools incentives to ‘teach to the rating’ and to shift resources towards marginal pupils and towards subjects that could best boost the school’s ratings. He analyses individual level test score data from Texas between 1994 and 1998. He exploits two types of discrete cut-offs used in Texas’s accountability system, the scores required for a student to pass the exams and the cut-offs for various standards that determine the schools rankings. Within a school, he examines whether students tend to make higher than normal improvements when small increases in the expected performance of these students and their classmates in the tested subjects would have a relatively large impact on the probability that a school earns a high rating. Across schools, he compares student performance at schools with and without strong immediate incentives to raise certain pass rates.

He finds that students whose performance could most influence their school's rating make better improvements than other students, but these effects are very small. He finds larger effects related to the fraction of students in the grade at the school whose performance could influence the schools ratings. Students who previously earned relatively low test scores make higher than expected improvements when their grade-mates may strongly influence the schools ratings. Further, the effects of having a greater fraction of classmates with positive accountability incentives are much stronger for lower achieving students.

2.2 The UK evidence

The volume of UK evidence on responses to increased accountability is far smaller[2]. A central component of the accountability system for state secondary schools, which educate 93 percent of children between the ages of 11 and 18, is the publication of information on the proportion of children gaining a certain number at a certain level of exam passes taken in key exams taken at the age of 16 (details are provided on these exams in Sections 3 and 5 below).

Performance – as measured by these raw exam scores – has certainly improved since the introduction of the present accountability arrangements in 1988. It is not clear to what extent such apparent improvements are due to a change in performance as opposed to changes in the difficulty of the test (Goldstein 2001: Tymms and Fitz-Gibbon 2001) and it is also difficult to isolate which elements of the programme of reforms instituted since 1988 have had an impact on outcomes (Bradley et al 2000).

[2] Evidence on the impact of school choice in the UK is growing: a review is provided by Burgess et al (2005).

Wilson (2004) identifies three possible types of responses to this type of target indicator. First, whenever an arbitrary dichotomy is introduced, it will focus agents' attention on the borderline (Fitz-Gibbon, 1996). Research has found these behaviours in practice include the use of volunteer helpers with weaker students, strategic mentoring by teachers, after-school coaching and holiday revision courses (West and Pennell, 2000). Wiggins and Tymms (2002) provide evidence of similar responses at primary school (age 5-11) level. Such practices may be employed in general if attendance on the course is compulsory. If this is not the case, the incentive would be to remove weak students from the course (Fitz-Gibbon, 1996). There is anecdotal evidence from England that schools are removing weak students from more difficult courses and putting them into other exams which count towards the pass rate but are easier or more vocational (Times Educational Supplement, 2002).

Second, there may be the incentive to reclassify weak students in order that they are not eligible for sitting the tests that are the subject of the target indicator. There is some indirect evidence that the publication of performance has created the incentive to exclude certain types of pupils from the school. Gillborn (1996) (quoted in West and Pennell, 2000) reports a tripling of permanent exclusions in the three year period from 1993/94 (the first year of publication of test results was 1992). Gerwitz *et al*. (1995) use the term 'constructive exclusion' to describe ways in which schools may put pressure on certain children in order that they leave 'voluntarily'. Third, schools have the incentive to also exclude weak students *ex ante*, i.e. to engage in selection or cream skimming at the point of admission. Gerwitz *et al*. (1995), Whitty *et al*. (1998) and West and Pennell (2000) discuss ways in which a school can design the procedure in order that only certain types of pupils (and parents) are attracted to the school.

In sum, while there is considerable discussion of the incentives, and limited qualitative evidence that schools target certain pupils in response to accountability standards based on summary test indicators, there has been no large scale investigation of the impact of such behaviour on pupils in English schools.

3. The impact of performance measurement on school behaviour

3.1 The English secondary school accountability system

The English education sector is large. In 2001, there were over 8.4 million pupils in 25,780 schools and over 400,000 full time teachers[3]. Schooling is compulsory up to age 16. State secondary schools, which educate children between the ages of 11 and 16 or 18, have operated within a 'quasi-market' system since 1988. There is local management of schools, with devolved budgets calculated on a per capita basis, overlapping catchment areas and open enrolment. To maintain resource levels, a school must attract sufficient pupil numbers; the overlapping catchment areas create the potential for competition for pupils (Glennerster 1991, Wilson 2004). The decision of which pupils to admit to a school is mainly made at the local educational authority (LEA) level, who set catchment areas for schools based on geography. So selection is primarily on location. However, a small number of LEAs (around 15 percent of all LEAs) allow selection on ability. Schools in this small number of LEAs are divided into grammar schools, which select on ability tests taken at the end of primary education and account for around 25 percent of schools in these LEAs, and the rest of the schools in these LEAs, which do not select on ability. In addition, a small number of schools within non-selective LEAs are allowed some control over selection and can select on criteria such as aptitude in certain subjects or religious faith[4] (West et al 2004).

Since 1992, state secondary schools have had to publish information on the proportion of children gaining 5 or more GCSEs at grade C or above. GCSEs are a national exam taken at age 16 (more details are provided in Section 5) and are used by employers to rate pupils who enter the labour market at age 16 and by schools and colleges to rate pupils entering post-GSCE education. Approximately 50 percent of pupils achieve 5 GCSEs at grade C or above, though there is considerable variation across schools in this outcome.

---

[3] Approximately 7 percent of children of secondary school age are educated in private schools (known as 'public' schools) for which parents pay full fees. We do not analyse the behaviour of these schools here.

[4] Approximately 5 percent of schools can select on some overt criteria.

The publication of this indicator and other summary indicators of performance (including absences) including information gathered by inspection[5] provides information on individual school performance. The percent of pupils passing 5 or more GCSEs at grade C or above, known as the 'school league tables', is the most widely disseminated piece of information. Better performance on the school league tables attracts more students or, if the school is at capacity, gives the school scope for selection (though grounds for selection and selection policies are set by the LEA). The league tables are also used by central government in determining whether a school is sanctioned for poor performance. Sanctions include 'name and shame' policies, replacement of management, and school closure. In summary, even if schools have other aims, the fact that the school league table system is the main source of information for parents and the government means schools have an interest in improving their ranking in the league table of 5 A-C grades at GCSE.

3.2 How might accountability affect the behaviour of schools with respect to their existing student body?

For student *i* in school *s* the score she achieves in the (high stakes) exams at age 16 will be determined by her ability, the school inputs she receives, peer and school effects, parental inputs plus noise. Value added at age 16 - the difference between performance in exam performance at 16 and some earlier date - will be:

$$(1) \quad y_{ist} = f(y_{ist-1}, R_{ist}, X_{ist}, Z_{st}, \varepsilon_{ist})$$

where y_{ist} is an achievement score for individual *i* in school *s* at time *t*, R_{ist} are the resources given to student *i*, X_{ist} is a vector of student characteristics, Z_{st} is a vector of school characteristics including measures of school peer group ability and ε_{ist} is noise.

A school rated on its league table performance will allocate resources across pupils to maximise some function of this performance and other goals it may have. The league table score at school level is non-linear in pupil performance: students only contribute to the school's performance in the league tables if they achieve grades equal to, or above, the passing threshold. Given (1), the probability of an individual student

[5] These other measures include in-depth reports on school performance conducted every 4 years or so, information on passing rates in tests taken at age 14 (known as key stage 3 tests).

achieving these grades depends on both on their prior performance and the school resources they receive. Students with low values of y_{ist-1} will require a high input of resources to hit the threshold, while those with higher y_{ist-1} will require fewer resources. In fact, unless the passing threshold is very high, pupils at the top of the distribution of y_{ist-1} may achieve the target without any resources from the school. This means that a school seeking to increase their performance against the target will focus resources on those students whose expected pass rate is close to, but below, the passing threshold. Unless the target is very high, these will be the pupils in the middle of the ability distribution. If only two levels of resources were given to pupils, the resource allocation would be of the form:

(2) $$R_{ist} = \begin{cases} R^h & \text{if } \boldsymbol{y}-\varepsilon < y_{ist-1} < \boldsymbol{y}+\varepsilon \\ R^l & \text{otherwise} \end{cases}$$

where $R^h > R^l$ and $\boldsymbol{y}$ is the t-1 achievement score which predicts performance just below the passing threshold.

The basic idea is depicted in Figure 1. The lower solid line represents the translation of prior ability into outcomes without any school resources; the upper solid line the effect of adding resources equally (under the assumption of a linear mapping from resources to outcomes). The dashed line indicates the impact of targeting resources at pupils close to the threshold; the effect is to raise the proportion of pupils whose outcome is above the target.

In England, as around 50% of pupils pass GCSEs, students who are in the top part of the ability distribution should get GCSEs at a C grade relatively easily. Conversely, students at the bottom of the distribution will not have much chance of getting a grade at or above C level. So, for a given budget, a school seeking to maximise their league table performance should focus resources on those students whose expected pass rate is close to the passing threshold: these will be the pupils in the middle of the ability distribution.

Given that schools can identify from prior scores which pupils are most likely to be close to this threshold, resources given to pupil i will be a function of their prior score at $t\text{-}1$: students whose expected scores, given their past scores, are close to the passing threshold will, for a given level of resources, get more resources than students whose expected scores are far away from the threshold. In addition, given that the pass rate is at school level and resources are limited, resources will also be a function of the distribution of achievement scores at time $t\text{-}1$ in the pupil body and in particular, in the number of marginal students. The more marginal pupils there are, the more it is worth the school focusing resources on these pupils. A higher proportion of marginal pupils will mean that a small increase in each of their scores will translate into a large increase in the school's passing rate. Thus the resources given to each pupil will depend on their past scores and the distribution of the pupil body:

$$(3) \qquad R_{ist} = f(y_{ist-1}, G(y_{st-1}))$$

where $G(y_{st-1})$ is the proportion of students who are marginal.

Which students benefit from this focus i.e. how this translates into outcomes, depends on how school resources are deployed and how resources impact on pupil attainment. While schools can identify marginal students and so choose to focus resources on them, the nature of the production process may be that there are externalities, such that resources directed towards one group benefit other students. If, at the limit, all pupils benefited equally from these spillovers, then a focus on marginal pupils would benefit all students (and, as a corollary, an examination of exam outcomes would not detect the resource focus on marginal students). However, provided that spillovers are less than full, then the targeted group should have gains in outcomes relative to other students.

As Figure 1 is drawn, the losers could be drawn from both ends of the ability distribution. But high ability students may be desirable to a school over and above their contribution to the current school league table position. High ability students may exert a positive peer group effect, making school resources more productive. Parents with high ability children are probably the marginal choosers in the English system[6]; such parents would not be attracted to schools where high ability children get fewer resources. Schools are rewarded on the basis of pupil numbers as well as pupil performance. So head teachers with longer term goals than the current league table performance may not wish to divert resources away from high ability pupils. We might therefore expect that the losers will be the very low ability students.

One other strategy a school might use to improve their measured performance is to enter students for exams that are easier to pass and/or count more towards the 5 passes required. Certain exams in the English curriculum are more vocational in nature and count towards more GSCEs points than others (see section 5). Therefore schools with high numbers of marginal pupils may do more of these exams.

Wilson et al (2004) interviewed around 20 headteachers, seeking their responses to the incentives given by the published league tables. It is clear that headteachers are not only acutely aware of their performance on this key measure, but also seek measures to improve their performance, conditional on the ability of the children in the school. Headteachers drew particular attention to the fact that, given the importance of the 5 A-C target, improvements in the performance of pupils who could improve school performance against the target were particularly important. These pupils are known as the C/D border since passes at a grade C or above count towards the schools target whilst those below do not. Response included:

> "We are very definitely targeting CD borderline pupils. That's the first thing we are doing and the second thing […] we introduced a GNVQ course which is a double award"
>
> "For example identifying the fragiles […], they are called borderliners in some schools […] youngsters who are going to get Ds, who with a bit of a push might get Cs […] we put in a separate form […], so if I get those 20 through along with my sort of 30 that we predicted to anyway that would be very nice thank you"

---

[6] If high ability is correlated with high income, such parents can afford to move into catchment area of good schools.

> "By choosing 5 or more A-Cs […] it's no surprise that most schools put a huge amount of energy and resource into those students who are on the C/D borderline"

4. Empirical approach

We focus on the effect of the incentive on pupil performance. Our primary interest is distributional: the impact on value added for individual pupils at different points in the ability distribution. But we also investigate whether there are impacts at school level.

The English reforms are not an experiment: all schools were subject to increased accountability in the form of published school outcomes. Data at pupil level do not exist for the period prior to the introduction of increased school accountability. So identification of the impact of the incentives must be made by examining the effects of cross sectional variation in the proportion of children in a school who are marginal pupils. However, the proportion of pupils who are marginal may also be a measure of the ability distribution of the children within the school. This means that analyses must control for this ability distribution. At school level, we are able to control for the ability of pupils at intake into secondary school, which is age 11.[7] At pupil level, we can allow for school fixed effects, which allows us to control both for the school ability distribution and also for any unobserved characteristics of the school that may be associated with the proportion of marginal pupils that there are in the school. We also distinguish situations in which these incentives may operate more sharply and test to see if outcomes are different in these situations[8].

We begin by using school level data for one cohort (pupils who took GCSE exams – the high stakes exams - in 2002) to see whether there is evidence that schools with high proportions of marginal pupils have different patterns of exam taking. We look to see whether schools with high proportions of marginal pupils have more value added, whether they get more pupils through the hurdle that makes up the

[7] Although we have data that allow us to control for school level ability at age 14, this will be collinear with the proportion of marginal pupils at age 14.
[8] In taking this approach we treat the ability distribution at ages 11 and 14 as exogenous. Our analysis therefore looks at the responses to a given stock of students. We also attempt to test this assumption below.

accountability standard, or they have students who take a higher proportion of their high stakes exams in the form of the easier GNVQ exams. We estimate at school level:

$$(4) \quad \% y_s = \beta \ \% cd_s + \alpha \ Z_s + \varepsilon_s$$

where *y* is one of three outcomes – the average value added between the ages of 14 and 16, or the proportion of pupils who hit the target (get 5 or more GCSEs at grade C or above), or is the proportion of exam entries that are accounted for by GNVQs. *Z* includes controls for the pupil type, including median ability at age 11 to control for peer group, and type of school.

We then turn to the effect of the incentive on value added at pupil level, controlling for other factors, including pupil level characteristics, and school characteristics including peer effects. We run OLS regressions, allowing for pupil clustering within school and then control for school fixed effects, so deriving estimates from the within school differences across pupils of different abilities.

We estimate models of the following form:

$$(5) \quad y_{is} = \beta_1 \% cd_s * I(low)_i + \beta_2 \% cd_s * I(cd)_i + \alpha_1 y_{ist-1} + \alpha_2 X_{is} + \alpha_3 Z_s + \varepsilon_{is}$$

where *y* is the outcome at age 16 for individual *i* in school *s*, *%cd* is the proportion of marginal pupils in the school, *I(low)* is an indicator variable taking the value of 1 if the ability of pupil I is low, *I(cd)* is an indicator variable taking the value of 1 if the ability of pupil I is marginal, y_{ist-1} is prior pupil attainment, *X* is a vector of student characteristics, *Z* is a vector of school characteristics including the measure of school peer group ability and geographical location. Our primary outcome measure is pupil level value added, but we also examine the probability that a pupil will achieve the target in the high stakes exams i.e. get 5 or more GCSEs at grade C or above.

The βs in equation (5) reflect the impact of the accountability policy – the impact of the percentage of marginal pupils on the attainment of the low and marginal pupils. Given the discussion above we expect low ability pupils to lose, but we have less clear priors to the effect of increasing the number of marginal pupils in the school on the performance of the marginal pupils themselves[9].

This estimation strategy relies on cross sectional variation in the proportion of marginal pupils. However, this proportion may be accompanied by other kinds of behaviour, which may counteract resource allocation aimed at improving school performance against the target. For example, as the size of the marginal group grows, there are fewer resources available for any one individual marginal student. This may dilute the impact of the resource allocation. Or the proportion of marginal pupils may have peer group effects that are not controlled for by measures of attainment of the schools pupils at 11. Or schools may be reluctant to take resources away from pupils who are high ability and have high prior scores. We therefore first examine whether the marginal pupils gain more when they increase relative to pupils with low prior scores (low ability pupils) compared to when they increase relative to pupils with high scores. We then examine the case where there are few high ability pupils so that the marginal pupils act as a positive peer group, so that if low ability pupils lose as the proportion of marginal pupils increase, this must indicate resource allocation, as the peer group effect will work in the opposite direction. Next, we seek to identify circumstance in which the incentives may be particularly sharp. The importance of performance on school league tables may be greater for schools that are in more competitive situations. These schools are likely to compete actively for pupils, and also may be closer to capacity, so are more able to select pupils. Those who are in less competitive positions will have less ability to select pupils and are more likely to be chosen by parents less interested in school performance. Improvements in performance may therefore matter more to the former than the latter. We estimate:

$$\text{(6)} \quad \begin{aligned} y_{is} = {} & \beta_1 \%cd_s * I(low_i) + \beta_2 \%cd_s * I(cd_i) + \\ & \beta_3 \%cd_s * I(low_i) * I(highcomp) + \beta_4 \%cd_s * I(cd_i) * I(highcomp) + \\ & \alpha_1 KS3_i + \alpha_2 X_{is} + \alpha_3 Z_s + \varepsilon_{is} \end{aligned}$$

where *I(highcomp)* is an indicator variable with value 1 if the school is in a competitive position. This is defined in several ways: details are provided below.

[9] In estimation, we actually distinguish between 5 groups of pupils based on prior attainment, the omitted group being pupils of highest ability.

Finally, as the distribution of pupil ability, and so the proportion of marginal pupils, may be endogenous, for example if schools seek to get better results so they can have fewer low ability pupils, we seek a method of dealing with this potential source of endogeneity. We cannot, in the absence of a natural experiment fully address this, but we make an attempt by instrumenting the distribution of pupils. There is evidence that within-year age of a pupil has an impact on outcomes, younger pupils performing worse than their older peers (see Alton and Masey (1998)), so we use month of birth as an instrument for the proportion of pupils who are marginal. The within-year age distribution predicts the proportion of marginal pupils well and is arguably less under the influence of schools than pupil ability. But it is clearly not a perfect instrument as, if month of birth is also correlated with measured ability at age 11, schools could select on ability and so on month of birth.

5. The data and details of the English exam system

The data we use is the Pupil Level Annual School Census (PLASC) dataset, combined with the National Pupil Database (NPD). This contains data on all pupils in both primary and secondary state schools in England, with approximately half a million pupils in each cohort. At pupil level, it provides linked histories of scores in national tests, plus some individual characteristics: gender and within-year age, ethnicity, eligibility for free school meals (FSM) (an indicator of low household income) and special educational needs. At school level, there is data on a range of school characteristics, including performance measures, geographical co-ordinates, school size, age range, religious denomination, funding status, gender mix and admissions policy.

Students in England take national tests (labeled key stages) at 4 ages: 7, 11, 14 and 16. At key stage 2 (taken at age 11) and key stage 3 (taken at age 14), pupils are tested separately in english, maths and science. At key stage four (KS4), when the pupils are 16, exams are taken in a variety of subjects. A majority of pupils take the GCSE exams, but some of these may be replaced or combined with vocational GNVQ qualifications. Pupils can be entered into different tiered papers according to ability in each exam, with the maximum achievable grade in lower tiers capped. The average number of exams taken by each pupil is nine. We use data for pupils who took their KS4 exams in the summer of 2002.

To define the proportion of marginal (CD border) pupils in a school, we define marginal as all pupils who have key stage 3 (KS3) scores for which any pupil has a predicted probability of between 40 – 60% of getting 5 A–C grades at KS4[10]. This group is 22% of all pupils[11]. From the rest of the average KS3 score distribution we define 4 groups: two below the marginal pupils and two above. Group 1 is the lowest ability group and is the bottom 10% of the KS3 average distribution. Groups 2 (32%) is the remainder of pupils between these very low ability pupils and the marginal pupils. Group 3 (24%) have scores between the marginal pupils and the very high ability. The top ability group accounts for 8% of pupils.

As well as the pupil level characteristics described above, we control for several school characteristics: cohort size, school gender, concentration of various pupil types, whether the school selects on ability (is a grammar school) and average school ability at age of entry (age 11) as a measure of peer effects. We also include dummies to control for whether the school is in London, a Non London Urban or Non London Rural area and local education authority (LEA) dummies capture any policy differences across LEAs, who are responsible for that part of education policy that is made at local level[12].

As noted above, most schools in England do not select on ability. However, a small proportion of LEAs allow selection on ability. Schools in LEAs with a relatively large concentration of selective schools will face different incentives to schools in other LEAs. We therefore define a selective LEA as having more than 10% of all schools selecting on ability and split our data into three school types: non selective schools in non selective LEAs (which excludes any selective schools located in non-selective LEAs), selective schools in selective LEAs and non selective schools in selective LEAs. Given selection policy, the concentration of marginal pupils will vary across selective and non selective schools. This can be seen in Figure 2, which

[10] Estimation results are robust to slightly narrower and slightly wider definitions of this band.
[11] This group, and the other groups, are not defined by exact deciles because of the discreteness of the KS3 average distribution.
[12] In our analyses we drop 3 LEAs (out of the 148) that have 4 or fewer schools.

presents the kernel estimate of the density of marginal pupil concentration for all schools and each school type. The mean proportion of marginal pupils is around 24 percent. The range is 0 to 1; schools at the 90th percentile have 30% percent of marginal pupils; schools at the 10th percentile have 15 percent. From the figure, it is clear that selective schools have fewer marginal pupils, while non selective schools in selective LEAs have a higher concentration of marginal pupils than non selective schools in non selective LEAs. Given this difference, the bulk of our analysis focuses on the dominant form of school – the non-selective school in a non-selective LEA. This type of school accounts for nearly 90 percent of all pupils.

This focus also reduces the potential for endogeneity of the distribution of pupil ability. Pupils in non-selective schools in non-selective LEAs will be assigned to school on the basis of geographical location rather than ability. It is true that pupils may move to be close to schools on the basis of the school's performance and this will mean that ability is not orthogonal to location, but the moving decisions of households are likely to have a smaller effect in the short term on the distribution of ability from which a school can draw (i.e. the ability distribution in their catchment area) than overt selection on ability.

We measure competition in three ways. The first is a count of the number of schools within a 10 minute drive time of the school that are above the national mean score at GCSE. This is a measure that combines both the quantity and the quality of schools in the vicinity, so picks up both density and the quality of the potential competitors. We use the top quartile of this distribution to identify those schools that are likely to engage in the most competition for pupils and parents. As a robustness test, we also use two other measures of competition. The first is a count of the number of schools within a 10 minute travel time of the school. This is simply a measure of the number of competitors, so is a measure of pupil density, rather perhaps than competition that is related to both density and quality. The second is a dummy variable that takes value 1 if a school is in the top quartile of their LEA league table. The majority of parents choose schools within the LEA in which they reside. Those schools that are in the top quartile of the distribution within LEA are likely to be the most sought after, and are therefore likely to engage in the most competition for pupils and parents[9].

The means and standard deviations of all variables used in the analysis are presented in Table 1.

6. Results

6.1 School level analyses

We first use data at school level to see whether there is evidence that schools with high proportions of marginal pupils have different patterns of exam taking. We estimate equation (2) to test whether schools with high proportions of marginal pupils have more value added, whether they get more pupils through the hurdle that makes up the accountability standard or they have students who take a higher proportion of their KS4 exams in the form of the more vocational GNVQ exams. These analyses are for the cohort of pupils who graduated in 2002. We include controls for pupil type at the school, including median ability at age 11 and type of school. To test whether this behaviour is sharpened where schools face greater competition, we repeat the analysis allowing for an interaction term with competition (defined as a dummy variable which measures the number of good schools in the 10 minute drive time).

Table 2 presents these results for all schools. The first two columns are for value added, the second two for performance against the accountability standard and the third for the proportion of exams made up by the more vocational GNVQs. Column 1 shows that as the proportion of marginal pupils increases so does school level value added. Column 2 shows that this effect is increased if schools are in competitive positions, although insignificant. Column 3 shows that performance against the accountability standard does not rise as the proportion of marginal pupils increases,

---

[9] Schools may be interested in improving their position in the national and/or their within LEA ranking. In favour of the national ranking, information is published widely each year (in national newspapers) on the ranking of the top 300 or so schools in the country. For these schools, and those close to this set, national rankings are probably the most important. In favour of local rankings, there are around 3000 secondary schools in England, and children almost always attend a school within LEA. Schools' positions in the national ranking is closely correlated with their positions in the within LEA rankings (corr=0.83)

but column 4 shows that where schools are in a competitive position the proportion of marginal pupils is associated with an increase in performance against target. Column 5 shows no overall effect on the proportion of more vocational exams taken and this remains unaffected by increases in competition. So these results show some support for an increase in performance as the number of marginal pupils in the school increases, particularly if schools are in a competitive position.

These school level analyses suffer from the fact that even with controls for pupil ability at age 11, the proportion of marginal pupils may be measuring pupil ability as well as the effect of having more marginal pupils. In addition, while they show school behaviour, they do not show whether there is differential impact across different types of pupils. To do this, we need to use pupil level data.

6.3 The impact on individual student value added

We focus on whether the proportion of marginal pupils affects the educational value added of different pupils. We begin by looking at the raw data. Figure 3 plots the value added (the difference between test scores in the high stakes age 16 and lower stakes age 14 exams) of students grouped into 5 groups on the basis of age 14 test scores. This is then normalised by the average school level mean value added to allow for differences in the levels of average VA across schools. The normalised value added for each of the 5 student ability groups is graphed for all 5 groups in two sets of schools: schools in the bottom quartile of the distribution of marginal (C/D border) students and schools in the top quartile. The figure shows that the value added, relative to the school average, of the bottom ability group is always positive, but is lower where these students are in a high marginal student environment. In contrast, the value added of students in the marginal group and above them is greater where the marginal students are a larger fraction of the student body. The raw data therefore provides some support for a change in the distribution of value added across students as the proportion of marginal pupils increases. The group that appear to be losing are the students at the bottom of the ability distribution.

To test this more formally we begin by estimating equation (5) with the dependent variable defined as the total points scored by the pupil at KS4. The controls include the pupil's score at KS3 (so the outcome is interpretable as value added between age 14 and 16), plus their ability group (the 5 groups defined by the KS3 score), the school peer group (as defined by the median ability of pupils at age 11) plus the other pupil and school controls listed in table 1 other than month of birth[13].

The first column of table 3 presents the coefficients of the interactions of the proportion of marginal pupils with pupil ability. The omitted group is the highest ability pupils. All the estimates include dummies for pupil ability groups and school KS2 median (not reported in the table). The results indicate that as the proportion of marginal pupils increases, the value added of all pupils relative to the top ability group falls, the fall being greatest for the lowest ability pupils and smallest for those pupils just below the top ability group.

To test whether this effect is sharpened where schools face greater competition, column 2 allows for an additional interaction term for schools located in areas of high competition. The estimates indicate that where competition is higher, the lowest ability pupils and those just above them gain relative to the high ability group as the proportion of marginal pupils increases. Column (3) repeats this, but focusing instead on schools which face little competition. In these schools, the lowest ability pupils now lose relative to the top ability group as the number of marginal pupils increases.

Columns 4-6 present the same specification, but allowing for school fixed effects, so that the impact of the proportion of marginal pupils is identified from within school variation[11]. The results confirm the analyses of columns 1-3: as the proportion of marginal pupils increases, all pupils lose relative to the most able, but the lowest ability group loses most. Strong competition ameliorates this loss for all ability groups; the lowest ability pupils gaining the most. Lack of competition hurts the lowest ability pupils the most at the margin.

[13] The Department for Education and Skills produces an age 14 to 16 value added measure. Our results also carry through using this measure

[11] All school level variables, including the proportion of marginal pupils and the school peer group effect, become part of the estimated school fixed effect.

As most schools are non-selective, located in non-selective LEAs, for the rest of the analysis we focus on this group, which accounts for 87% of all pupils. This focus reduces the possible endogenity. Columns 7,8 and 9 present fixed effect results for these schools corresponding to columns 4-6. Only fixed effects estimates are presented, as these are very similar to results without school fixed effects. The results reflect those for the full sample: the lowest ability group lose the most as the proportion of marginal pupils increases. In competitive environments, all ability groups gain relative to the high ability, the lowest ability gaining most at the margin. When competition is low, this group loses out most as the proportion of borderline pupils increases.

While statistically significant, the size of these effects is not large. Table 4 presents the estimated effect of moving, for the different ability groups, from the 10th to the 90th decile of the distribution of marginal pupils. This uses the estimates for non-selective schools in non-selective LEAs from Table 3[12]. The first row of the table presents the estimates for all such schools, the second for such schools in the top quartile of the number of good schools within a 10 minute travel zone, the third for schools in the bottom quartile. The estimates indicate that all pupils lose to the most able group as the proportion of borderline students in a school increases. But the size of the loss is small. In percentage terms the loss is at most just under 15 percent, and in absolute terms it is around 1 GCSE point, which is equivalent to 1/20th of a standard deviation in the GCSE score distribution for this group. Where there is high competition, the value added of all groups of pupils rises relative to the most able as the proportion of marginal pupils increases: again the size of these gains is small.

6.4 The impact on hitting the school league table target

Table 5 repeats the analysis of table 3 using the individual probability of hitting the target (5 GCSEs at grade C or above) as the outcome variable. The estimates include the same pupil and school controls as the value added analyses above. The patterns for non selective schools are similar to those for value added: as the proportion of marginal pupils rises the chances that low ability pupils will hit the target falls. If a

[12] The estimates for each ability group are calculate using the characteristics of the mean student in that ability group.

school is in the top quartile of the competition distribution, this raises the chances of pupils in the lowest two ability groups relative to the top ability group.

Again, these effects are not large. Table 6 shows the estimated effects for pupils in the different ability groups. The top panel shows that the net impact of moving from the 10^{th} percentile to the 90^{th} percentile of the proportion of marginal pupils has a positive effect on the chance of hitting the threshold for the marginal pupils. But the increase in probability is again small, being only 3 percent, which is equivalent to around a 25th of the standard deviation of the probability of 5 A-Cs for this group. Again, the losers, as the proportion of borderline pupils increases in a school, are the lower ability students.

6.5 At whose expense do marginal pupils gain?

The results so far indicate that as the proportion of marginal pupils increases compared to all pupils, the losers are the low ability pupils. Marginal pupils may increase in a school relative to either low ability or high ability pupils (or to both). But if schools do not wish to harm high ability pupils, the focus of resources on marginal students may be less as this group increases relative to high ability pupils compared to where this group increases relative to low ability pupils. Table 7 compares the impact of an increase in marginal pupils relative to the two groups above on the value added for each pupil type with the impact of an increase in marginal pupils relative to the two groups below them. The results are presented for non-selective schools in non-selective LEAs only. The first column indicates that as the fraction of marginal students to those above them increases, there is little effect on the value added of any pupil type. In contrast, column 2 shows that as the marginal group increases relative to the lower ability groups, the groups in the middle of the ability distribution gain relative to both groups at the extremes, and the biggest gain is by the marginal students. These results suggest that marginal students only gain where they increase relative to low ability pupils. This may be indicative that schools do not wish to push resources towards marginal pupils as their numbers grow large relative to high ability pupils.

6.6 Robustness checks

Marginal pupils are roughly midway along the distribution of pupils within schools. It is possible that the interaction terms in the proportion of marginal pupils are picking up some complex negative peer group effect (on top of the school fixed effects that includes the effects of average school ability at entry (age 11) and the proportion of marginal pupils). To examine this, we re-estimated using only those schools where the marginal pupils were at or near to the top of the distribution of pupil ability. In this case, the marginal pupils should act as a positive peer group for all groups of pupils, so the interaction terms cannot be interpreted as a negative peer group effect caused by an increase in the number of marginal pupils. Table 8 presents results for only those non-selective schools in non-selective LEAS where the marginal pupils form the top of the school ability distribution at age 14. Columns 1 and 2 present the value added results; columns 3 and 4 present the results for hitting the target. The pattern of coefficients for each ability group is the same as for the much larger sample of all schools: as the proportion of marginal pupils increases the lower ability pupils lose most. These differences are not statistically significant for value added. The impact of being in a school where competition is low is similar to all other schools: as the proportion of marginal students increases in schools where there is low competition, the lowest ability pupils lose the most in terms of value-added and hitting the target.

The final four columns of Table 8 present estimates using the two other measures of competition. The first is based on the number of schools within a 10 minute travel time of the school, so is a measure of school and pupil density. The second is based on a school's position in its own LEA in terms of the 5 A-C league table. Both measures show that as the number of marginal pupils increases the lowest ability group loses most. The results in columns 4 and 5 are very similar to those when competition is defined using the number of good schools in a 10 minute travel zone: as the number of marginal pupils increases, the loss of the lowest ability is ameliorated where competition is highest and is worsened where competition is lowest. In columns 4 and 5, where high competition is defined as being in the top quartile of the LEA distribution, competition significantly improves the position of the marginal pupils as the proportion of marginal pupils also increases. In schools with little competition, the marginal pupils lose, but the lowest ability group gain.

If a school's behaviour with respect to marginal pupils attracts pupils to the school, then the distribution of pupil type within schools will be endogenous. We use month of birth as an instrument for the proportion of pupils who are marginal. The first column of Table 9 presents the first stage of this instrumental variables (IV) approach, which shows that a higher proportion of younger pupils within year is significantly associated with a higher proportion of marginal pupils. Column 2 reports the second stage of the IV estimates and Column 3 presents the non-IV estimates for comparison. Our previous results remain robust to the use of the instrument: low ability pupils lose the most as the number of marginal pupils increases.

Discussion

This paper examines the distributional impact of a central element in the English school accountability policy. The use of threshold passing rates gives schools incentives to target the marginal pupils: the paper asks who gains and who loses from this. Taking the ability distribution of pupils in a school as given, we first show that as the proportion of marginal pupils at school level increases, so does the performance of that school in terms of value added at school level. Performance is increased particularly where schools face the added incentive of being in competition with other local schools for pupils. Pupil level analyses make it clear that the losers from this are the lowest ability pupils. In almost all types of school, as the number of marginal pupils increases, these pupils get less value added and have a lower chance of getting the qualification required for post-16 study and access to reasonable school leaver jobs. The marginal pupils themselves do not appear to gain, except where the schools they attend are in competition with other schools for pupils. In this case they do gain. The pattern for performance in the high stakes exams at pupil level follows that for value added: the low ability pupils lose the most as the proportion of borderline pupils in the school rises.

Whilst these effects are present, their quantitative impact is quite small. Moving from the 10^{th} to the 90^{th} decile of the marginal pupil distribution for the lowest ability student lowers the total number of points achieved in the age 16 exams by just under 1 point. In comparison, coming from a low income household (as measured by being eligible for free schools meals) has a marginal negative effect of 3 points achieved in the age 16 exams; being male has a similar negative impact.

Recent evidence from the US which has examined the distributional effect of accountability at pupil level also finds that the relative importance of a student's performance within a school has only a very small, positive, effect on that student's performance relative to his or her peers (Reback 2004). Reback (2004) does, however, find larger distributional effects when schools have a strong incentive to improve the performance of many students in a grade, but even this effect is not large in quantitative terms. The incentives in the English system are probably rather weaker than in the American states that have been heavily studied. While the school league table is undoubtedly important to headteachers, as it determines their standing and that of their school, there are no clearly identified thresholds at which schools get penalised or rewarded as in the Texan or Chicago public school systems. In addition, there is less focus on explicit financial rewards in the English system. So while marginal pupils do appear to affect the pattern of value added it is perhaps not surprising that this effect is small.

Finally, the fact that marginal pupils do not themselves gain unless the incentives are ramped up by competition perhaps suggests something about the kind of behaviour undertaken by schools in response to the policy. The average school's policies to target this group may be diffuse rather than focused. Higher ability students may be more able to profit from these activities, whilst the less able ones lose. However, where incentives are sharper, the marginal pupils do gain in terms of value-added, suggesting that in these schools the activities undertaken target this group rather than the high ability students.

References

Alton, A and Massey, A (1998), Date of Birth and Achievement at GCSE and GCE A-Level, Educational Research, 40(1): 105-109

Bradley, S., Crouchley, R., Millington, J. and Taylor, J. (2000), Testing for quasi-market forces in secondary education. *Oxford Bulletin of Economics and Statistics*, 62, 3, pp. 357-390.

Burgess, S, Propper, C and Wilson, D (2005) Will more choice improve outcomes in education and health care: the economic evidence. University of Bristol, CMPO, April 2005.

Cullen, J B and Reback, R (2002) Tinkering towards accolades: School Gaming under a performance accountability system. Department of Economics, University of Michigan (mimeo).

Deere, D and Strayer, W (2001) Putting Schools to the Test: School Accountability, Incentives and Behaviour. Working Paper, Department of Economics, Texas A and M University.

Figlio D N and L S Getzler (2002) Accountability, ability and disability: Gaming the system? NBER W9370 (November)

Fitz-Gibbon, C.T. (1996), *Monitoring Education: Indicators, Quality and Effectiveness* (Cassell, London).

Fitz-Gibbon, C.T. and Tymms, P. (2002), Technical and ethical issues in indicator systems: doing things right and doing wrong things. *Education Policy Analysis Archives*, 10, 6, http://epaa.asu.edu/epaa/v10n6, 12/02/02.

Gerwitz, S., Ball, S.J. and Bowe, R. (1995), *Markets, Choice and Equity* (Oxford University Press, Milton Keynes).

Gillborn, D. (1996), *Exclusions from school.* Viewpoint Number 5, Institute of Education, University of London.

Glennerster, H. (1991), Quasi-markets for education. *Economic Journal*, 101, pp. 1268-1276.

Goldstein, H. (2001), Using pupil performance data for judging schools and teachers: scope and limitations. *British Educational Research Journal*, 27, 4, pp. 433-442.

Hanushek E and Raymond M (2004) Does School Accountability lead to improved student performance? NBER Working Paper 10591

Jacob B A and Levitt S D (2003) Rotten apples: An Investigation of the Prevalence and Predictors of Teacher Cheating. *Quarterly Journal of Economics* CXVIII (3): 843-878.

Jacob, B (2004) Accountability, Incentives and Behaviour: The Impact of High Stakes Testing in Chicago Public Schools. Kennedy School of Government, Harvard University mimeo (January).

Reback, R (2004) School Accountability and the Distribution of Student Achievement. Barnard College Economics Department (mimeo)

Times Educational Supplement (2002), *League table bonus attracts schools to vocational option*, 23 August.

Tymms, P. and Fitz-Gibbon, C. (2001), Standards, achievements and educational performance: a cause for celebration? In Phillips, R. and Furlong, J. (Eds), *Education, Reform and the State: Twenty-five Years of Politics, Policy and Practice* (Routledge, London), pp. 156-173.

West, A., Hind, A. and Pennell, H. (2004) School Admissions and 'selection' in comprehensive schools: policy and practice. *Oxford Review of Education* vol. 30 no. 3 pp. 347 – 369

West, A. and Pennell, H. (2000), Publishing school examination results in England: incentives and consequences. *Educational Studies*, 26, 4, pp. 423-436.

Whitty, G., Power, S. and Halpin, D. (1998), *Devolution and Choice in Education: the School, the State and the Market* (Open University Press, Milton Keynes).

Wiggins, A. and Tymms, P. (2002), Dysfunctional effects of league tables: a comparison between English and Scottish primary schools. *Public Money and Management*, 22, 1, pp. 43-48.

Wilson, D (2004) Which Ranking? The Impact of a Value_added measure of Secondary School Performance. *Public Money and Management,* 24, 37-45.

Wilson, D et al (2004) What Gets Measured Gets Done: Headteachers' Responses to the English Secondary School Performance Management System. University of Bristol CMPO Discussion Paper 04/107.

Figure 1: Model of relation between test score output and prior ability

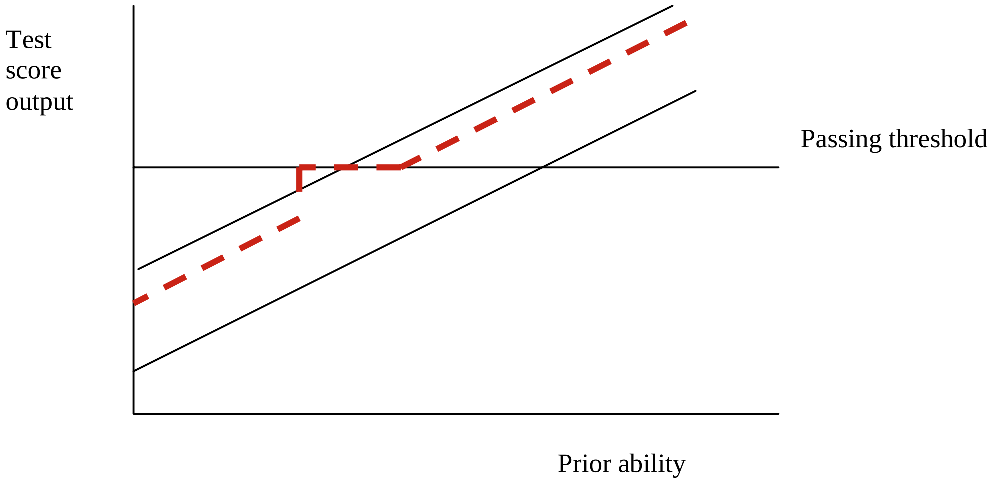


Figure 2: Concentration of marginal pupils by school type

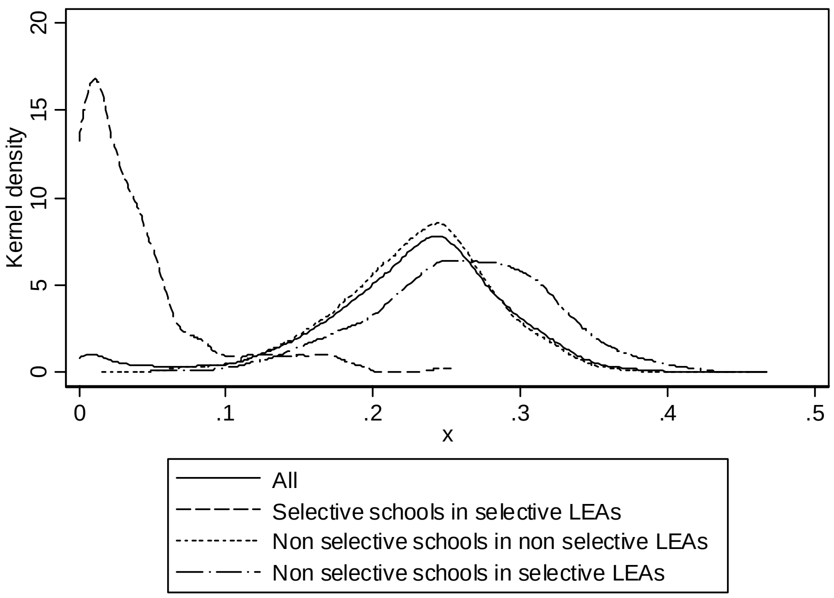
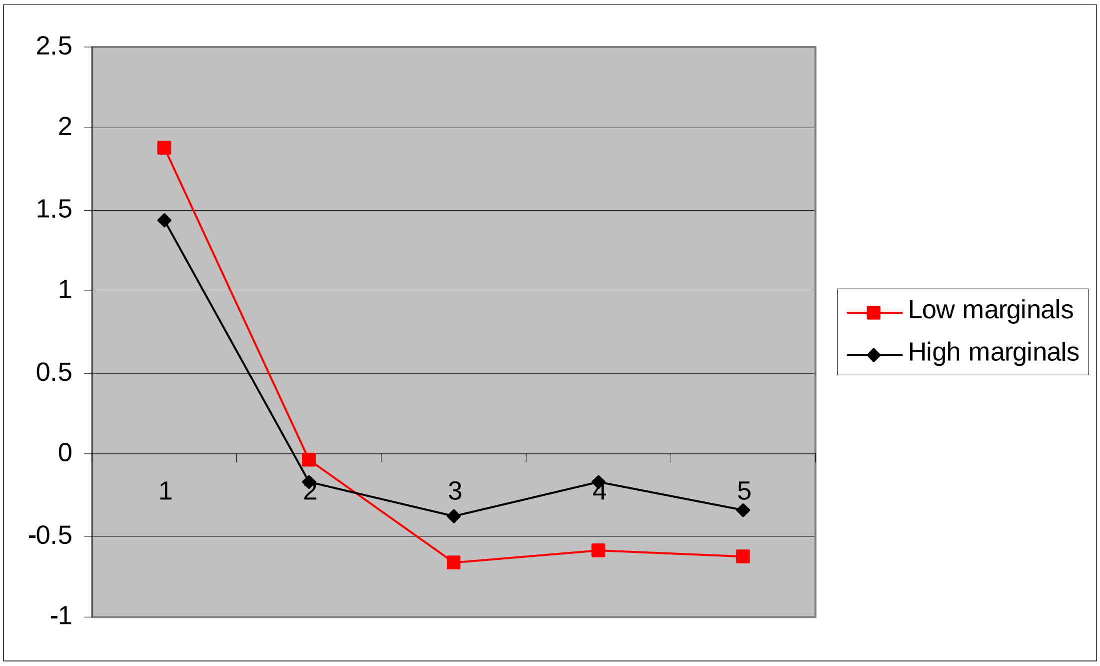


Figure 3: Total VA relative to school mean VA by KS3mn band



Notes

Low marginals are the bottom quartile and high marginals are the top quartile of schools classified by the across school distribution of cd border pupils.

Table 1: Means and Standard Deviations

Variable	**Mean**	**Std. Dev.**
Pupil characteristics (n=560209)		
Lowest ability (Group 1) dummy	0.102	0.302
Group 2 dummy	0.315	0.465
Marginal pupil dummy	0.218	0.413
Group 3 dummy	0.243	0.429
Highest ability dummy	0.080	0.272
KS3 mean	32.924	6.661
Total point score GCSE/GNVQ (dependent variable)	40.759	19.242
Male dummy	0.507	0.500
FSM dummy	0.135	0.342
English as a second language dummy	0.085	0.279
Ethnic minority dummy	0.182	0.386
SEN (non statement)	0.145	0.352
SEN (with statement)	0.024	0.154
Born in September	0.083	0.275
Born in October	0.081	0.273
Born in November	0.075	0.264
Born in December	0.074	0.262
Born in January	0.079	0.269
Born in February	0.072	0.259
Born in March	0.083	0.275
Born in April	0.081	0.272
Born in May	0.085	0.279
Born in June	0.082	0.274
Born in July	0.083	0.276
Born in August	0.084	0.277
School characteristics (n=3092)		
Proportion of marginal pupils	0.223	0.071
Median KS2 score	25.994	2.001
Boys school dummy	0.059	0.235
Girls school dummy	0.073	0.261
Selective school dummy	0.053	0.224
Cohort size	181.209	60.582
Percentage of students with special needs (SEN)	0.181	0.116
Percentage of students low income (eligible for FSM)	0.148	0.135
Percentage of students with english as a second language	0.091	0.183
Percentage of students from an ethnic minority	0.180	0.226
London	0.130	0.336
Non London Urban	0.608	0.488
Non London Rural	0.262	0.440

Source: Department for Education and Skills

Table 2: School level analyses of impact of % marginal pupils

Explanatory variables	Value added		% 5 A-C		GNVQ/GCSE ratio	
% Marginal pupils	3.447	3.297	-0.006	-0.022	-0.004	-0.006
	(2.43)**	(2.30)**	(0.16)	(0.55)	(0.24)	(0.30)
% Marginal pupils * Top quartile of competition		0.774		0.082		0.006
		(0.73)		(2.71)***		(0.44)
School controls						
KS2 median			0.051	0.051	-0.001	-0.001
			(29.22)***	(29.26)***	(0.79)	(0.79)
KS3 mean	2.720	2.719				
	(64.43)***	(64.38)***				
Cohort Size	0.002	0.002	0.000	0.000	0.000	0.000
	(1.50)	(1.55)	(2.94)***	(3.11)***	(0.46)	(0.49)
Percentage FSM	1.861	1.853	-0.477	-0.476	0.061	0.061
	(1.61)	(1.61)	(15.79)***	(15.79)***	(4.28)***	(4.28)***
Percentage ESL	3.258	3.260	0.138	0.138	-0.012	-0.012
	(4.89)***	(4.89)***	(7.31)***	(7.32)***	(1.35)	(1.35)
Percentage SEN	-1.839	-1.833	-0.182	-0.181	0.021	0.021
	(2.48)**	(2.47)**	(8.77)***	(8.72)***	(2.12)**	(2.13)**
Percentage ethnic minority	1.564	1.556	0.029	0.028	0.004	0.004
	(2.80)***	(2.78)***	(1.80)*	(1.75)*	(0.59)	(0.58)
Selective school dummy	-2.547	-2.517	0.144	0.147	-0.016	-0.016
	(4.16)***	(4.10)***	(8.63)***	(8.77)***	(2.03)**	(2.01)**
Boys school dummy	-0.634	-0.648	0.007	0.006	-0.003	-0.003
	(2.04)**	(2.08)**	(0.80)	(0.63)	(0.80)	(0.82)
Girls school dummy	1.789	1.781	0.077	0.076	-0.005	-0.006
	(6.25)***	(6.22)***	(9.55)***	(9.44)***	(1.44)	(1.46)
London	0.346	0.202	0.093	0.077	-0.022	-0.023
	(0.18)	(0.10)	(1.66)*	(1.38)	(0.83)	(0.87)
Non London Rural	0.194	0.204	0.020	0.021	-0.006	-0.006
	(1.01)	(1.06)	(3.72)***	(3.90)***	(2.39)**	(2.35)**
Constant	-51.094	-51.054	-0.851	-0.850	0.034	0.034
	(23.04)***	(23.01)***	(12.67)***	(12.67)***	(1.08)	(1.08)
Observations	3071	3071	3071	3071	3071	3071
R-squared	0.88	0.88	0.78	0.78	0.14	0.14

Notes:

1. Standard errors in parenthesis. * Significant at 10%; ** significant at 5%; *** significant at 1%
2. Regressions include LEA dummies
3. Competition is a dummy variable taking value 1 if the schools in the 10-minute drive time are in the top quartile of the distribution of good schools.

Table 3: Value-added at pupil level

	OLS			Fixed effects			Fixed effects		
Lowest ability * % marginal pupils	-15.995	-15.915	-14.351	-8.680	-8.646	-6.767	-9.349	-8.509	-7.308
	(2.572)***	(2.570)***	(2.614)***	(1.109)***	(1.110)***	(1.181)***	(1.545)***	(1.547)***	(1.629)***
Group 2 * % marginal pupils	-9.375	-9.339	-8.594	-4.943	-5.116	-4.043	-5.330	-4.871	-4.589
	(2.236)***	(2.231)***	(2.278)***	(0.874)***	(0.875)***	(0.924)***	(1.352)***	(1.352)***	(1.410)***
Marginal * % marginal pupils	-5.233	-5.166	-5.204	-3.332	-3.576	-3.576	-1.980	-1.739	-2.442
	(1.937)***	(1.930)***	(1.969)***	(0.883)***	(0.885)***	(0.933)***	(1.376)	(1.376)	(1.434)*
Group 3 * % marginal pupils	-3.647	-3.479	-3.589	-2.306	-2.552	-2.664	-1.330	-1.130	-1.991
	(1.262)***	(1.263)***	(1.326)***	(0.624)***	(0.627)***	(0.681)***	(1.336)	(1.336)	(1.392)
Lowest ability * % marginal pupils		5.616			9.659			10.151	
* Top Quartile of competition		(1.213)***			(0.849)***			(0.930)***	
Group 2 * % marginal pupils		3.458			8.081			7.711	
* Top Quartile of competition		(1.102)***			(0.712)***			(0.776)***	
Marginal * % marginal pupils		0.785			5.390			5.054	
* Top Quartile of competition		(1.151)			(0.719)***			(0.787)***	
Group 3 * % marginal pupils		-1.214			3.500			2.892	
* Top Quartile of competition		(1.182)			(0.712)***			(0.777)***	
Lowest ability * % marginal pupils			-2.540			-3.048			-2.956
* Bottom Quartile of competition			(0.744)***			(0.604)***			(0.639)***
Group 2 * % marginal pupils			-1.308			-1.603			-1.218
* Bottom Quartile of competition			(0.706)*			(0.481)***			(0.505)**
Marginal * % marginal pupils			-0.068			-0.110			0.011
* Bottom Quartile of competition			(0.760)			(0.485)			(0.509)
Group 3 * % marginal pupils			-0.086			0.380			0.673
* Bottom Quartile of competition			(0.809)			(0.479)			(0.502)
Constant	-49.975	-49.816	-50.018	-26.123	-26.158	-26.199	-26.762	-26.805	-26.837
	(2.326)***	(2.326)***	(2.328)***	(0.344)***	(0.344)***	(0.345)***	(0.374)***	(0.374)***	(0.375)***
Observations	508276	508276	508276	511055	511055	508276	441677	441677	439353
R squared	0.73	0.73	0.73	0.70	0.70	0.70	0.71	0.71	0.71

Notes:
1. Standard errors in parenthesis. Clustered at school level for OLS regressions. * Significant at 10%; ** significant at 5%; *** significant at 1%
2. OLS regressions include ability dummies, pupil and school level controls as in Table 1 and LEA dummies
3. Competition is a dummy variable taking value 1 if the schools in the 10-minute drive time are in the top quartile of the distribution of good schools.

Table 4: Estimated value-added for different ability groups

Panel A: Non selective schools in non selective LEAs

	Lowest ability		Group 2		Marginal		Group 3		High ability	
	Low	High	Low	High	Low	High	Low	High	Low	High
Without competition	10.87	9.52	30.82	30.05	43.22	42.93	56.24	56.05	70.39	70.39
In Top Quartile	11.99	12.23	31.69	32.10	43.80	44.27	56.56	56.81	70.36	70.36
In Bottom Quartile	10.55	9.07	30.69	29.85	43.23	42.88	56.30	56.11	70.39	70.39

Table 5: Probability of getting 5 A-Cs (probit analysis)

	Non selective schools in non selective LEAs		
Lowest ability * % marginal pupils	-4.150 (1.048)***	-4.173 (1.052)***	-4.125 (1.098)***
Group 2 * % marginal pupils	-2.153 (0.739)***	-2.136 (0.737)***	-2.023 (0.741)***
Marginal * % marginal pupils	-1.120 (0.708)	-1.132 (0.707)	-1.132 (0.709)
Group 3 * % marginal pupils	-0.836 (0.696)	-0.851 (0.695)	-0.946 (0.698)
Lowest ability * % marginal pupils		0.377 (0.449)	
Group 2 * % marginal pupils* Top Quartile of competition		0.362 (0.156)**	
Marginal * % marginal pupils* Top Quartile of competition		-0.070 (0.139)	
Group 3 * % marginal pupils* Top Quartile of competition		-0.392 (0.152)***	
Lowest ability * % marginal pupils* Bottom Quartile of competition			-0.059 (0.367)
Group 2 * % marginal pupils * Bottom Quartile of competition			-0.179 (0.095)*
Marginal * % marginal pupils* Bottom Quartile of competition			0.011 (0.083)
Group 3 * % marginal pupils * Bottom Quartile of competition			0.150 (0.087)*
Constant	-11.002 (0.333)***	-11.006 (0.333)***	-11.004 (0.332)***
Observations	439353	439353	439353
Pseudo-likelihood	-141457.92	-141421.63	-141443.26

Notes:
1. Clustered standard errors at school level
2. Standard errors in parenthesis. * Significant at 10%; ** significant at 5%; *** significant at 1%
3. Includes ability dummies, pupil and school level controls (except a selective school dummy) as in Table 1 and LEA dummies

Table 6: Predicted probability of getting 5 A-C grades for different ability groups

Panel A: Non selective schools in non selective LEAs

	Lowest ability		Group 2		CD border		Group 3		High ability	
	Low	High	Low	High	Low	High	Low	High	Low	High
Without competition	0.00	0.00	0.13	0.11	0.63	0.66	0.98	0.98	1.00	1.00
In Top Quartile	0.00	0.00	0.14	0.13	0.63	0.65	0.97	0.98	1.00	1.00
In Bottom Quartile	0.00	0.00	0.13	0.11	0.63	0.66	0.98	0.98	1.00	1.00

Table 7: Marginal pupil performance relative to other ability group

	Non Selective Schools in Non Selective LEAs	
Lowest ability*(% marginal / % top)	0.186	
	(0.176)	
Group 2*(% marginal / % top)	0.029	
	(0.172)	
Marginal *(% marginal / % top)	-0.043	
	(0.174)	
Group 3*(% marginal / % top)	0.043	
	(0.176)	
Lowest ability*(% marginal / % bottom)		-0.367
		(0.228)
Group 2*(% marginal / % bottom)		0.728
		(0.147)***
Marginal *(% marginal / % bottom)		0.858
		(0.137)***
Group 3*(% marginal / % bottom)		0.451
		(0.126)***
Constant	-26.662	-26.537
	(0.377)***	(0.375)***
Observations	441461	441502
R-squared	0.71	0.71

Notes:
1. Standard errors in parenthesis. * Significant at 10%; ** significant at 5%; *** significant at 1%
2. Regressions include ability dummies and pupil level controls as in Table 1.
.

Table 8: Robustness checks

Explanatory variables	Schools where marginal pupils are the top of the pupil distribution				Competition as number of schools		Competition as position of school in the school % 5 A-C distribution within LEA	
	Value-added (FE)		Probit 5 A-C		Value-added (FE)		Value-added (FE)	
Lowest ability * % marginal pupils	-8.629	-5.387	-2.021	-1.765	-8.555	-7.836	-9.315	-8.862
	(5.404)	(5.632)	(1.440)	(1.472)	(1.549)***	(1.607)***	(1.569)***	(1.554)***
Group 2 * % marginal pupils	-6.498	-5.193	-2.824	-2.874	-5.028	-4.876	-7.145	-5.517
	(5.128)	(5.334)	(1.146)**	(1.157)**	(1.352)***	(1.389)***	(1.365)***	(1.360)***
Marginal * % marginal pupils	-0.605	-2.018	-1.227	-1.377	-2.066	-2.386	-3.703	-2.956
	(5.625)	(5.838)	(1.116)	(1.194)	(1.376)	(1.413)*	(1.390)***	(1.386)**
Group 3 * % marginal pupils					-1.413	-1.914	-2.013	-1.494
					(1.335)	(1.370)	(1.345)	(1.342)
Lowest ability * % marginal pupils* Top Quartile Comp					8.050		-0.250	
					(0.859)***		(0.791)	
Group 2 * % marginal pupils * Top Quartile Comp					6.057		4.800	
					(0.728)***		(0.529)***	
Marginal * % marginal pupils * Top Quartile Comp					3.883		4.542	
					(0.739)***		(0.524)***	
Group 3 * % marginal pupils * Top Quartile Comp					3.024		2.382	
					(0.735)***		(0.510)***	
Lowest ability * % marginal pupils* Bottom Quartile Comp		-7.013		-0.907		-2.714		4.163
		(3.452)**		(1.372)		(0.667)***		(0.954)***
Group 2 * % marginal pupils * Bottom Quartile Comp		-2.823		0.108		-1.032		-0.607
		(3.218)		(0.650)		(0.512)**		(0.863)
Marginal * % marginal pupils * Bottom Quartile Comp		3.607		0.328		-0.134		-3.065
		(3.496)		(0.604)		(0.515)		(0.878)***
Group 3 * % marginal pupils * Bottom Quartile Comp						0.707		-0.630
						(0.507)		(0.888)
Constant	-23.795	-23.854	-7.761	-7.761	-26.859	-26.842	-26.636	-26.813
	(1.285)***	(1.285)***	(0.841)***	(0.840)***	(0.374)***	(0.375)***	(0.374)***	(0.375)***
Observations	27350	27350	27350	27350	441677	439353	441677	439353
R-squared	0.63	0.63			0.71	0.71	0.71	0.71
Psuedo-Likelihood			-7727.9481	-7726.8927				

Notes:
1. Standard errors in parenthesis, clustered at pupil level. * Significant at 10%; ** significant at 5%; *** significant at 1%
2. FE regressions include ability dummies and pupil level controls as in Table 1
3. Probits include ability dummies, pupil and school level controls as in Table 1 (except the selective school dummy), LEA dummies and clustered standard errors at school level

Table 9: Instrumental Variable

Explanatory variables	1st stage (OLS)	2nd stage IV (FE)	No IV (FE)
School % (October- December)	0.290		
	(8.26)**		
School % (January – April)	0.301		
	(9.08)**		
School % (May - August)	0.317		
	(9.95)**		
Voluntary Aided	0.015		
	(5.36)**		
Voluntary Controlled	0.009		
	(1.50)		
Foundation	0.020		
	(5.69)**		
City Technology College	0.041		
	(2.94)**		
Group 1 * % marginal pupils		-39.110	-9.349
		(5.466)***	(1.545)***
Group 2 * % marginal pupils		-12.565	-5.330
		(4.813)***	(1.352)***
CD border * % marginal pupils		4.174	-1.980
		(4.897)	(1.376)
Group 3 * % marginal pupils		-6.071	-1.330
		(4.834)	(1.336)
Constant	-0.042	-26.759	-26.762
	(1.73)	(0.374)***	(0.374)***
Observations	2598	441589	441677
R-squared	0.23	0.71	0.71
Number of schools		2598	2599

Notes:
1. Standard errors in parenthesis. * Significant at 10%; ** significant at 5%; *** significant at 1%
2. FE regressions include ability dummies and pupil level controls as in Table 1
3. In column 1: omitted birth month is September and school type is community. LEA dummies are included
4. One school is dropped in the second stage IV as all pupils have missing age information.