

**Education, location, education:
a spatial analysis of English secondary school public examination
results¹**

Ian Gordon* and Vassilis Monastiriotis**

Published November 2006

ISBN 07530 2009 2

*Department of Geography and Environment
** European Institute
London School of Economics
Houghton Street
London WC2A 2AE

Tel: +44 (0)20 7955 6180
Email: i.r.gordon@lse.ac.uk
v.monastiriotis@lse.ac.uk

Abstract

This paper addresses the question of how location affects educational outcomes, through analyses at two scales (the school and the sub-region) of the relationships between English secondary school exam results/value-added and local socio-demographic characteristics. As a complement to standard micro-analytic approaches to identifying significant peer group or neighbourhood 'effects' on pupils' attainment, it takes the existence of significant non-linearities in aggregate relations as its key indicator of these. At the school level, strong evidence is reported of such effects, both in relation to the ethnic/class mix of local families, and to the 'quality' of the school's own intake, with its disciplinary order (proxied by low absence rates) identified as a key intervening variable, conferring strong advantages on pupils in the most advantaged areas and schools. At a broader level, both social class variations and access to selective schools appear as important influences. Controlling for these, results appear significantly worse in those economically successful regions where teacher recruitment / retention is hardest.

1. Introduction

'Education, education, education' was Tony Blair's 1997 formulation of the priorities of a prospective New Labour administration in the UK, reflecting a clear view about the strategic significance of schooling to achievement of its 'modernising' goals of competitiveness and of social inclusion (where the focus was to be on shifts from 'welfare to work'). The strategy involved 'modernising' initiatives within education itself, where outcome targets and competition between schools were to play key roles. Rhetorically, the phrase clearly echoed the estate agents' traditional 'location, location, location' explanation of the key to property values, standing perhaps for access, reputation and neighbours, each of which plays important roles in the new quasi-market for local schooling. In this quasi-market – where competition is mediated by factors other than a direct service price – location is perceived as making a substantial difference to educational attainment, while educational factors are recognised as of increasing importance to location decisions and social/economic geographies. A more specific link between schooling and estate agency has come to the fore since 1997, with the very widespread dissemination of spatially indexed data on local school results, together with more discursive inspection reports, alongside other neighbourhood indicators, including house prices – with an official web-site now making available kinds of 'neighbourhood data' relevant to home-movers pioneered by private agencies. And, while links between middle class housing choices and perceived school quality are by no means new, their implications in terms of the implicit price of access to 'better schools' have come to be much more clearly documented in this period, through a series of increasingly sophisticated hedonic analyses of house prices (Gibbons and Machin, 2003; Cheshire and Sheppard, 2004; Brasington and Haurin, 2005).

The basis of these analyses is the fact that residential location has a major impact on the set of schools to which children have effective access – both through bureaucratic rules determining formal catchment areas for particular schools (within the state sector), and in terms of feasible/acceptable travel to school (within either the state or independent day-school sectors). Where school quality varies appreciably, this becomes a source of differential locational value, at least for those who have (or anticipate) children in a relevant age group. Parents with concerns about their children's education are then faced with trade-offs between: access to good schools;

satisfying other residential and/or consumer preferences; and pursuing other routes to better education (via purchase of non-local independent schooling, or supplementary tutoring, and/or greater personal involvement).

Beyond this, the logic of the residential choice models implies the possibility of strong influences on the social composition of areas around relatively 'good' and 'bad' schools. The former should attract more of those with children, a strong interest in their educational attainment, and greater financial resources - while the latter are liable to be populated with more of the childless, the less educationally-concerned, the poor and/or those lacking control over their area of residence. A very likely consequence then is that the 'good' schools come to display even better examination results (and the 'bad' schools even worse), whether simply because pupils entering the former enjoy other educational advantages (including parental interest and affluence), or perhaps also because their presence tends to raise the performance of all in the school - with a corresponding vicious circle of outcomes in the schools originally identified as 'bad'. If this is the case there are a number of important implications, including the possibility of progressive polarisation over the long run in local population mixes, school performance and local property values. But it also raises questions as to how well-founded (or not) might be the underpinning judgements about school 'quality', where these rest simply on published exam results that may be heavily influenced by differing mixes of more/less advantaged pupils (Brasington and Haurin, 2005)

At this neighbourhood-level more is currently known about how education affects locational behaviour and outcomes, than about how locational attributes affect education. This is also true at the broader scale of cities and regions. Here there is now a strong body of evidence about the crucial importance of human capital stocks for growth and productivity. Glaeser and Saiz (2004) show (for the US) that these have particular positive effects in otherwise disadvantaged regions, perhaps because education enhances adaptation to change. For UK regions, Monastiriotis (2002) found positive spillover effects from human capital to productivity. There is evidence also of strong divergence across (US) cities during the 1990s in levels of human capital, presumably mainly as a result of selective migration of the more skilled toward areas with an already better educated population (Berry and Glaeser, 2005). But much less

seems to be known about the geography of human capital production at this scale, particularly as it relates to schooling (as distinct from skill development within the workforce).

In this paper the focus is on identifying and accounting for the geography of achievement patterns (at school and area level) in the General Certificate of Secondary Education (GCSE) exams taken by English students at the end of their last compulsory year of schooling (at age 16). It builds on preliminary analyses for schools in the London region (within Gordon and Monastiriotis, 2006) but extends the scope across the country, using more recent data to address questions about value-added, as well as about sources of regional variation in performance. In the following sections of the paper: we first outline the theoretical relationships between school inputs, outputs and value-added in a spatial context (section 2); then we present cross-sectional evidence on the relationship between GCSE results/value-added and a series of school and neighbourhood characteristics (section 3); the spatial dimension of these is examined with specific reference to inter-school competition, neighbourhood effects and broader regional influences (section 4); and finally some conclusions are drawn about the interaction between location and education in this country (section 5).

2. Relating School Inputs, Outputs, Value Added and Spatial Context

Schooling is recognised to have a great variety of effects (mostly, if not all, positive) on those who go through it, including not only the development of skills and knowledge of many kinds, but also of fitness, attitudes/tastes and social networks. However, in this paper, as in most work on geographies of education, the focus will be exclusively on formally certificated outcomes. This is primarily because of data availability, but similar informational constraints in the real world mean that these results play key roles in mediating both the effects of local education on individuals' competitive prospects (in work and further education) and parental judgements about which locations offer those 'better' schools that are worth expending resources to access.

In the tradition of education production function analyses (Hanushek, 1979) we may think of such ‘outputs’ from a neighbourhood school as being produced with a combination of six sets of input:

- (a) individual pupil endowments (potential intelligence, motivation etc.);
- (b) family inputs (time, money, aspirations, human/social capital and strategies);
- (c) interaction of pupils or parents with peers in the community;
- (d) interaction of pupils with peers in school/class;
- (e) teacher quality;
- (f) school management and other school resources.

These can be grouped in several ways: for example, (a) to (c) and possibly (d) may be taken to represent *community* inputs; (d) to (f) and possibly (c) may be seen as contributing to *school quality*, of the kind worth paying for in one way or another; while, for local pupils, (c) and (d) together represent the *neighbourhood effect* in a strict sense – i.e. the impact of local population composition on outcomes for individuals, over and above the direct influence of their own background and characteristics. A value-added version of this approach, for a particular stage of schooling, simply relates outcomes of that stage to the *stock* of learning/skills brought from the previous stage and the *flow* of inputs of each kind during the current stage. If the inputs are additively separable (i.e. if there is no substantial interaction between them), then value-added during this stage could be seen as the sum of elements associated with each of the input types, with overlapping groups of these again defining the contribution from the community, from school quality, and from any neighbourhood effects.

This kind of framework underlies a vast number of studies – especially in the US but also in the UK – which have examined the relation between schooling outcomes, those educational inputs open to manipulation by policy, and uncontrollable background factors. For the manipulable inputs, a number of reviews have suggested substantial dissensus about which are really significant – reflecting problems of both collinearity and control in real world analyses – though a few basic findings do seem well established at least in the American literature. Thus, Barrow and Rouse (2005) report clear evidence that teacher quality matters, with experience and prior attainment as possibly relevant factors, and that smaller class sizes also tend to

improve outcomes, though the impact seems to vary greatly, probably being stronger for otherwise disadvantaged groups. In England, recent multi-level analyses of secondary school performance find significant effects from both behaviour and teaching quality, as rated on OFSTED inspections (Benton et al., 2003). Little or no evidence has been found, however, for effects of pupil-teacher ratios on performance in English state schools (see e.g. Dearden et al., 2002). School type has been shown to make a big difference, but this is very much bound up with issues about selection. At the family level, British studies show strong impacts of parental interest/time inputs² (Feinstein and Symons, 1999), together with both parents' education level and family income (Blanden and Gregg, 2004). Ethnic origin also has important effects, although after even a crude control for economic circumstance³, pupils from all non-white origins apart from Afro-Caribbean and 'black other'⁴ actually exhibit *better* GCSE results, while all are apparently associated with greater progress during secondary school (Wilson et al., 2005). In terms of peer group effects, there is strong evidence from British secondary school pupils of positive externalities, from having classmates who are abler, and/or from families of high socio-economic status (Feinstein and Symons, 1999), while Gibbons (2002) finds significant *additional* benefits at a neighbourhood level from living in areas with better qualified residents. The quality of educational inputs too may be affected by pupil or area mix factors (Lupton, 2005, 2006), whether through varying stress levels impacting on teacher performance, or through greater difficulties in teacher recruitment and retention in those schools/areas with more 'challenging' pupil mixes (Smithers and Robinson, 2001).

Much of the cited work operates within a value-added framework, in that it relates secondary school achievement at a given stage *both* to variables relevant to that stage *and* measures of attainment at an earlier stage. Outside such a framework of analysis, measures of 'value-added' during part of the education process have an inevitable degree of arbitrariness to them, since educational development cannot be measured simply in terms of differences in a stock of pupil knowledge between two points of assessment. In England, where value-added indicators (VAI) have recently been added to raw measures of exam performance in published school league tables – to create a more level competitive 'playing field' – they simply represent differences between pupils' actual results at a particular stage and expectations of these based on

relative achievement levels in the previous stage. If, as we should expect, the flows of various kinds of ‘input’ are positively correlated across successive stages (so that many of the same children are (dis)advantaged in each), this method will clearly produce biased measures of value-added, obscuring continuing disparities in educational progress across successive stages⁵. Other more specific kinds of bias can arise at points where the basis of assessment changes, as at age 16 within the English system, where it shifts from the earlier standard tests (KS1 to KS3) on a common curriculum of English, maths and science to public examinations (for the so-called GCSE), where students may take and pass quite disparate mixes of exams⁶. In these, the core subjects can be replaced, as well as augmented, by a wide choice of others, including many vocational subjects, now treated as fully equivalent to traditional academic disciplines. In this respect it seems that measures of ‘output’ from schooling now reflect political values (of inclusivity) more than those of the labour market. In any case the shift in bases of valuation confuses the significance of value-added measures. Empirically, this is reflected in a marked disjunction between the patterns of VAI found for the first two years of secondary schooling (between KS2 and KS3), where measured value-added was significantly higher in academically selective schools, and those for the next three years (between KS3 and GCSE), where these schools did significantly worse than average, while those with arts and business specialisms⁷ achieved higher VAI scores. Schools with less advantaged pupil mixes can also be seen to have done absolutely or relatively better in these terms after KS3 (Taylor and Nguyen, 2006)⁸. Similar discrepancies are evident in local authority league tables, where the most deprived areas achieve higher VAI scores *after* KS3⁹. At an individual pupil level, there is no comparable evidence of pupils from poor families starting to do better when the basis of assessment shifts, but there does appear to be an upsurge in performance among the least advantaged ethnic minorities. This pattern of findings supports the notion that value-added estimates linking GCSE/NVQ results to KS tests are subject to significant biases because they do not actually compare like with like, reflecting a shift to more inclusive bases of valuation after KS3 as much as real additions to value. Though far from perfect, these new data do, however, cast some important additional light on the pattern of variability in secondary school outcomes and the extent to which this reflects differences in intake mix, as between both types of school and geographic areas.

How far residential social geographies affect these intake mixes should depend on three sets of factors: the scale and rigidity of school catchment areas; what autonomy schools have in selecting pupils, and how they use it; and parents' varying capacity to pursue/access preferred schools outside their immediate area. In England the catchment area system has always operated differently for selective, voluntary aided and grant maintained schools than for community comprehensives, though the formal scope for parental choice in these was greatly extended by the 1988 Education Reform Act. An indicator of the varying extent to which this has been taken up is the proportion of parents appealing against initial failures to access a first-choice school, which ranges between 20% in the top quartile of authorities and 5% in the bottom quartile¹⁰. Another indicator is that across England as a whole 8% of pupils now attend a secondary school outside their home local authority (over 20% for London), with net shifts evident out of areas with larger proportions of black or lower class families. Generally, we would expect that middle class groups were better able (for both financial and informational reasons) to exploit the full range of school choices potentially available (Stillman, 1990; Taylor, 2002), though those locating in traditional middle class suburbs for housing/environmental reasons, as well as those who have deliberately moved to 'good' school catchment areas will have less reason to do so. For those among the middle class with more 'urban' tastes, as well as strong educational interests, accessing non-local schools provides an alternative means of exploiting their advantage in the quasi-market, increasing the potential tendency toward polarisation of school intakes.

3. Regression Evidence on School Performance

Our empirical analyses focus on school level performance (rather than measures for individuals or areas), relating GCSE success rates, and the associated VAI, to attributes of schools themselves (on the one hand) and of surrounding areas (on the other). The basic data set is taken from the DFES's published 2004 school performance tables for English secondary schools¹¹, which include information on the proportions of pupils in their last year of compulsory education who meet minimal, basic or median standards of achievement¹², together with VAI scores for the period since the KS2 tests at age 11, the normal point of transfer to secondary education. Whether for historic or political (anti-elitist) reasons, the official tables omit any parallel measure of numbers achieving high standard passes, though this gap can

partially be filled from press sources. The additional background information which the official tables provide covers the formal status of the school (whether independent or more/less state controlled, its gender admissions policy, whether academically selective or not, and whether catering for special needs pupils), together with data on pupil numbers, age ranges and absence rates. Presumably for policy reasons the public tables include none of the information on pupil characteristics from the Schools Census (now included in the PLASC research database). However, that census (despite useful information on ethnic and language background) includes only a single binary indicator of family economic status. In any case, our primary interest is in the relationship between school outcomes and characteristics of the local population, rather than with schools' own pupil mixes. For this we rely on Population Census data for families in wards close to the school in question.

The omission of measures of high individual achievement from the published tables is a significant gap, both because distinct influences may come to the fore at this level, and because it obscures differences among the high achieving schools: in 2004 for example, 5% of pupils went to schools which had success rates of 97-100% by the median (5 A-C) standard. The recording of overall 'points', based on a weighted sum of successes at different grades, does not fill this gap because they are too sensitive to varying school practices in entering students for additional subjects. We have therefore made use of partial data on results at A grade and above compiled by two of the broadsheet national newspapers for sets of high achieving schools. One set from the *Daily Telegraph* provides figures on high performance success rates measure in terms of the proportion of pupils achieving at least 5 high level (A/A* grade) passes – directly paralleling the published data for median and basic successes – but only for a set of (some 500) leading state schools. The *Times*' list includes schools from both state and independent sectors (though 500 of its 840 entries come from the latter, and it includes many fewer comprehensives than does the *Telegraph* list), but presents results in terms of the proportion of *passes* achieved at high grades¹³. Fortunately, from the 273 schools appearing in both lists the two measures are well enough correlated to allow reasonably reliable conversion from one to another¹⁴, while prediction of higher level success in other schools is possible from relationships with achievement rates at the lower levels and school point scores in relation to school type observed in this data set. These are used in this paper to generate two sets of

estimates for the proportion achieving 5 high level passes, one quite direct set for all of the 1077 schools covered by one or both of the *Times/Telegraph tables* (covering about a quarter of the age cohort), and a second less direct set for the full range of schools, most of which are expected to have success rates at this level well below those of the bulk of schools covered by these press tables¹⁵.

Additionally, we have computed a summary measure of overall GCSE/NVQ performance for each school intended to approximate the (unpublished) capped points score from which the DfES's VAI are computed. The difference between the capped figures and the published totals is that the former take account only of grades achieved in up to eight subjects. To achieve a similar standardisation we have used regressions of total points on schools' estimated success rates at high, medium, basic and minimal levels (none counting grades in more than 5 papers), and then rescaled these to match the VAI¹⁶. We shall refer to this summary output measure as 'adjusted points' (AP).

For our statistical analyses we have excluded schools catering solely for pupils with 'special educational needs' (SEN), together with all schools that cater wholly for boarding pupils, while schools which are part-boarding were down-weighted pro rata to the boarding proportion¹⁷. We also dropped schools with missing values for any of our variables and those with only 30 or less pupils in the GCSE cohort, leaving us with a set of some 3300 schools, between them covering 93% of GCSE entrants (98% of those from state schools and 39% of those from the independent sector). After exploratory analyses, schools were grouped into four broad types - selective schools (and City Technology Colleges); non-selective independent schools; directly funded or voluntary aided comprehensives (from the state sector); and community/local authority controlled comprehensives and secondary moderns – with a cross-cutting distinction between boys, girls and mixed schools. Three other school variables were derived from the DfES tables: pupil numbers in the GCSE cohort; day pupil absence rates (whether authorised or not); and the proportion of SEN pupils in the GCSE cohort (with a conservative 50% down-weighting of those without formal need statements).

Compositional measures for the local population were all derived from the 2001 Census, relating to proportions: from lone parent families; with parents born abroad and/or with potential language needs¹⁸; from households with nobody employed; from Black or Asian ethnic origins; from professional/managerial (SEC1-2) or working class (SEC 6-8,U) families; and of parents with degree level, or no, qualifications¹⁹. For each of these we computed spatially averaged measures using three alternative distance weights to reflect more or less localised effective catchment areas. In contrast to earlier work focused on the London region (Gordon and Monastiriotis, 2006) we found that the most localised version²⁰ function provided the best fits for just about every variable, with no clear class variation, and this was adopted for all. Exploratory analyses showed a quite strong pattern of inter-correlation among the socio-demographic variables, with two principal components associated respectively with social class/qualifications and ethnicity/immigration. At least initially, however, we included the full set of these variables in the regression analyses.

All regressions were estimated by Weighted Least Squares except for those relating to the highest level of achievement, where a maximum likelihood procedure²¹ was used to counter obvious selection biases in the ‘sample’ derived from the *Times* and *Telegraph* lists of top schools. This procedure involved first modelling the chances of schools being selected for these lists – in relation to their success rates at basic and median performance levels together with the size and type of school, its location and the local population mix – and then including predicted probabilities of selection (the ‘inverse Mills ratio’) as a control in the main regression²². This Heckman procedure was also used – with a regression equation including just success rates at the basic and median levels, interacted with school type (independent/state by selective/non-selective) – to generate predicted ‘top level’ success rates for schools outside these lists.

The base set of regression results are presented in Table 1 for four levels of attainment: the minimal (1 or more passes at grade G or above); basic (5 or more passes at this level); median (5 or more passes at grade C or above); and top (5 or more passes at grade A or A*). These are nested, with the same set of independent variables used for each, so that looking along the row one can see the incremental effect of each in lifting attainment from one level to the next. Strong associations are

evident between performance and both school and area characteristics. For schools, the strongest effects, unsurprisingly, tend to be academic selection on the positive side and the proportion of SEN pupils, on the negative side. But results also appear to be significantly better in girls-only schools, those with larger pupil numbers and independent or voluntary/directly funded (non-selective) schools. Apart from the school-size effect, results are generally consistent as between the more/less restricted regressions for high level results, and suggest a tendency for gender-mix and selection/control both to have stronger effects above the basic qualification level.

Table 1. Regressions of school-level GCSE results on school and area characteristics

	Qualification Level				
	Minimal	Basic	Median	High (published)	High (published + predicted)
Constant	103.9 (69.7)**	104.5 (35.7)**	55.31 (8.8)**	36.43 (2.0)*	8.53 (2.3)*
School characteristics					
Girls school	0.89 (3.2)**	2.18 (4.0)**	5.88 (5.0)**	4.73 (2.8)**	5.48 (7.8)**
Mixed school	-0.48 (2.0)*	-0.76 (1.6)	-1.18 (1.2)	-0.78 (0.5)	-0.70 (1.2)
Selective/CTC	1.10 (4.5)**	4.55 (9.5)**	37.82 (36.8)**	20.75 (11.2)**	35.44 (58.3)**
Voluntary aided /directly funded	0.80 (6.9)**	2.41 (10.7)**	7.01 (14.4)**	-0.35 (0.3)	2.45 (8.5)**
Independent non- selective	-2.80 (4.3)**	-0.22 (0.2)	22.82 (8.3)**	5.60 (1.3)	13.10 (8.1)**
SEN pupil %	-0.15 (14.3)**	-0.43 (21.5)**	-1.06 (24.7)**	-0.36 (2.9)**	-0.37 (14.4)**
Pupils in year (logged)	0.54 (3.5)**	1.53 (5.0)**	4.95 (7.6)**	-1.25 (0.9)	1.27 (3.3)**
Area characteristics					
Lone parents %	-0.17 (7.9)**	-0.42 (9.8)**	-0.58 (6.3)**	0.04 (0.2)	-0.16 (3.0)**
Born abroad %	0.07 (1.4)	0.15 (1.6)	0.54 (2.6)**	-0.78 (1.8)	0.12 (1.0)
Language need %	-0.08 (2.3)*	-0.24 (3.4)**	-0.75 (5.1)**	0.62 (2.0)*	-0.17 (1.9)
Workless household %	0.01 (0.2)	0.09 (1.7)	0.30 (2.7)**	0.07 (0.3)	0.09 (1.3)
Asian %	0.09 (3.1)**	0.26 (4.6)**	0.68 (5.6)**	-0.65 (2.5)*	0.13 (1.7)
Black %	0.13 (4.9)**	0.32 (5.9)**	0.51 (4.4)**	-0.19 (0.8)	0.05 (0.7)
No qualifications %	-0.16 (6.0)**	-0.38 (7.4)**	-0.35 (3.1)**	0.21 (0.8)	0.05 (0.8)
High qualifications %	-0.01 (0.7)	0.01 (0.2)	0.43 (6.3)**	0.46 (3.9)**	0.33 (8.3)**
Professional and managerial %	-0.05 (2.1)*	-0.07 (1.6)	-0.11 (1.2)	-0.21 (1.3)	0.004 (0.1)
Working class %	0.03 (1.0)	0.06 (1.1)	-0.23 (2.1)*	-0.16 (0.8)	-0.10 (1.5)
Observations (schools)	3304	3304	3304	759	3304
R-squared	0.25	0.40	0.64	0.63	0.72

Notes: 1. Bracketed values are t statistics; 2. Asterisks denote significance levels: * = 5%; ** =1%; 3. Omitted categories are: boys and community/controlled non-selective schools; 4. Observations are weighted by the square root of the number of day pupils in the year.

In terms of the characteristics of local populations, it is qualification levels in the potential parental population which emerge as important, rather than occupationally-based indicators of class – with absence of qualifications mattering up to the median level, and degree-holding from there upwards. The proportion of lone parents also exerts a strong negative effect (though less at the top end), while the only significant relation with worklessness is perversely positive. Among the ethnic/migration variables, the proportions from Asian and black origins appear to exert strong positive effects up to the median level, while the proportion born abroad also shows a significant positive effect at that level, whereas there is a significant negative relation with the language need indicator at least up to the median level. At the highest level, the results for these groups are less significant, and differ somewhat between the more/less restricted regressions.

These summary regressions leave a large part of the variance across schools in GCSE success rates unaccounted for, particularly at the basic and minimal levels – in other words this combination of school and population characteristics is not particularly successful in explaining the cases where a substantial minority of pupils fail to reach the ‘basic’ level of GCSE attainment. With this problem particularly in mind, we turn to an extended model including information on school absence rates, which we take as a potentially important intervening variable between population/school characteristics and performance, since it can represent not only a constraint on achievement by the absentees but also an aspect of schools’ general disciplinary order. In Table 2 we report results for the summary ‘adjusted points’ performance measure both with and without inclusion of the absence rate, together with a regression of this absence rate on the other attributes. The results bear out our hypothesis, with the absence rate emerging (in column 2) as the strongest influence on overall performance, while its inclusion substantially reduces the effects attributed to many of the other variables which appear to operate in part at least through the absence rate (column 3). In particular, it appears that the selective/independent/voluntary school advantage is largely attributable to their lower absence rates (and/or stronger school order), while the effects associated with variations in the local population mix also appear to run in large part through this behavioural factor. This is most strikingly the case in relation to the proportion of lone parent families, which emerges as the strongest influence on

absence rates. But all of the effect of parental lack of qualifications, most of the negative effect of language need and the positive effects of non-white/migrant groups also seem to be mediated in this way. The exception is the positive effect of a graduate presence among the local population, which emerges as the strongest influence on results after controlling for the absence rate – maybe because its main effects are at the higher levels of GCSE performance among pupils less susceptible to absenteeism. After controlling for the absence rate, the working class share appears as a significant negative influence (previously off-set by its association with lower absence rates). More detailed results for each of the separate qualification levels (not shown here) indicate that the absence rate has a significant negative impact at each level, and is the most significant of all the independent variables except on success rates at the highest achievement level. Its inclusion does particularly improve the regression fit for the minimal level of achievement, but this still remains much the least satisfactory. Overall, this set of results suggests the existence of two rather separate patterns of association between the local population mix and school exam performance: one contributing to better/worse disciplinary orders in the school; and the other perhaps more directly associated with variations in individual pupils' educational (dis)advantage.

Table 2. Regressions of school-level GCSE and absence rates

	Adjusted points	Adjusted points	Absence rate
Constant	302.9 (18.4)**	351.9 (26.6)**	4.50 (5.0)**
School characteristics			
Girls school	17.59 (5.7)**	16.91 (6.8)**	-0.06 (0.4)
Mixed school	-4.03 (1.5)	-0.32 (0.2)	0.34 (2.4)*
Selective/CTC	97.27 (36.2)**	61.95 (26.9)**	-3.25 (22.0)**
Voluntary aided /directly funded	18.62 (14.7)**	10.27 (9.9)**	-0.77 (11.0)**
Independent non- selective	48.09 (6.7)**	12.50 (2.2)*	-3.28 (8.3)**
SEN pupil %	-2.91 (26.0)**	-1.78 (19.0)**	0.11 (16.0)**
Pupils in year (logged)	12.71 (7.4)**	10.52 (7.7)**	-0.20 (2.1)*
Absence rate %		-10.87 (42.9)**	
Area characteristics			
Lone parents %	-1.90 (7.9)**	-0.25 (1.3)	0.15 (11.4)**
Born abroad %	1.37 (2.5)*	0.56 (1.3)	-0.08 (2.5)*
Language need %	-1.91 (5.0)**	-0.70 (2.3)*	0.11 (5.3)**
Workless household %	0.75 (2.6)**	0.54 (2.3)*	-0.02 (1.2)
Asian %	1.81 (5.6)**	0.53 (2.0)*	-0.12 (6.7)**
Black %	1.53 (5.1)**	-0.05 (0.2)	-0.15 (8.8)**
No qualifications %	-1.28 (4.4)**	-0.04 (0.2)	0.11 (7.1)**
High qualifications %	0.99 (5.6)**	0.79 (5.5)**	-0.02 (1.9)
Professional and managerial %	-0.34 (1.4)	-0.33 (1.7)	0.00 (0.0)
Working class %	-0.40 (1.4)	-0.96 (4.2)**	-0.05 (3.3)**
Observations	3304	3304	3304
R-squared	0.65	0.77	0.46

Notes: see Table 1.

Table 3. Regressions of school-level results and value added

	GCSE Result (Adjusted Points)	Entry Quality	Value-Added since KS2
Constant	376.2 (47.1)**	-672.2 (117.9)**	1048.5 (189.3)**
School characteristics			
Girls school	16.91 (6.8)**	-6.15 (3.5)**	23.05 (13.4)**
Mixed school	-0.09 (0.0)	-7.89 (5.3)**	7.80 (5.4)**
Selective/CTC	61.73 (27.0)**	47.10 (28.9)**	14.63 (9.2)**
Voluntary aided /directly funded	9.98 (9.7)**	7.62 (10.4)**	2.35 (3.3)**
Independent non- selective	12.34 (2.1)*	1.11 (0.3)	11.24 (2.8)**
SEN pupil %	-1.82 (19.5)**	-1.40 (21.0)**	-0.42 (6.5)**
Pupils in year (logged)	9.97 (7.4)**	9.56 (9.9)**	0.41 (0.4)
Absence rate %	-10.84 (48.4)**	-4.64 (26.0)**	-6.20 (35.8)**
Area characteristics			
Lone parents %	-0.16 (1.5)	0.26 (3.5)**	-0.42 (5.8)**
Low class factor	-1.09 (15.6)**	-0.80 (16.0)**	-0.29 (6.0)**
High ethnic factor	-0.09 (1.5)	-0.57 (13.0)**	0.48 (11.3)**
Observations	3304	3304	3304
R-squared	0.77	0.69	0.60

Notes: see Table 1.

For a second peek into the black box of school performance, we explored relationships with the DfES' measure of 'value-added' between KS2 and GCSE. Despite the potential sources of bias in this indicator, noted in the previous section, we take it at face value here as indicative of progress made by pupils during the previous five years of (normally) secondary schooling. We also use it together with the adjusted points score to recover the implicit measure of average entry quality for each school (estimated as the difference between adjusted points scored and the VAI)²³. Three parallel regression analyses were undertaken (with results in Table 3) using as independent variables the same set of school characteristics, but now with just three

indicators of local population characteristics: the lone parent rate, and two summary measures of class/parental education and of ethnic minority/migrant presence²⁴. The first of these regressions simply relates the adjusted point scores for each school to this set of characteristics. The second and third regressions, taking the entry quality and value-added measures (respectively) as their dependent variables, then partition these relations between effects identifiable with patterns of recruitment into the schools (column 2) and with progress in the school (column 3).

Overall, the results tend to show that variables associated with a stronger intake into particular schools are also linked with greater value-added during time in the school. This is notably the case for the main school types (in terms of selectiveness and control) where, though intake effects are dominant (except for independent non-selective schools), they tend to be reinforced through value-added. In the case of gender mix, there is a turnaround in the relative position of the two single sex types, with girls schools reversing an initial disadvantage²⁵. School size is exceptional in a different way, in that the strong positive association between larger numbers of pupils in the GCSE year and better results seems to be entirely a matter of larger schools attracting a better quality intake, rather than of them contributing to higher value-added. The SEN and absence rate variables do conform to the dominant pattern, by showing similarly negative relations both for entry quality and value-added, but the meaning of the entry quality relationship is not entirely obvious, since absences occur *after* entry, as may SEN recognition. What this result actually shows is that schools with high current absence (or SEN) rates tended to recruit pupils of lower quality (in relation to KS2 performance). This might possibly reflect a position of competitive weakness for schools with poor disciplinary reputations, or alternatively perhaps a continuity of disciplinary problems affecting performance of some groups of pupils through both primary and secondary schooling. One partial test is to see how far entry quality may affect (secondary school) absence rates, by adding this variable into the absence rate regression (from Table 2). When this is done (column 4), it emerges as much the strongest of the independent variables, taking over the bulk of the effect associated with the selective and voluntary aided state school types. This suggests that the effect does not run simply from discipline to recruitment, but rather that intake characteristics have a major effect on school absence rates – though, with still only

53% of variance accounted for, there still seem to be important school-specific factors in these.

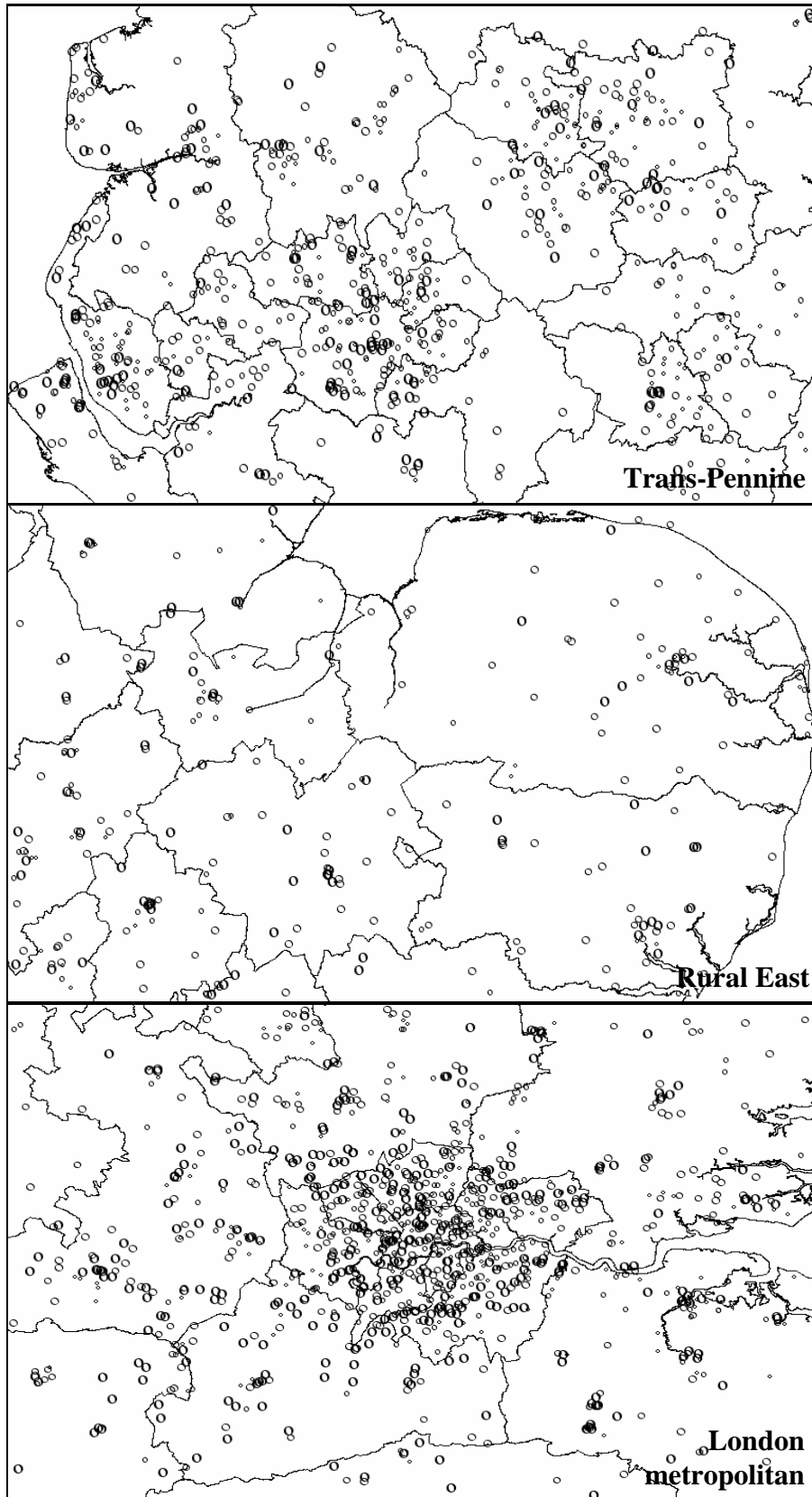
In terms of the area-related influences on entry quality and progress in secondary school, the low class factor conforms to the general pattern, with similar strong negative effects on both. For both lone parenthood and ethnicity there is a clear reversal of sign, with an initially positive association in the former case being succeeded by larger negative effects on progress after KS2, while a strong negative effect of minority/migrant presence at KS2 was apparently overcome through stronger progress thereafter²⁶. A more detailed analysis with the full set of population indicators suggests that larger proportions born abroad, or of Asian and black ethnic origins were all associated with higher value-added (in the last case reversing an apparent disadvantage at the intake stage). This is not inconsistent with the pattern shown at the individual level (where data are disaggregated by ethnic origin and first language) so we cannot draw any simple inferences about possible neighbourhood effects or other geographic influences on performance. These are followed up rather more closely in the next section, starting with an examination of the spatial pattern of results within two urban and one rural region.

4. Spatial Patterns and Spatial Processes

4.1 Spatial Correlation Analysis of School Level Results

The extent to which good and bad schools are clustered together is of some importance, both in terms of whether (or not) families in particular areas have an effective choice and for the dual question of how much competitive pressure might be exerted on the weaker schools to either improve their performance or shrink. This begs a question which we cannot yet answer about which are actually good and bad schools in terms of what the institutions themselves contribute to pupils' varying exam achievements. But one point of entry into this issue is to look at whether schools with above (or below) average GCSE results tend to be located close to similar schools - as might be the case either if residential segregation operated at a scale much broader than individual school catchment areas, *or* if schools learned from their neighbours. Alternatively, we might find a negative pattern of spatial correlation if some successful schools, with a degree of control over their intake, managed to 'cream off' the more promising pupils, to the disadvantage of their neighbours.

Figure 1. School GCSE Results for three English urban/rural regions



Legend: Quintiles of school AC5 scores: ◦ Bottom; ◦ Second; ◦ Median; ◦ Fourth; ◦ Top

How far any of these processes can be expected to operate depends on settlement scale and density, since in small towns and rural areas there will be fewer reachable neighbours for any given school, and because residential segregation tends to operate across much broader areas in more populous city-regions (Gordon and Monastiriotis, 2006). Available transport options and possible administrative obstacles to 'out of area' school access might also be of significance. Bearing these in mind, we decided to explore the spatial pattern of association of secondary school results within three broad areas: a version of the London metropolitan region; a trans-Pennine region (of the same scale) encompassing the four Yorkshire and Lancashire conurbations; and a more rural region covering East Anglia and a small part of the East Midlands.

Simple mapping of school results (in Figure 1) suggests little discernible pattern in the Rural East, while in the two urban regions some patterns of positive association seem evident. These are hard to characterise, however, at least without more geographical cues in the map; even for London, where a cluster of low performing schools seems evident along Thameside, but where the traditional inner-outer pattern is not obvious. It is in any case notoriously difficult to spot patterns of negative spatial association, if these were present. For a less subjective test we computed a series of measures of spatial autocorrelation (Moran's I statistic) relating results for individual schools, on the adjusted points measure, to more or less tightly defined sets of neighbours. These neighbours were identified both in terms of those within a certain distance range and in terms of the nearest so many: broad results were very similar but somewhat stronger in the latter case, so we focus here on the associations between the results for a school and those of its nearest 4 or 9 neighbours. The basic results, presented in Table 4, for the overall AP measure of GCSE performance, display clear evidence of positive correlation (especially for the 4 neighbour comparison), across the country as a whole and in the two urban regions, though not within the rural east. Because this might simply reflect the extension of areas of social (dis)advantage across multiple school catchments, obscuring more subtle competitive or complementary relationships, the analysis was repeated after adjusting school outcomes for three different sets of controls. These involved: average performance across the relevant Training and Enterprise Council areas (of typically around half a million population); our school and neighbourhood measures; and these plus absence rates.

Table 4. Moran I statistics for spatial association of school outcomes

Variable	Neighbour schools	Rural East	Trans-Pennine	London Metro	England
Adjusted Points (AP)	4	0.025	0.163 [#]	0.126 [#]	0.163 [#]
	9	0.005	0.105 [#]	0.086 [#]	0.114 [#]
AP relative to TEC mean	4	-0.007	0.126 [#]	0.114 [#]	0.101 [#]
	9	-0.021	0.064 [#]	0.070 [#]	0.055 [#]
AP controlled for school/area chars.	4	0.022	0.022	0.092 [#]	0.075 [#]
	9	0.007	0.003	0.079 [#]	0.067 [#]
AP controlled for school/area/absences	4	0.118 [#]	0.034 ⁺	0.057 [#]	0.058 [#]
	9	0.070 [*]	0.031 [*]	0.043 [#]	0.050 [#]

Notes: Significance levels for 10%, 5% and 1% are denoted by +, * and #, respectively. Moran's I statistics have been calculated in GeoDa using the k-nearest neighbours criterion (alternatively for k=4 and k=9).

Including these controls served progressively to reduce the measured strength of positive spatial association (except in the Rural East where the last control actually disclosed such a pattern), but without ever suggesting any significant negative spatial dependence. This suggests a couple of implications. First, the potentially available range of choice in terms of simple quality indicators *is* generally dependent on residential location, even for those able to travel beyond a local school's immediate neighbours. And, secondly, that there is no obvious general evidence of spatial competition working to the disadvantage of a successful school's near neighbours. In order to explore this competitive issue in a more focused way, we turn now to consider the relations between different types of school in an area, and evidence on how the travel to school behaviour of different social groups responds to the local pattern of opportunities.

4.2 The Effects of Spatial Competition Between Schools

Public discussion of the implications of increased choice and competition in a quasi-market within the state sector of education has tended to polarise between two different visions of how these processes can be expected to operate. One of these focuses on the exercise of choice by well-informed parents shifting demand toward schools with generally preferred characteristics, while creating positive incentives to the others to either raise their game or else be squeezed out of the system. The other assumes that the power of choice rests rather with schools, looking for better/less challenging pupils, and/or a minority of parents possessing both more effective choice

and children with more of the desired characteristics, leading to an increasing concentration of advantage in one set of schools at the expense of the rest and of those who have to rely on them. If the latter model applies, the net overall effect of educational outputs may be either better or worse, depending on how peer group effects actually operate, but in any case outcomes are likely to be less equal and to reproduce existing patterns of educational/social advantage.

One version of the unequal choice model rests simply on the processes of school-oriented residential choice that we discussed in relation to the hedonic models. This is the only operative form of competition in contexts where all schools are community comprehensives with rigidly enforced catchment areas, though it operates simply between parents rather than between schools, who cannot effectively alter their intakes. Typically we might expect it to reinforce the effects of other sources of residential differentiation. With independent, selective or voluntary controlled schools exercising more freedom in relation to the origin and attributes of their intake, there is scope for direct inter-school competition across wider areas. And, since the 1988 Education Reform Act, there is also potential for such competition between 'bog standard' comprehensives²⁷, with the more successful drawing pupils from further afield, selected for their suitability and/or self-selected by the efforts of their parents. Given the difficulty of distinguishing school quality from pupil quality, it remains very much easier, however, to identify the effects of competition across school types, except in cases where there are good a priori expectations against preference for the neighbourhood school.

To examine the first of these, we start by comparing the effects of local population characteristics on intake quality in schools of different types, and then looking at how proximity to one or other of the school types with some effective power to select their pupils affects outcomes in other schools. As the results reported in the first 3 columns of Table 5 show, for the selective schools/CTCs and the other independent ones there are actually no significant relations between intake quality and local population characteristics. For the voluntary aided/directly funded schools there is a significant negative relation with the low class factor (only), whereas for the community/controlled schools the low class factor has a significantly stronger effect,

and both lone parent families and minorities/migrants in the local population also serve to significantly depress intake quality.

Table 5. School type and school competition

	Selective and Independent Day Schools	Voluntary Aided and Directly Funded Schools	Community Comprehensives and Voluntary Controlled	All Schools	
	Entry Quality	Entry Quality	Entry Quality	Entry Quality	Value-added
Constant	-789.8 (65.6)**	-656.2 (52.8)**	-689.0 (90.8)**	-673.8 (108)**	1031.1 (155)**
School characteristics					
Girls	-3.19 (1.2)	-5.72 (1.5)	-3.06 (0.9)	-5.32 (2.8)**	23.65 (11.7)**
Mixed	-16.51 (6.2)**	-10.06 (3.5)**	-2.16 (0.8)	-10.44 (6.5)**	5.44 (3.2)**
Selective/CTC	36.52 (8.9)**			69.58 (39.6)**	37.88 (20.3)**
Independent non-selective				16.12 (3.7)**	30.37 (6.5)**
VFDCOM				15.69 (17.0)**	8.79 (9.0)**
Number of 15-year old pupils in school (log)	24.37 (9.9)**	3.93 (1.8)	11.00 (7.9)**	10.12 (9.8)**	1.24 (1.1)
SEN pupil %	-1.47 (5.4)**	-2.65 (18.5)**	-1.49 (18.5)**	-1.69 (24.0)**	-1.00 (13.3)
Area characteristics					
Low class factor	0.18 (1.2)	-0.90 (8.2)**	-1.23 (19.3)**	-1.26 (22.3)**	-0.70 (11.6)**
High ethnic factor	-0.07 (0.5)	-0.16 (1.8)	-0.47 (8.3)**	-0.26 (5.6)**	0.84 (17.0)**
Lone parent %	0.11 (0.5)	0.21 (1.3)	-0.43 (4.5)**	-0.12 (1.6)	-0.94 (11.4)**
SELEC/CTC pupils in surrounding area (%)				-49.37 (10.0)**	-24.71 (4.7)**
VFDCOM pupils in catchment area (%)				-20.07 (7.8)**	-8.63 (3.2)**
Observations	410	980	1915	3304	3304
R-squared	0.551	0.368	0.640	0.648	0.442

Notes: 1. See Table 1; 2. In column 2 the omitted school type is voluntary/directly funded.

The last pair of columns in this table reports comparable regressions for all (day) schools, adding measures of accessibility to two of the more advantaged types of school to test for ‘creaming effects’. The measure used here represents the proportion of all pupils within a quite broadly defined hinterland who attend the type of school in

question. As with the local population characteristics this ‘hinterland’ is actually defined in distance decay terms, as a weighted average across all surrounding schools, with distance related weights – though in this case the rate of distance decay used is only a third of that used to measure local population characteristics²⁸. In the reported regressions these competition measures are presented only for selective/CTC schools and for voluntary aided/directly funded schools, since there was no evidence of a competitive effect from the independent non-selective group. The results in columns 4 and 5 show clearly, however, that in areas where more children go to either of the advantaged school types there is a clearly negative (creaming) effect both on intakes and value added for other schools²⁹.

Since the scaling of these competition variables is directly comparable with the dummies for school types, adding the pair of coefficients (dummy and area average) should provide an estimate of the net effect on overall performance in an area of having (some or more) pupils going to one of the advantaged school types. From column 4, we can see that the net effect on school intake quality is clearly positive for the selective/CTC category, but not statistically significant for the voluntary/funded group³⁰. Clearly these schools cannot be directly responsible for the higher quality of the local intake as indicated by KS2 tests taken at the end of primary education. Unless this reflects their attraction of able pupils from an even wider area than we have considered, the explanation would have to be in terms of a motivational effect of selection at 11 on primary school achievements. From column 5, there is a suggestion (at the 5% significance level) of net positive effects in relation to value-added too for selective schools/CTCs, though not for voluntary/funded schools³¹. In sum then, the evidence from these spatial analyses is that the presence of selective schools and CTCs has a clear effect on intake characteristics – creaming from other schools, but also raising the average intake quality in the area – and also possibly the level of value-added. Voluntary/funded schools also benefit from creaming but do not add to overall outcomes, while (the much smaller category of) non-selective independent schools, which do not appear involved in creaming, also seem to make a positive contribution to overall local performance.

Within the community school sector there should not be the same potential for creaming through the exercise of choice by schools as to who to admit. A similar kind

of effect can, however, be produced not only through residential moves - by those who care enough and can afford it – into the catchment areas of schools which are preferred because of their results, reputation or the ethnic/social mix of their local recruits, but also through a similar group of parents patronising non-local schools. The most likely context for such extended school travel to be generated on a socially selective basis would be that of middle class enclaves within generally disadvantaged (or simply working class) catchment areas, from which potentially ‘better’ schools can be accessed through extended travel. How important this may (or may not) be in practice will depend both on the ease with which this option can be pursued in particular places, and on the extent to which it is families with already advantaged children who exploit it.

In order to investigate whether secondary school usage by middle class children is significantly less localised, either in general, or in the cases where a local school could be seen as inferior (because its catchment area is not preponderantly middle class), we examined micro-data from the 1999-2001 waves of the National Travel Survey on distances of school trips³². Overall it appeared that children of professional and managerial parents (in the 11-15 age range) did have longer median travel distances than others (2.5 versus 2.0 miles). But this pattern (which holds strongly in Inner London and in more rural areas) did not apply everywhere, and seemed to be very largely due to the fact that this social group tends to live in areas of lower density; controlling for urban/rural area type and residential density, the significance of class differences vanishes. In a simple sense then, it is not clear that middle class children (apart from those attending boarding schools) are less localised in their pattern of secondary school attendance. The real question, however, is whether they respond differently in situations where *local* schools are perceived as inferior to those available a bit further away. To investigate this with the NTS data (which does not identify specific localities), we took recorded population densities as a proxy for the proportion of lower class (or otherwise disadvantaged), and thereby as a predictor of likely perceptions about local school quality. Distances of travel to school were then regressed on a set of variables including both the density of the immediate local area (at ward level) and the local authority as a whole³³, together with parental social class and some other indicators of local accessibility.

The results (reported in Table 6) show incidentally that distances travelled tend to be substantially greater not only in rural areas and in those (low density) suburbs characterised by lack of nearby centres and infrequent public transport – but also in London, where the issue is clearly not distance to the nearest school (as in these areas) but rather readier access to a wide range of schools outside the immediate locality (or LEA).

Table 6. Regressions of travel distance on family and locality characteristics

Variable	Mean	S.D.	Regression coefficient
Dependent			
Logged distance (miles)	0.828	0.926	..
Independent			
Constant	1.000	0.000	-0.22 (0.9)
<i>Class & Density</i>			
Professional/ Manager (ProfManH)	0.39	..	-0.04 (0.5)
Density difference (pop. per hectare)	8.49	15.75	-0.01 (3.2)**
Density difference * ProfManH	3.48	10.41	0.01 (3.0)**
Rural area	0.09	..	0.23 (1.6)
Rural * ProfManH	0.04	..	0.62 (3.0)**
<i>Controls</i>			
Age (years)	13.37	1.63	0.08 (4.6)**
London	0.12	..	0.29 (3.2)**
Distant from local centre (>=27mins from a chemist)	0.14	..	0.46 (4.9)**
Public transport (>=1 bus per half hour)	0.70	..	-0.26 (3.6)**
Moved house in last 2 years	0.13	..	0.14 (1.6)
Statistics			
R ²		0.18	
N		865	

Notes: 1. *Source:* National Travel Survey 1999-2001 micro-data set from UK Data Archive; 2. Standard deviations reported only for continuous variables; 3. Data relate to day pupils in Great Britain aged 11-15.

In terms of our hypothesis, however, the key finding is that for professional/managerial families school travel distances were significantly longer from those

localities where densities were above the local authority average, whereas for other families the reverse was the case. In other words, where there is an incentive not to use the local school, there is evidence that this only affects the school choice behaviour of professional/managerial families, while others take up places in the local schools. In rural areas (where density differences may be a less useful indicator) there is also evidence of a clear class difference in behaviour, with professional/managerial families in this case sending their children over longer distances to a preferred school, possibly a selective, independent or voluntary aided school in a nearby centre, serving a wider catchment area than the local comprehensive. Both in rural areas and in the inner city, one significant implication is that an influx of gentrifiers cannot be relied on to raise the social class or educational profile of local schools. This is consistent with qualitative evidence from inner London in particular, where studies have reported very few of these incomers using local state schools for *secondary* education of their children (Butler, 2003; Buck et al., 2002)

4.3 Identifying Neighbourhood Effects

A key question in relation to the geography of school examination results is the extent to which their association with characteristics of local families simply reflects household level influences on individual performance or some additional influences arising from the interaction between local children whether inside or outside school. This question has both individual and social significance: in the first case, the relevant issue is how far the premium paid by some parents for living in catchment areas of schools which are high in local 'league tables' could actually be justified via the benefits of mixing with other able/disciplined children; in the second case, it is whether such social segregation serves to significantly widen overall disparities in educational attainment.

The obvious approach to analysing such effects is to start from pupil level data and combine this with neighbourhood/school level indicators of social/demographic composition. However, there are substantial difficulties and limitations with this approach – arising from data availability – particularly for any study attempting broad geographic coverage on recent cohorts of pupils (cf. Gibbons, 2002, using NCDS data relating to the 1970s). This justifies pursuit of a complementary approach via aggregate data.

The basis of this approach is the argument that any plausible model involving neighbourhood effects necessarily implies non-linear relations between exam success rates and the proportion of pupils/families with a relevant social/demographic characteristic – whereas, in the absence of such effects these relationships would be linear (Gordon and Monastiriotis, 2006). The non-linearity has two aspects. The more basic is that, even if the compositional influence on individual performance operates in a simply proportionate way – so that the presence of 20% of (dis)advantageous peers has just twice the effect that 10% would have – it generates an S-shaped (more or less logistic) aggregate relationship. Effectively this is because it tends to steepen the (positive or negative) relation between success rates and the proportion of pupils with this characteristic, necessitating a ‘flattening’ of the line/curve as the 0% and 100% limits are approached. The second aspect is that there may be quite disproportionate effects from the presence of *either* advantageous peers (‘good apples’) *or* disadvantageous ones (‘bad apples’): depending on which of these predominates, the effect is to add a general convexity or concavity to the shape of the relationship, notably to what would otherwise be its near linear segment.

In order to test for significant non-linearities of either kind, we return to examination of success rates for a specific attainment level (where clear 0% and 100% limits apply), specifically the median level, for which the overall success rate is closest to 50%. Where significant, but simply proportionate, neighbourhood effects operated, yielding an S-shaped relationship, we should expect here that relationships with population characteristics appeared more or less linear for most of the range – inflecting only as the upper/lower limits were approached – but that the slope of the linear segment would predict success rates outside these limits for feasible combinations of characteristics. Over the whole sample, we would not expect the simplest tests of non-linearity – via addition of a quadratic term – to produce significant results in this case, unless there was a disproportionate effect from presence of either good or bad ‘apples’.

Results of this analysis are presented in Table 7, with a data-set restricted to community/controlled comprehensives, as the school type drawing most directly from the local population. First we simply regressed success rates on a summary measure

of school characteristics plus the proportion of lone parents and the pair of composite ‘ethnic’ and ‘lower class’ variables (column 1). As earlier analyses would imply, we find strong relations with each of these neighbourhood indicators – but now we can see that the slopes are such that in combination they imply the possibility of success rates either well above or well below the logical limits of 0% and 100%. For an ‘average’ school (in terms of size and gender mix), the regression predicts a success rate of 130% for an area with no lone or working class parents, but all of them graduates, born abroad and of South Asian or black ethnic origins. At the other extreme, it predicts a success rate of –53% for an area with the reverse combination, i.e. all working class, lone parents, lacking educational qualifications, born in the UK from white³⁴ ethnic origins. These differences are highly statistically significant³⁵, and imply that at least 45% of the variation in results associated with population mix in the local area must represent neighbourhood rather than individual level effects³⁶. These results are consistent with an S-shaped relation between population mix and success at AC5 level.

Table 7. Regressions of AC5 success on rates in Community/Controlled Comprehensives on non-linear functions of neighbourhood characteristics

	1	2	3	4	5	6
Constant	73.1 (35.3)**	92.4 (50.5)**	100.7 (13.9)**	111.9 (41.1)**	132.5 (21.8)**	98.0 (13.3)**
School Characteristics	0.34 (19.6)**	0.20 (12.9)**	0.33 (19.0)**	0.19 (12.8)**	0.19 (12.7)**	0.33 (19.2)**
Lone Parent Rate (LP)	-0.62 (9.9)**	-0.11 (2.1)*	-0.75 (2.2)*	-0.11 (2.0)*	-0.11 (0.4)	-1.23 (2.9)**
Low Class Factor (Class)	-0.94 (22.3)**	-0.63 (17.4)**	-2.01 (5.8)**	-0.59 (16.4)**	-1.57 (5.6)**	-1.62 (4.6)**
High Ethnic Factor (Ethnic)	0.19 (5.0)**	0.01 (0.2)	-0.31 (4.2)**	-0.000 (0.0)	-0.17 (1.9)	-0.50 (1.9)
Absence Rate (AR)		-3.14 (25.0)**		-7.05 (14.2)**	-6.82 (13.6)**	
Weighted SEN rate (SEN)		-0.46 (10.1)**		-0.97 (8.0)**	-0.96 (7.9)**	
LP squared			0.003 (0.5)		0.001 (0.0)	0.02 (1.8)
Class Squared			0.01 (3.2)**		0.01 (3.57)**	0.01 (1.1)
Ethnic squared			0.01** (5.8)		0.01 (1.8)	0.02 (5.4)**

AR squared				0.19 (8.2)**	0.18 (7.7)**	
SEN squared				0.02 (4.6)**	0.02 (4.5)**	
LP * Class						0.000 (0.0)
LP * Ethnic						-0.02 (3.3)**
Class * Ethnic						0.02 (2.8)**
N	1915	1915	1915	1915	1915	1915
Adjusted R ²	0.43	0.61	0.44	0.63	0.63	0.45

Notes: See Table 1.

When we include absenteeism and SEN pupil rates in the regression (column 2), coefficients on all three neighbourhood variables are (again) greatly reduced, but strong slopes on the two additional variables now generate predictions with an even wider range of outcomes (over 350%!). In this case, however, the excess relates solely to the lower limit, with a linear prediction of only about 80% for schools with zero absenteeism, and negative values predicted for schools with more than about 25% absences, even when these are combined with average values for other variables. This suggests a strongly convex relationship, not simply an s-shape.

The third and fourth models test directly for non-linearity by adding squared terms, for population characteristics and absence/SEN rates respectively, to these first two regressions. In both cases the results are positive, though especially so in the latter case. Of the three socio-demographic variables, significant squared terms are found both for (low) class and ethnicity (column 3): for class this yields a relationship which is consistently negative but which weakens as the proportion of working class/unqualified residents rises; for ethnicity, however, the positive relationship only starts to take effect when at least 25% of the local population are immigrants or of black/Asian origin. For the absence and SEN variables the evidence of non-linearity is even stronger, with a convex form attenuating the initial negative impacts of both as the proportion of pupils rises (column 4). Both these effects and those of social class remain when the models are combined (column 5), but the non-linear relation with ethnicity no longer appears significant. Finally, since both absence and SEN rates are known to be affected by population composition (positively by low class and lone parenthood and negatively by ethnicity)³⁷, we fit a reduced form model relating

success rates just to neighbourhood characteristics, including squares and cross-products - and the summary measure of school characteristics (column 6). The lone parent variable now returns to significance, with a suggestion for it, as for low class, that negative effects are stronger in the lower ranges. But the only non-linear terms which are actually statistically significant again involve ethnicity, both on its own and interacted with lone parenthood or class. For comparison, column 7 presents a complementary model taking entry quality (rather than neighbourhood context) as its starting point, again showing a strongly non-linear and convex relationship, indicative of a peer group effect operating most strongly in schools achieving high quality intakes: the statistical fit of this model is evidently very much closer.

These analyses seem to provide clear evidence for presence of neighbourhood effects, accounting for a substantial part of the associations between outcomes and population characteristics reported in the previous section, partially channelled through schools' absence and SEN rates. In both of these cases the non-linear effect seems to operate *within* the school, and to show diminishing marginal effects as the absence rate, or the proportion of SEN pupils, rises. Interpreting the two variables together as representing the impact of disruptive influences on the effectiveness of schooling, these appear to follow the 'bad apple' model, with significant negative effects following from even low levels of disruption, but increasing rather less than proportionately as these rise – with little compensation to be expected from the presence of even a substantial minority of less difficult/demanding pupils or positive role models (i.e. 'good apples'). Consistent with this, the effect of the variations in the local class mix also appears to be stronger in the least working class areas. Gordon and Monastiriotis (2006) reported a similar pattern in relation to lone parenthood for schools in the London region (i.e. stronger effects in areas with fewer lone parents). However, this effect appears much weaker in the national analyses, because of significant differences between London and the rest of the country in the impact on exam success of lone parenthood (much stronger in London) and ethnicity (generally stronger in London). Within London, having a larger proportion of minority/immigrant families is associated with higher AC5 success rates, in a simple linear fashion; outside London, however, a non-linear relation is evident, with positive effects only emerging in the areas of highest concentration.

4.4 Regional and sub-regional variations

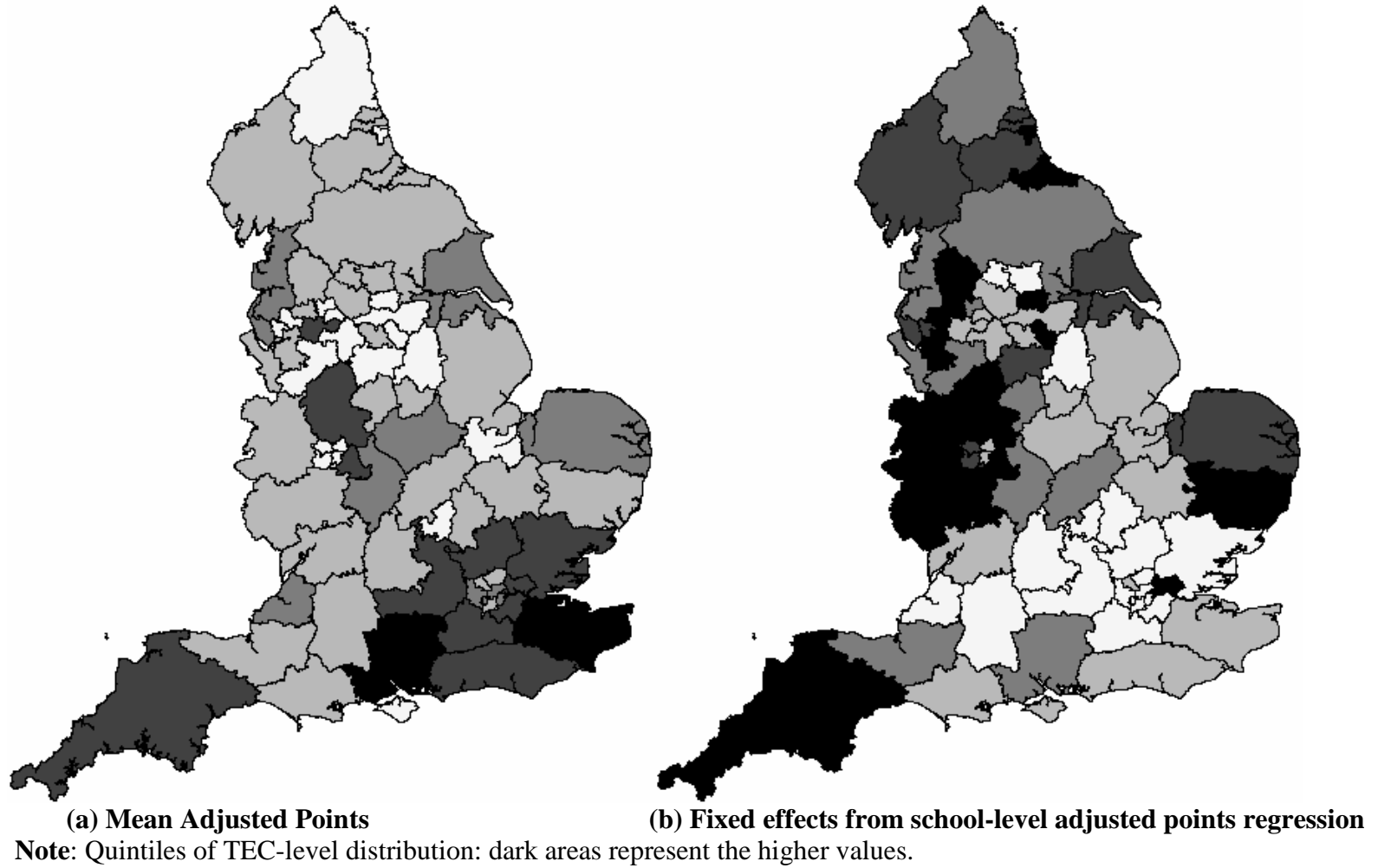
Interest in the geography of school exam performance has been concentrated up to now very largely at the intra-urban scale, rather naturally since most of the variation is to be found at this level, and because it relates most obviously here to choices that individuals (and schools) can make. However, as we have seen, much of this local variation is actually social rather than geographical in character, and more modest differences between urban regions may actually have more long run significance. Taking travel to work areas (TTWAs) as a frame of reference, it appears that a bit less than 10% of the (pupil weighted) variance in adjusted mean points across secondary schools is to be found between rather than within such areas. Though proportionately small, this still means that pupils in the top 6 TTWAs had success rates at the median qualification level (5 A-C grades) double those in the bottom 6 (74% against 38%).

A more convenient (if less functionally significant) basis for comparison is that of the former Training and Enterprise Council (or TEC) areas, which are both fewer in number and more comparable in population size. Mapping average adjusted point scores at this scale reveals a basic pattern of high values across most of southern England, together with a few more rural areas in the north, while the lowest values are found across the northern industrial cities (Figure 2a). In large part, as at the local level, this seems to strongly reflect the map of social advantage and disadvantage. This impression is largely borne out if we include (and then map) TEC level fixed effects in the school level regression analyses, to provide controls for the effects of local school and population characteristics (Figure 2b). Indeed this second map of area-specific effects is not only quite different in its basic pattern from the simple map of averages, but in some key respects actually seems to reverse it, with most of the booming 'western crescent' in and around London now figuring among the worst rather than the best performing areas. This is contrary to most of the plausible hypotheses which one might generate about likely regional influences on educational performance – for example that greater returns to skill in areas of post-industrial and high-tech opportunity would provide a stronger motivation to study.

The one clear exception to this is the proposition that in these areas commercial success is 'crowding out' public sector activity, in terms of its ability to attract and retain appropriately skilled labour for jobs which are less well rewarded relative either

to the cost of living or available private sector opportunities, and that constraints of this kind are affecting educational productivity (i.e. the ‘value’ that schooling, as distinct from the social environment, ‘adds’ to assessed performance). Possible mediating variables here would include the ability level/qualifications of the labour force attracted into public sector activities such as teaching (a factor which Nickell and Quintini (2002) have shown to operate over time at the national scale) and turnover or vacancy rates, affecting the continuity of teaching and the stock of relevant professional experience available in schools.

Figure 2. TEC-level school performance indicators



An alternative hypothesis is that the partial reversal of the performance map after controlling for influences evident at the school level occurs because some of the factors which vary much more strongly *within* sub-regions simply do not operate with the same force *between* sub-regions. The obvious example, which we have already tried to address, is that of the ‘creaming’ of talent from other nearby schools by those with some institutional advantage allied to a degree of control over pupil selection. Even in relation to school type, this spatial externality may have been incompletely controlled for in our school-level analyses. It is also possible that there are other comparable effects linked to the reputation and pupil composition of particular schools with indirect effects on others which we have been unable to consider.

To explore these issues we have undertaken some regressions at TEC level³⁸, again taking adjusted points as the dependent variable. For a baseline regression we related this simply to a pair of composite variables derived from the school level regressions, respectively representing effects expected from:

- school characteristics (other than absence rates and the proportion of SEN pupils) and accessibility to advantaged school types; and
- the population mixes of surrounding areas.

If there are no additional area effects, we should expect the coefficients on each of these in adjusted score regressions across TECs to be 1.0. However, that on the composite school variable proved to be effectively zero, while that on the neighbourhood composite, though still very strong, was significantly reduced, by about a quarter (Table 8, column 1).

Table 8. TEC-level regressions of average adjusted points on school, neighbourhood and labour market characteristics

	Model 1	Model 2	Model 3	Model 4	Model 5
Constant	44.5 (1.4)	122.6 (1.7)*	140.4 (2.5)**	..	-50.5 (4.3)
School characteristics					
Composite	0.09 (0.3)				
Selective schools (prop. of pupils)		52.71 (1.1)	68.19 (2.1)*	77.59 (2.3)*	96.71 (3.4)**
Co-educational schools (prop. of pupils)		12.03 (0.4)			
Independent Schools (prop. of pupils)		-45.93 (0.7)			
Mean pupils in year (logged)		-22.52 (2.0)	-17.78 (1.7)	-5.21 (0.4)	
Neighbourhood characteristics					
Composite	0.73 (10.0)**				
Low Class Factor		-21.81 (5.6)**	-22.91 (9.0)**	-22.26 (7.9)**	-23.02 (8.9)**
High Ethnic Factor		1.35 (0.5)			
Lone parents (%)		0.34 (0.5)			
Labour Market			52.22 (2.2)*	44.50 (1.9)	50.27 (2.1)*
Teacher pay ratio			-2.35 (2.2)*	-2.58 (1.2)	-2.78 (2.6)*
Teacher turnover rate (standard region)			-5.85 (3.0)**	-4.05 (1.9)	-6.31 (3.3)**
Change in teacher vacancy rate 2001-4				0.98 (1.2)	
Region dummies (GO regions)					
Observations	72	72	72	72	72
Adjusted R ²	0.60	0.65	0.68	0.69	0.67

Note: See Table 1, note 1.

Splitting each up to distinguish some or all of their major components, we found (column 2), among the school characteristics that: gender and independent school effects were quite insignificant at this level; there was still a net positive effect associated with the presence of selective schools (if not statistically significant); while school size now seems to be exerting a negative rather than positive effect (at the

margin of significance). Among the grouped neighbourhood characteristics: both lone parenthood and the presence of ethnic minorities lost their role at this more aggregate level; but the low class factor displayed very much the same effect as at school level, emerging as clearly the dominant influence. A third regression (column 3) retaining the more significant of these elements and adding three indicators relating to teacher supply – their pay relative to all local professionals, their turnover rate at regional level, and local increases in the teacher vacancy rate - finds some significance for each of these (as well as now for the selective school variable). Inclusion of a set of (government office) regional dummies³⁹ as a check on the robustness of these results (column 4) unsurprisingly lowers the significance of the each of the supply variables (whose variance is largely inter-regional), but its only substantive effect is to reduce the school size effect to insignificance. Removing this (and the regional dummies) leaves a model in which TEC level school performance is still primarily influenced by their social class composition, but also by the presence of selective schools and by teacher supply factors (column 5).

5. Conclusions

The relationship between schooling and geography is a two-way one. On one side, higher rates of exam success are believed to be associated at a local level with a socially superior residential mix, and at a broader level with stronger economic performance. On the other side, the local population mix, labour market pressures and local educational policy affect both school level success rates and the conditions under which schools operate. In this paper we have focused on evidence relating to the latter side – how location affects education – from English secondary school exam results. We have explored this issue at two different spatial scales, firstly in relation to the effect on individual school results of characteristics of families resident in immediately surrounding areas from which pupils may be drawn, and secondly in relation to the aggregate pattern of results across much broader sub-regions.

As with the array of studies attempting to identify the effect of school types, schooling inputs and/or school ‘effectiveness’, the basic problem remains that of reliably controlling for the impact of personal/family background factors on the performance of individual children, in order to distinguish the contribution of the other factors of interest (i.e. here that of geographic context). There is no ‘royal route’

to achieving this, with 'state of the art' micro-methods such as multi-level analysis or controlled experiments still being open to significant biases and limitations in practice. Here we have pursued a complementary approach, based on looking for non-linearities in aggregate relationships between school outcomes and community characteristics. Despite serious reservations about the direct meaningfulness of the newly published official estimates of secondary-school 'value-added', we have made use of these to recover (reasonable) measures of school intake quality for use in these analyses.

At the individual school/neighbourhood level, we have found strong evidence of non-linear relations; firstly with the ethnic and class composition of the surrounding area; secondly with school absence and SEN pupil rates; and in a third analysis with the pupil intake quality of the school. Both the second and third of these represent channels through which neighbourhood effects are mediated, as well as schools' differential successes/failures in pupil management and recruitment. The 'social class' relation seems primarily to reflect the educational background of local parents rather than relative affluence, with the effect of better qualified local parents being realised primarily through stronger pupil intakes into neighbourhood schools. The 'ethnicity' relation was also generally positive (perhaps less predictably) in the sense that, above a certain threshold level, schools in areas with more black or Asian families and parents born abroad tended to show better GCSE results. In this case the key intervening variable was not intake quality but school absence rates, which tended to be significantly lower in areas with substantial minority populations, and which appeared to have a powerful, non-linear effect on results, whether directly or as an indicator of broader disciplinary problems in the schools concerned. As more qualitative evidence from interviews with parents and teachers also suggests (Buck et al., 2002) the disciplinary order of schools is a crucial influence on what pupils can achieve, and one which reflects local social conditions and family structures – though competitively successful schools, in terms of achieved intake quality have fewer disciplinary challenges to face. Indeed the one clear conclusion which can be drawn from the 'value-added' statistics is that school outcomes at GCSE level are largely predictable from the quality of their intakes, with strong peer group effects operating in those schools with the better intakes in terms of their primary school (KS2) results. The convexity of the relation between GCSE results and intake quality, as in those

with attendance rates and with better qualified local parents, indicates that the stronger neighbourhood/peer group effects operate among the most, rather than the least, advantaged areas and schools. This is more consistent with the hypothesis that a few disruptive 'bad apples' have more impact on the achievement of others than an equivalent number of positive role models. It may also offer an explanation of Cheshire and Sheppard's (2005) finding that proximity to the *most* successful schools is associated with a disproportionate mark-up in house prices, since at the top-end of the distribution of results (and intake quality) a greater part of the differential in performance would reflect externalities available to those able to gain access⁴⁰.

At the sub-regional level, the one socio-demographic characteristic of the area's population which retains its significance is the parental 'class' (or qualification) variable. This goes a long way toward explaining the generally superior performance in GCSE terms of schools in South East England (as well as those outside the conurbations). Another significant factor tending to reinforce the regional contrast, however, seems to be that of selective education, with better overall results appearing in sub-regions with significant numbers in selective schools – implying, as the neighbourhood analyses also suggested, that gains to those attending such schools are larger, at least in terms of overall 'points' than losses to those in the other ('creamed') schools. However, when population and school characteristics are controlled for, we find that results are significantly worse than expected in the most prosperous areas of southern England, apparently for reasons of constrained teacher supply, in areas where their pay rates are uncompetitive. Thus while the general feedback relation between education and locational success seems to imply a widening gap between north and south in both terms, this crowding out factor seems to imply either a constraint on the south's growth potential, or a reason for it to rely increasingly on human capital produced in the schools of the north, to fill its skilled jobs.

References

- Atkinson, A. and Wilson, D. (2003) 'Measuring Pupil Attainment in English Secondary Schools: A Preliminary Analysis', *CMPO Working Paper Series*, 03/063, University of Bristol.
- Barrow, L. and Rouse, C.E. (2005) 'Causality, causality, causality: the view of education inputs and outputs from economics', *WP-05-15*, Federal Reserve Bank of Chicago.
- Benton, T., Hutchison, D., Schagen, I. And Scott, E. (2003) *Study of the performance of maintained secondary schools in England: report for the National Audit Office*, London: National Foundation for Educational Research.
- Berry, C.R. and Glaeser, E.L. (2005) 'The divergence of human capital levels across cities', *Papers in Regional Science*, 84, 407-444.
- Blanden, J. and Gregg, P. (2004) 'Family Income and Educational Attainment: A Review of approaches and Evidence for Britain', *CMPO Working Paper Series No. 04/101*, University of Bristol.
- Brasington, D. and Haurin, D.H. (2005) 'capitalization of parent, school, and peer group components of school quality into house price', Economics Department Working Paper 2005-04, Louisiana State University.
- Buck, N., Gordon, I., Hall, P., Harloe, M. and Kleinman, M. (2002) *Working Capital: life and labour in contemporary London*, London: Routledge.
- Burgess, S.A., Gossage, D. and Propper, C.(2003) 'Explaining differences in hospital performance: does the answer lie in the labour market?', *CEPR Discussion Paper*, 4118, Centre for Economic Policy Research.
- Butler, T. (2003) *London Calling: the middle classes and the re-making of Inner London*; Oxford: Berg.
- Cheshire, P. and Sheppard, S. (2005) 'Capitalising the value of free schools: the impact of supply characteristics and uncertainty', *Economic Journal*, 114, F397-F414.
- Dearden L., Ferri J.and Meghir C. (2002) 'The effect of school quality on educational attainment and wages', *Review of Economics and Statistics*, 84, 1-20.
- Feinstein, L. and Symons, J. (1999) 'Attainment in Secondary School', *Oxford Economic Papers*, 51, 300-321.
- Gibbons, S. (2002) 'Neighbourhood effects on educational achievement: evidence from the Census and the National Child Development Study', *Centre for the Economics of Education Discussion Papers*, LSE18, London School of Economics.
- Gibbons, S. and S. Machin, S. (2003) 'Valuing primary schools', *Journal of Urban Economics*, 53, 197-219.

Glaeser, E.L. and Saiz, A. (2004) 'The Rise of the Skilled City', Brookings-Wharton Papers on Urban Affairs.

Gordon, I.R. and Monastiriotis, V. (2006), Urban size, spatial segregation and educational outcomes, *Urban Studies*, 43 (1), 213-236.

Hanushek, E.A. (1979) 'Conceptual and empirical issues in the estimation of educational production functions', *Journal of Human Resources*, 14, 351-388.

Lupton, R. (2005) 'Social justice and school improvement: improving the quality of schooling in the poorest neighbourhoods', *British Educational Research Journal*, vol 31 (5): 589-604

Lupton, R. (2006) 'How does place affect education', in S. Delorenzi (ed) *Going Places: neighbourhood, ethnicity and social mobility*, London: Institute of Public Policy Research.

Monastiriotis, V. (2002) 'Human capital and wages: evidence for external effects from the UK regions', *Applied Economics Letters*, 9, 843-846.

Nickell, S. and Quintini, G. (2002) 'The consequences of the decline in public sector pay in Britain: a little bit of evidence', *Economic Journal*, 112, F107-F118.

Reed, J. and Hallgarten, J. (2003) 'The time to say goodbye? The future of school performance tables', London: Institute of Public Policy Research.

Smithers, A. and P. Robinson (2001) *Coping with Teacher Shortages*, report for National Union of Teachers, Centre for Educational and Employment Research, University of Liverpool [<http://www.data.teachers.org.uk/resources/pdf/Shortages.pdf>]

Stillman, S. (1990) 'Legislating for choice', pp. 87-106 in Flude, M. and Hammer, M. (eds.) *The Education Reform Act, 1988*, London: Falmer Press.

Taylor, C. (2002) *Geography of the 'New' Education Market: secondary school choice in England and Wales*, Aldershot: Ashgate.

Taylor, J. and Nguyen, A.N. (2006) 'An analysis of the value added by secondary schools in England: is the value-added indicator of any value?', *Oxford Bulletin of Economics and Statistics*, 68 (2), 203-224.

Wilson, D., Burgess, S. and Briggs, A. (2005) 'The Dynamics of School Attainment of England's Ethnic Minorities', *CMPO Working Paper 05/130*, University of Bristol.

Notes

¹ An earlier version of this paper was presented to the Geography of Education session, at the Institute of British Geographers annual conference, London, September 2005. We are very grateful for the helpful comments of other participants in this session.

² Encapsulating an important lone parent effect, whether through lower time availability or problems in securing discipline over homework.

³ A binary measure of free school meal entitlement.

⁴ A category largely comprising Black British with one or both parents of Caribbean origin.

⁵ Because the method is predicated on value-added within a stage being independent of measured performance at the end of the previous stage.

⁶ We are grateful to Steve Gibbons for this point.

⁷ Those with technology specialisms gained in both stages, but by significantly more in post-KS3.

⁸ The negative effects associated with having more pupils eligible for free school meals or with Special Educational Needs (SEN) appear to reduce by three quarters between these two stages; those with more Afro-Caribbean pupils start to gain, while positive effects associated with the proportion of pupils from Bangladeshi, Pakistani or 'other' ethnic origins more than double.

⁹ And (in the 2005 data, where a distinction is made) to have their 5 A-C grade pass rate lowered most when this is required to include English and maths passes.

¹⁰ With some tendency for higher appeal rates in more ethnically diverse education authorities.

¹¹ For reasons explained below (and in footnote 17) our statistical analyses omit about one third of the schools, while still covering 93% of pupils in the age cohort. Comparable data are not available on a comprehensive basis for schools in Scotland, Wales or Northern Ireland, where the devolved administrations had all decided by 2003 not to publish secondary school results in league table form. Scottish results *are* still available on a school by school basis, but relate to a different set of examinations and standards than the English and Welsh GCSE system.

¹² The basic and median standards involve 5 or more passes at or above particular grade thresholds: G (the former CSE pass standard) in the 'basic' case and C (the former GCE pass standard) in the 'median' case; the minimal standard involves at least one CSE level pass.

¹³ Actually there are two indicators, one for passes at A and above, and one for A* passes; we concentrate here on the former, as closer to the *Telegraph's* measure of high performance

¹⁴ The simple correlation between the two is 0.964 rising to 0.986 when six apparently deviant observations are removed from the *Telegraph* data set.

¹⁵ On our estimates, around 80% of schools outside the *Times/Telegraph* lists would have success rates below 90% of those included there, though the *Telegraph* list in particular includes a small number of schools with success rates toward the bottom of the overall distribution, down to 6% - against a size weighted average of 36% for the two lists, and (we estimate) about 12% for other schools.

¹⁶ Predicted values from the regression were deflated by a factor (1.224) representing the coefficient on VAI in a regression of predicted points on school/area characteristics and VAI: this compares with a national average ratio between average total and capped scores of 1.202.

¹⁷ This down-weighting, and the exclusion of wholly boarding schools were based on the number of day pupils recorded in the tables as a denominator for the absence rates. Use of this criterion (and of the absenteeism rates in our analysis) led to the omission also of some schools in both independent and state sectors that seem simply to have failed to return this information, though formally required to do so.

¹⁸ Where the 'household representative person' was born outside the UK, Ireland, USA or Old Commonwealth.

¹⁹ The qualifications measure actually related to adults aged 35-49 (the most likely age band for parents of 16 year olds), while the rest were based more directly on households, or their 'representative person' with dependent children aged 12-14.

²⁰ With a parameter of -0.9 on distance in miles for an exponential distance decay function.

²¹ The ML version of the Heckman procedure in Stata, with a probit selection equation.

²² The selection equation showed a predictably strong bias toward inclusion of schools with strong GCSE records, with biases also toward larger schools, in areas with highly educated populations.

²³ This description implies that all pupils entered the schools concerned at age 11. In fact, not all pupils actually change school at this point, and among those who do some will subsequently have changed school and/or area. This complicates the picture for individual schools, without distorting the main patterns evident in the data.

²⁴ These are based respectively on the principal component of the two qualification variables and the working class share (low class factor) and of the Asian/black minority variables plus the proportion born abroad (high ethnic factor). They were scaled so as to yield values of 0 and 100 for the logical extremes of each.

²⁵ In part because, at an individual level, girls' academic progress is faster in this phase of development (Atkinson and Wilson, 2003).

²⁶ Though this results hinges on the appropriateness of official measures of value-added between KS3 and GCSE.

²⁷ This dismissive expression came from Tony Blair's press spokesman in 2001.

²⁸ i.e. the distance parameter in the exponential function is -0.3 rather than -0.9 .

²⁹ Actually as specified this negative effect applies to all schools, so that the basic advantage of (say) a selective school in intake terms will be reduced to some extent in areas where more children go to such schools.

³⁰ The t statistics for the sums of the paired coefficients are 4.1 (selective), 2.6 (CTC) and -1.9 (voluntary/funded).

³¹ The relevant t statistics for the paired coefficients are: $+3.9$ in the entry equation and $+2.4$ in the VA equation for selective/CTC; and -1.6 (entry) and $+0.1$ (VA) for voluntary/funded schools.

³² The survey covers a week of travel by individuals coded by trip purpose and time/distance. In order to cut out travel to boarding schools or travel linked to weekend/holiday excursions, we focused on trips made to school on the Wednesday morning of the recorded week, and included only individuals for whom there was a record of at least three such trips during the week. This left data on school travel by 865 individuals in the relevant age band (11-15).

³³ These are actually coded categorically, into 13 and 10 groups respectively; for this analysis we have used midpoint densities for each category.

³⁴ The defined category actually includes 'other' non-black, non-South Asian ethnic origins (e.g. Chinese), but the results basically relate to the white population, and the association may well be reversed for some or all of the 'other' groups.

³⁵ Since we are extrapolating well outside the observed range, we use a prediction error measure based on the standard errors of the three slope coefficients and the correlation between these.

³⁶ The proportion is *at least* 45% since no more than 100 points out of the 183 point range predicted from extreme values of the three variables could be generated from individual level effects.

³⁷ Though in the case of SEN rates, this is only true after controlling for a (positive) effect from the incidence of language problems.

³⁸ TEC averages were constructed using weights based on the estimated numbers of 15 year old day pupils in each school, and omitting special schools (for comparability with the school level analyses).

³⁹ None of which were individually significant.

⁴⁰ This is an alternative to the authors' own explanation in terms of the 'positional' value ascribed to the topmost level of achievement.