School Performance, Score Manipulation and Economic Geography*

Erich Battistin†

Queen Mary University of London, CEPR, IZA and IRVAPP

Lorenzo Neri‡

Queen Mary University of London

October 2017

Abstract

We show that grading standards for primary school exams in England have triggered an inflation of quality indicators in the national performance tables for almost two decades. The cumulative effects have resulted in significant differences in the quality signaled to parents for otherwise identical schools. These differences are as good as random, with score manipulation resulting from discretion in the grading of randomly assigned external markers. We find large housing price gains from school quality improvements artificially signaled by manipulation, as well as lower deprivation and more businesses catering to families in local neighborhoods. The design ensures improved external validity for the valuation of school quality with respect to boundary discontinuities, and has the potential for replication outside of our specific case study.

JEL Classification: C26, C31, I2

Keywords: School quality, House prices, Score manipulation

---

*Special thanks go to Barbara Donahue, Rebekah Edgar, Rebecca Evison, Martin Harris and Tim Leunig for advice and guidance in interpreting the results from the National Pupil Database. Our thanks to Ghazala Azmat, Jad Beyhum, Dan Black, Francesca Cornaglia, Francesco Fasani, Seth Gershenson, Stephen Gibbons, Ellen Greaves, Rachel Griffith, Stephen Machin, Marco Manacorda, Marco Pagano, Elisabetta Pasini, Barbara Petrongolo, Steve Pischke, Imran Rasul, Enrico Rettore, Olmo Silva, Jeffrey Smith, Andrea Tesei, and seminar participants at the University of Naples Federico II, the 2016 Workshop on Labour and Family Economics (WOLFE), 2017 RES Symposium of Junior Researchers, 2017 IZA Summer School in Labour Economics, the 16th LAGV Conference in Public Economics, 8th International Workshop on Applied Economics of Education (IWAEE) and the 70th European Meeting of the Econometric Society (ESEM) for helpful discussions and comments. The views expressed here are those of the authors alone.

†School of Economics and Finance, Queen Mary University of London, Mile End Road, London E1 4NS, UK. Contact: e.battistin@qmul.ac.uk.

‡School of Economics and Finance, Queen Mary University of London, Mile End Road, London E1 4NS, UK. Contact: l.neri@qmul.ac.uk.

1 Introduction

Understanding household preferences for neighborhood attributes is of fundamental importance. A substantial portion of household expenditure is devoted to housing. Differences in house prices and economic activities between locations shape residential choices, influencing urban and suburban sprawl and the social consequences associated with such expansion. Among a number of local amenities, there are good reasons to value proximity to good schools in particular. In many countries, home residence is the way that most children gain access to public schooling. In England, the context considered here, boroughs require evidence of tax payment and electoral-roll registration as proof of address. As a consequence, parents may be prepared to pay a substantial premium to secure an address within a desirable school catchment. Mobility motivated by school quality results in residential sorting into communities with similar tastes for area characteristics such as open space and child-friendly amenities.

Disentangling the effects of area composition from those of school quality on house prices and local development is the subject of a sizable empirical literature (see Black and Machin, 2011, and Duranton and Puga, 2015). We know that academic standards are valued by parents (Hastings and Weinstein, 2008), and these standards have important effects on the willingness to pay for a good school. The review in Machin (2011) suggests a 3% house price premium from a one-standard deviation increase in test scores. The identifying variation for assessing causality in most related studies stems from price differences across school admission boundaries. Regression discontinuity-style approaches have been used to identify the effects of school quality in many countries, including the US (Black, 1999, Bayer et al., 2007), France (Fack and Grenet, 2010) and the UK (Gibbons and Machin, 2003, and Gibbons et al., 2013). Alternative quasi-experimental designs exploit school openings and closures or changes to school boundaries (Ries and Somerville, 2010, and Tannenbaum, 2015).

The fact that parents value quality and that this value is capitalized in house prices have increased the focus on the role of information provided about schools. Though in many countries, performance indicators are readily accessible through local authorities or central governments, whether or not parents' perceptions of school performance reflect the actual school quality is a question that challenges how comparative information is acquired and interpreted. There is evidence that the housing market responds to information beyond the signaling value of standardized testing. For example, Figlio and Lucas (2004) find that news about school grades, the evaluation issued by the education department using test scores and other public available data, have an effect on local house prices and transactions. Mizala and Urquiola (2013) show that flagging top-performing schools using a measure of value added, as opposed to absolute performance, does not have an effect on school enrollment. The housing market in England also responds rapidly to new information about school performance and ratings, as we show below (see also Hussain, 2017).

Standardized testing for evaluation purposes has been in place in England since the early 1990s and provides the yardstick by which schools are assessed and compared. Examination results are used by the Department for Education to form performance tables, an accessible source of comparative data to which considerable attention is given by the media and local authorities. These tables are updated every year and contain indicators of performance along with contextual data on school environment, staff and finances. Household preferences recovered from survey data confirm that among school attributes, parents strongly value academic quality in determining the choice of school (Burgess et al., 2015 and Hussain, 2015). Performance tables report the fraction of students scoring above subject-specific national targets at the end of primary school, as we explain in the next section. The differences between schools in students attaining these targets are the indicators of quality used in our analysis (and the headline figure published by the media; see Black and Machin, 2011).[1]

We use unintended and unforeseen consequences of the marking scheme for standardized tests to study how information on the quality of primary schools affects residential sorting, house prices and local development across narrowly defined neighborhoods in England. Exams are proctored locally and marked externally by an agency appointed to maintain and develop the national school curriculum and the educational assessments. Exams are randomly assigned to markers, grading is blind and the score thresholds used to award national targets are disclosed by the Department for Education only at the end of the process. Since the mid-1990s, students who narrowly missed the target had their exams reviewed by the same marker, but those who barely passed the borderline kept their result without additional scrutiny. This re-scoring, known as borderlining, was limited to exams within three marks of

[1] Table layouts and the information reported underwent minor changes over the years considered in our analysis. A detailed description of the content and changes to the information published can be found at http://www.education.gov.uk/schools/performance/archive/index.shtml.

the proficiency cutoff while leaving the rest of the distribution unaffected (a similar procedure is implemented for New York's Regents exams, as explained in Dee et al., 2016).

This practice, originally introduced to avoid unfair denial of levels because of low marking quality, was dismissed in 2007 because it was found to be responsible for boosting the results for thousands of students and overstating school standards for over two decades (Department for Children, Schools and Families, 2009). Evidence of score inflation around achievement thresholds is shown in Figure 1, which presents the language score distributions computed from national tests in selected years before and after the removal of borderlining.[2] Continuous lines are obtained from a local-linear fit estimated excluding scores that are within three points of the Level 3 ("working towards expected level"), Level 4 ("expected level") and Level 5 ("exceeded expected level") thresholds, which are denoted by vertical lines. The dots represent the share of exams with given scores on the national test. Bunching is evident to the right of the critical pass thresholds, with no student downgraded. Our calculations show that students who took the English exam were 4 times as likely to receive the minimum passing score to attain Level 5 than to receive the score one point lower. Our calculations also show that approximately 19% of the exams that were re-scored below Level 5 were eventually inflated. This manipulation fades away after 2007, when re-scoring was abolished.

We argue here that score manipulation reflects marker behavior - specifically, a leniency in grading for borderline students. Blind marking, combined with the random assignment of markers, yields score inflation independent of demographics and school characteristics, adding random noise to the share of students achieving at the critical level and changing the perception of school quality. We study how this noise affects local house prices, residential sorting and, ultimately, the beautification of neighborhoods and the opening of brick-and-mortar businesses (retailers, restaurants and recreational facilities) catering to the needs of families. The identifying variability follows from the random assignment of exams to markers, year-to-year variations in a school's number of students scoring below critical thresholds, and changes made by the Department for Education to achievement targets over time and across school subjects that cannot be anticipated by markers.[3] The width of the manipulation

[2] The KS2 data used here are described in the next section and in documents and links at https://www.gov.uk/government/collections/national-pupil-database. Figure C.1 and Figure C.2 in Appendix C present the distributions for math and science scores.

[3] Chetty et al. (2013) use bunching in income distributions across neighborhoods to proxy for local information about the marginal rate structure of the Earned Income Tax Credit. Fuzziness in the quality of

region is the result of grading protocols, and this feature simplifies estimation with respect to other empirical research (Diamond and Persson, 2016 is an example).

We use a number of administrative sources and exploit the multi-layer geographic hierarchy developed by the Office for National Statistics for the production of their official statistics. This approach allows us to measure residential sales, the socio-economic characteristics of the population and the economic activity of business organizations across narrowly defined areas of the same neighborhood and to assess the quality of accessible schools on a block-by-block basis. The analysis controls for unobserved attributes of the local neighborhood (as in Bayer et al., 2007, and Caetano, 2016), and offers a new research design that has the potential for replication outside of our specific case study and yields causal estimates with improved external validity compared to those from boundary discontinuities. Indicators of school quality and the geocoding of residential locations are derived from the National Pupil Database, a dataset with standardized scores for all students in England over the past two decades.

We show that the manipulation-induced noise that accumulated in the performance tables over the decade from 1998-2007 is capitalized in future house prices. A three-standard deviation increase in the signal artificially generated by borderlining (approximately 3 percentage points in the share of students attaining above expectations in English over a baseline of 26%) yields an house price premium between 1% and 1.5% depending on the specification considered. Our calculations below show that these estimates suggest a 5% house price premium from a one-standard deviation increase in the share of students attaining Level 5 in language, a valuation larger than that found in previous studies for England (see Gibbons et al., 2013). This result likely follows from improved external validity and from the fact that the returns considered here are for performance above expectations as opposed to average performance.

We investigate the role of alternative drivers of residential sorting, and conclude that school quality is the main determinant of our findings. We document the displacement of residents by an influx of more economically affluent households with school-age children, targeting neighborhoods independently of trends in local development. For example, we don't find strong evidence for differential effects of the noise created by manipulation in the most up-and-coming blocks for economic activities and lower deprivation. On the other hand,

schools signaled by matriculation rates in Israel is considered in Lavy (2009).

we find that the effects of borderlining on house prices increase to 3% for blocks belonging to a number of school catchments above the sample average, a monetary equivalent of £7,110 using 2015 figures. This is consistent with the idea that parents may hedge to widen the set of possible choices at school enrollment. In addition, council tax records reveal that the house price premium is not related to any improvements or beautification of existing dwellings.

Finally, we show that the effects of score manipulation spill over to the demography and the activity of local businesses. A three-standard deviation growth in the quality artificially signaled through borderlining increases by 15% the number of retailers, restaurants and coffeehouses surrounding schools. Census data on business activities show that workplace employment increases by approximately 22% (or 27 employees), across a number of economic activities. Other research has demonstrated that household consumption has an important local component (see, for example, Agarwal et al., 2017, and references below); in England, in particular, the majority of customers visiting convenience stores are households living within 0.25 miles.

Why should the fact that the effects of manipulation spill over to house prices and local neighborhoods be of interest to policymakers and researchers? As testing regimes have become the accepted means of measuring school quality, so have concerns about the reliability and fidelity of assessment results (see Neal, 2013, and Battistin, 2016). Substantial cheating on standardized tests was documented by Jacob and Levitt (2003) for Chicago schools. A recent system-wide cheating scandal in Atlanta, described as one of the largest in United States history, sent some school administrators and teachers to jail. Dishonesty is not the only force driving manipulation, and the proficiency levels used for school accountability lead schools to focus more on the "marginal student" (Neal and Schanzenbach, 2010) sometimes for reasons unrelated to NCLB-style accountability. Studies have documented that discretion in grading may favor students with certain characteristics, yielding discontinuities around pass-or-fail thresholds similar to those documented here (see Dee et al., 2016 for New York Regents Examinations, and Diamond and Persson, 2016 for Sweden).[4] Our analysis uncovers a new substantive problem inherent to the reputation of school districts when admission is based on school boundaries and shows that score manipulation can trigger social

[4]Additional examples of student discrimination by teachers were documented for England (Burgess and Greaves, 2013), France (Terrier, 2016), Greece (Lavy and Megalokonomou, 2017), Israel (Lavy, 2008, and Lavy and Sand, 2015) and India (Hanna and Linden, 2012, and Bhattacharji and Kingdon, 2017).

inequality across local neighborhoods, possibly with consequences for students displaced to schools in higher-poverty areas (Chetty et al., 2016).

The remainder of the paper is organized as follows. The next section presents the institutional background on schools and tests in England. Section 3 describes our data and the sample selection criteria. Following a graphical analysis, Section 4 documents the effects of borderlining across schools. Section 5 discusses the source of the identifying variability. The empirical specifications and results for house prices, school composition and local development are presented in Section 6. The conclusions and directions for further work are presented in Section 7.

2 Background and Context

The National Curriculum Assessments in England

School age in England begins with the term following a child's fifth birthday, and education is compulsory until age 16. Primary education consists of two blocks of years: Key Stage 1 (KS1; ages 5 to 7) and Key Stage 2 (KS2; up to 11). The former phase includes reception, which is delivered as pre-school, and two years of formal education known as Year 1 and Year 2. KS2 runs from Year 3 to Year 6. The National Curriculum establishes standardized programs of core knowledge and attainment targets for all subjects at both cycles.

Our analysis considers public schools, which enroll over 95% of students in the country (Department for Education, 2016), and for which coordination and financial support lies with the local authority (LA).[5] Community schools, by far the most common, are established and fully funded by LAs. Faith schools, originally established by voluntary or religious bodies (e.g., churches), are more independent but still largely funded by LAs. Among the remaining state-funded schools, foundation schools and academies have the greatest freedom in management. In particular, academies (akin to charter schools in the United States) are independent of LA control and do not have to follow the National Curriculum. With this exception, all remaining schools must follow precise guidelines for core subjects. Our working

[5]Fack and Grenet (2010) show that private schools in the neighborhood decrease the house price premium associated with public schools quality. In addition to low enrollment, private schools in England do not have to follow the National Curriculum or participate in the national KS2 evaluation. Because of this, data on these schools are not published in performance tables.

sample retains community, faith, and foundation schools and excludes a limited number of institutions providing education to children with special needs. Importantly, the number of school openings and closings in the time window considered was negligible, as were all major changes following the introduction of academies in the late 2000s (see Eyles and Machin, 2015).

The criteria for school entry are regulated by LAs. Priority is given to children with special education needs or with siblings at the school. Faith schools are also allowed to enroll students on the grounds of the religiosity of the parents. Other than these pathways to entry, the most common way of prioritizing applications is geographical proximity to the school. It follows that school catchments may vary across areas, with an average radius of 0.67 miles in our sample, and are not rigidly defined attendance zones as are common in the United States. However, catchment areas are by and large within the LA: only 3.6% of students in our data attend schools other than in their home LAs. Although there are no legal restrictions on school choice, applications outside the LA are burdensome, and LAs do not have the statutory requirement to find a school for children from a different district (see Burgess et al., 2015, and Gibbons et al., 2013).

Academic assessment is statutory at the end of each stage of education, and attainment targets are set using six progressive levels of learning (from Level 1 to Level 6). Level 2 is expected by the end of KS1, and Level 4 is expected at KS2. Important changes to the measurement tools used to assess progress have been made in the past decades. Nationwide standardized testing in KS1 was phased out in 2004 in favor of decentralized assessments from teachers. The Standards and Testing Agency (STA), a government body charged with educational assessment, provides teachers with standards against which a child should be assessed in KS1. Nonetheless, teachers are free to make judgments based on their knowledge of the student.

Standardized testing in KS2 has been conducted continuously in the three core subjects (English, mathematics and science). Because of this, KS2 results are key for accountability purposes. Minimum levels of quality are regularly set by the Government using KS2 scores to hold schools responsible for their performance. In addition to a number of contextual indicators, performance tables report the share of students at or above Level 4 and Level 5 every year for all state-funded schools as well as an overall score obtained by combining these shares across subjects.

It is well documented that high academic standards are considered to be the most important school attribute by parents in England, followed by socio-economic composition and proximity (see Burgess et al., 2015 and Hussain, 2017).[6] School quality is quite heterogeneous within LAs, suggesting that parents may selectively target neighborhoods in specific catchments rather than just the school district: for example, the within-LA variability in the share of students performing above expectations is 87% of the total variation in our working sample.

Grading Protocols and Borderlining

KS2 tests are proctored locally and marked externally by an agency appointed by the Department for Education.[7] LAs or the STA can make unannounced visits on the test day to ensure that test protocols are implemented correctly. Grading is blind and carried out without knowledge of the thresholds required to award achievement levels.[8] Mark boundaries are set by senior examiners and disclosed only at the end of the process. Thresholds change every year and across subjects, and official documents and our data offer no evidence that these boundaries can be predicted. Once marking is concluded, schools can request a review of their scripts for a fee, which is refunded in the event of a successful appeal. Tests are a combination of multiple choice questions and open-response items, for which a more intense grading effort is required. This opens the door to interpretation and opinion, perhaps differentially across school subjects (we show examples of items in Appendix C). Scoring materials and instructions are provided to markers to enforce consistent grading, including

[6]KS1 results were used in performance tables to publish indicators of value added in selected years. The computations, however, did not follow a consistent methodology in the time window considered here. The evidence on whether parents value, or fully understand, value-added indicators of school quality compared to test scores is mixed (Black and Machin, 2011). Mizala and Urquiola (2013) show that providing parents with school rankings based on value added does not lead to shifts in choice toward higher-scoring schools. Gibbons et al. (2013) cannot reject the hypothesis that the coefficients for value added and scores in hedonic regressions are equally valued. In contrast, the results in Wilson et al. (2006) suggest that value added is not a determinant of residential choices in England.

[7]The agency has changed over time but did so after the period relevant to our analysis. The National Assessment Agency was in charge until 2007, at which point the Qualifications and Curriculum Development Agency took over.

[8]Burgess and Greaves (2013) refer to KS2 as "almost-blind" grading because markers can see the name of the student on the script and infer ethnicity. However, we do not find discontinuities around attainment thresholds considering a long array of demographics, including ethnicity (see Table 3 and appendix Figure C.3).

examples and precise guidelines on how to interpret potentially ambiguous answers. Graders have no relationship with the school and are paid to complete training. Graders can expect to earn between £500 and £1000 per examination series, although these fees are per script or item and are determined by the size of the paper.

Since tests were instigated in the 1990s, to avoid students being unfairly denied a level, all exams falling *three* points or less below the pass-mark were revisited by the original marker; exams falling above were not. This re-scoring, known as borderlining, was abolished in 2007. It has been estimated that 300,000 students were upgraded between 1996 and 2007, with more pronounced effects at the Level 4 and Level 5 cutoffs.[9] In Figure 1, for example, the share of students scoring above Level 5 in 2007 exceeds the value extrapolated through the continuous line by approximately 3%. One year later, this same quantity is below 1%. The remaining discontinuities are the result of school appeals, which increased substantially after 2008. This result can be seen in appendix Figure C.4, which reports the percentage of exams for which schools appealed (dashed line) and the percentage of successful appeals (continuous line).

This evidence suggests that the abolition of borderlining made schools more liable for correcting errors around thresholds. At the same time, the limited effect of the increased number of appeals on score distributions suggests that borderlining is the prime suspect for discontinuities until 2007. As the performance tables report the share of students attaining at each target, not the average score by subject, there may be a large signal change in perceived school quality caused by such discontinuities.

Parents' Responses to News on School Ratings

The dates when LAs begin accepting applications for the new year are on different days depending on the LA, usually at the start of the autumn term of the year before the child is due to start school. The deadline is in early January and, when applying, parents must provide proof of address. Because buying and selling a property are time-consuming tasks, the house transactions of parents purposefully targeting a school catchment are most likely concluded by the end of the summer. This expectation is borne out by Panel A of Figure 2,

[9] See http://www.standard.co.uk/news/marking-fiddle-has-boosted-sats-results-6918127.html and Department for Children, Schools and Families (2009). The results, which are available upon request, show that the size of the discontinuities in Figure 1 does not differ by school type.

which shows systematic above-trend increases in the time series of transactions during the second and third quarters (the shaded areas of the graph) for all house types.

Panel B of Figure 2 indicates the importance of year-to-year changes in school quality in explaining seasonal fluctuations, suggesting that parents respond to new information on school ratings. The vertical axis reports the difference in house prices paid for transactions in the second and third quarters of two consecutive years in the school district (LA). The horizontal axis considers the one-year difference in the share of students attaining at Level 5 (above expectations) for the same district. Performance tables are published in December using the latest available KS2 results, meaning that information on schools available to parents wishing to enroll their child in September 2018, for example, comes from the December 2016 tables. The figure adjusts for this time lag and also shows the fit from a linear regression that controls for the effects of LA and house type using the working sample described in Section 3, below.[10] Our regression estimates suggest that a one-standard deviation (σ hereafter) increase in the share of children attaining above expectations yields a 2.3 percentage change in prices.

3 Data

Geographic Hierarchies and Sample Selection

We use administrative records from the National Pupil Database (NPD) on primary school students in England (approximately 600k per year). Data include scores and progression (i.e., attainment level awarded) though key stages along with school and student characteristics, such as gender, ethnicity, first language, eligibility for free school meals and special educational needs. The first wave of the NPD was conducted in 2002 by linking national tests to the school census, although scores in English, mathematics and science have been collected since 1998.[11] The availability of students' residence and school postcodes allows for linkages with small-area statistics (e.g., crime rates, social homogeneity, labor market participation

[10] Panel B of Figure 2 is constructed by computing the share of students performing above expectations by district and year. The log-transformed real prices are collapsed to the district-year-house type level. The scatterplot uses one-year differences for both variables, and the linear fit controls for the effects of district and house type.

[11] Science tests were discontinued in 2010.

and land use) produced by the Office for National Statistics (ONS) using the 1991, 2001 and 2011 censuses.

The geography considered has a very fine resolution and consists of areas of compact shapes, fitted within LA boundaries, with a target population of 400 households. Given this size, we will conventionally call these areas "blocks". We use the variability within homogeneous neighborhoods consisting, on average, of an aggregation of 5 adjacent blocks.[12] The right-hand panel of Figure 3 presents an example of the geographic hierarchy for the borough of Tower Hamlets, which is to the East of the City of London and includes the redeveloped Docklands region. This borough is organized into 31 neighborhoods and 130 blocks with a population of 254,100 (listed in the 2011 census). Blocks have an average size of just above 0.05 square miles, and the neighborhoods are akin to squares each with sides 0.5 mile long.

We keep all neighborhoods within metropolitan areas and all urban neighborhoods within non-metropolitan areas of England. Our primary sample consists of all blocks with at least one school belonging to the LA within a 0.4-mile radius of the block's centroid. A similar geographic width was used in other studies (see Machin, 2011, Gibbons et al., 2013 and Burgess et al., 2015), and represents the 50th percentile of the student-school distance distribution in the NPD. Our primary sample consists of 5,009,817 students in 11,484 schools across 23,774 blocks of 6,030 neighborhoods in England (the areas considered are shown in the left-hand panel of Figure 3). In the tables, we check the robustness of our conclusions considering two alternative samples defined by 0.3-mile and 0.6-mile radii centered on block centroids.

Residential Sales and Businesses

We use administrative records from the Land Registry of all residential sales between 1995 and 2015. Each transaction reports the sale price, date of transfer, property type (detached, semi-detached, terraced, flats/maisonettes) and property age (newly-built property or established

[12] Our definition of "neighborhoods" uses Middle Layer Super Output Areas (MSOAs) as defined by the Office for National Statistics. There are 6,781 such neighborhoods in England, aligning to LA boundaries, with population sizes between 5,000 and 15,000 and an average of 3,000 households. What we call "blocks" are instead Lower Layer Super Output Areas (LSOAs), a set of 32,482 narrowly defined areas across England used by the ONS for the computation of small-area statistics. See Appendix A for details.

residential building).[13] Addresses are geocoded (within blocks) and linked to census statistics and council tax records for the area in which each sale took place.

Business data from 1997 to 2015 have been collected by the Office for National Statistics and HM Revenue and Customs and are available in the Business Structure Database (BSD). The businesses listed are obtained from the Inter-Departmental Business Register (IDBR), which captures approximately 99% of the economic activity in the UK. Each business is divided between the "enterprise", which represents the overall business organization, and "local units" (e.g., stores, bank branches). For each business industrial classification, the incorporation and termination year, among other information, are reported. Descriptive statistics for school characteristics, demographic composition and residential sales and businesses across blocks are presented in Table 1 and Table 9.

4 Graphical Analysis

The Effects of Borderlining

We begin with non-parametric plots that quantify the bunching in score densities near achievement cutoffs by pooling data from 1998 (the first available year of data from the NPD) to 2007 (when re-scoring was abolished). Figure 4 is obtained considering scores in the $[-8, 7]$ window centered on the relevant cutoffs. We compute f_{scjt}, the share of students scoring $s \in [-8, 7]$ around cutoff c (Level 3, Level 4 and Level 5) for subject j (English, mathematics and science) in year t (between 1998 and 2007). The residuals from the regressions of f_{scjt} on a full set of subject and time dummies are plotted separately for the three achievement thresholds. The continuous lines are the fitted values generated by local linear regressions (LLRs), and the smoother uses data on one side of the cutoff only with a normal kernel. Consistent with our expectations, the LLR fits show discontinuities around the cutoffs, and the sharpness of the breaks varies with the attainment level. The regressions show a drop in score densities beginning at three points below the cutoffs, which is compensated by bunching above the achievement thresholds for $s \leq 1$. A number of students are clearly moved from below to just above the thresholds, with otherwise smooth score distributions

[13] Land Registry data do not provide information on the size of the property. In an effort to control for the average house size, all regressions below adjust for neighborhood fixed effects.

away from these critical points.

The effect of borderlining is obtained by contrasting f_{scjt} to the value $\hat{f}_{scjt}$ that would have been observed in the absence of notches and bunching around the proficiency cutoffs. Such a counterfactual distribution is retrieved borrowing from the literature on bunching (see Kleven, 2016 for a review). The idea is to fit a flexible polynomial to the observed score distribution while excluding data in a window around the thresholds. Knowing how borderlining is implemented and using the non-parametric evidence from the LLRs, we consider the area ranging from three marks below to two marks above the thresholds (that is, [-3,1]), as excess bunching fades out after this point.[14] The following equation is estimated separately by the attainment threshold (the index c on parameters is omitted for simplicity):

$$f_{scjt} = \alpha(j,t) + \sum_{i=0}^{2} \beta_i s^i + \sum_{i=-3}^{1} \gamma_i 1(s=i) + \varepsilon_{scjt}, \tag{1}$$

where $\alpha(j,t)$ is shorthand for a full set of subject and time effects centered at zero, a second-order polynomial in s is used to approximate the counterfactual densities, and the γ_i values represent the score-specific effects of notches or bunching (below and above the thresholds, respectively). The equation is estimated by imposing that "missing mass" equals "bunching mass", which implies a linear restriction on the estimated γ_i values. The dashed lines in Figure 4 represent the predicted values of $\hat{f}_{scjt}$ implied by the equation above.

The size of this drop is shown in Panel A of Table 2, which also reports the estimates of γ_i by attainment threshold in columns (1) to (5). The value $\sum_{i=-3}^{-1} \gamma_i$ is in column (6) and represents our estimate of the notch induced by bordelining. Consistent with Figure 1, the notch in score densities at Level 5 is much larger and is estimated to be approximately 1.5%, twice that at Level 4. The discontinuity at Level 3 is negligible. Importantly, equation (1) does not detect discontinuities away from the scores that are relevant to borderlining, which can be seen in Panel B of Table 2, where the estimates using a $[-8, 7]$ window centered ten points below the critical thresholds are reported (see Appendix B for additional details). All γ_i values and the missing masses are precise zeros across the three attainment levels.

[14] Our conclusions are robust to the choice of this interval and to the order of the polynomial used in equation (1), below.

The Anatomy of Discontinuities

A closer look at the score distributions before 2008 reveals that the discontinuities are not the result of adjustments to random errors in marking. Random errors are symmetrically distributed and would result in some scores being adjusted downwards. Figure 1 weighs against this hypothesis, showing that the density of marks greater than four points away from the achievement levels is not affected. At the same time, the drop in score distributions becomes more evident near the relevant thresholds, again suggesting that random errors are not likely to be the explanation. A visual inspection of Table 2 reveals this gradient, with larger values for γ_{-1} (in absolute terms) than for γ_{-3}.

What is the origin of score manipulation? The simplest explanation seems to be the most likely: motivated by a "genuine" willingness to help, markers manipulate scores, and the students falling just below an important grade boundary may benefit from having their score manipulated upwards. The fact that score densities are smooth across achievement thresholds after 2008 suggests that manipulation is unrelated to accountability incentives as these were not changed by the abolition of borderlining. Examples of manipulation that is unrelated to NCLB-style accountability pressure and arising from the grading protocol have been found in other contexts (see Angrist et al., 2017 for Italy, Diamond and Persson, 2016 for Sweden, and Dee et al., 2016 for the United States).

The theoretical case for manipulation can be fleshed out using a stylized model of grading behavior for exams reviewed with borderlining. Assume that the utility of external markers is linear in the number of exams upgraded. The cost of upgrading increases with the distance between a student's mark before re-scoring and the threshold. It follows that the upgraded students will have their marks moved to the threshold, implying a spike in the score distribution at that value. Upgrading should be more likely for marks that were originally closer to thresholds, implying more pronounced drops in the distribution near critical values. These are the empirical regularities observed in Figure 1 (indeed, in all years until 2007).

Table 3 shows that manipulation takes place across the board and that exams are not selectively targeted. These findings are consistent with the fact that exams are graded anonymously, implying that markers are not under pressure to improve the standing of schools in the performance tables or the reputation of students, as demonstrated in other institutional contexts. We first collapse data to the score-cutoff-subject-year level, and then estimate equation (1) using the number of students, school and area characteristics at the Level 5 attainment threshold on the left-hand side (the same conclusions hold up for Level 4). We consider gender, ethnicity, eligibility for free school meals, the language ability of students and the type of school from the NPD data. Additionally, we use the characteristics of the area where the student lives as derived from the 1991 census. Estimates of the γ_i values by the attainment threshold are exactly zero, suggesting that the upgraded students are not selectively different in terms of family background and the type of school attended. The statistical analysis in Table 3 mirrors the conclusions from a visual check for discontinuities in appendix Figure C.3.

5 School Quality and Manipulation

Empirical Specification

The quality of schools in block b of neighborhood n, q_{bn}, is proxied by indicators constructed as in the performance tables. We first consider the share of students scoring above expectations (i.e., at or above Level 5), a measure of school excellence. In addition, we derive a summary score calculated by assigning points to each student's results using an equivalence scale provided by the Department for Education.[15] We construct both indicators of school quality by pooling exams from 1998 to 2007 in schools within a fixed distance of a block's centroid, mirroring the procedure discussed in the data section. It follows that q_{bn} represents the average quality of schools around block b in the years before the abolition of borderlining (see Panel C of Table 1 for summary statistics).

A measure of how borderlining affects the perceptions of parents in block b of neighborhood n, z_{bn} is constructed in a similar manner. We start by computing f_{scjt} using all schools within a fixed distance of a block's centroid (the 0.3, 0.4 or 0.6-mile radii described in the data section). We then estimate equation (1) by block and attainment threshold to obtain the values of the missing mass, $\sum_{i=-3}^{-1} \gamma_i$.[16] By iterating over blocks in the sample, this pro-

[15] In the performance tables for 2003, for example, each student is awarded a number of points depending on the achievement level attained (15 points at Level 2; 21 points at Level 3; 27 points at Level 4; and 33 points at Level 5). The average school score is calculated by adding the total points across subjects. A score of 30 would therefore mean that on average, students attained more than Level 4 but less than Level 5.

[16] In the interest of precision, our preferred specification pools scores from 1998 to 2007, collapses data to the score, cutoff and subject level and estimates a version of equation (1) without the index t. The results

cedure yields our proxy for the noise in the measurement of school quality. The quantity z_{bn} is defined as (for simplicity, we avoid indexing coefficients to b and c):

$$\frac{\sum_{i=-3}^{-1} \gamma_i}{\sum_{s=-3}^{-1} \sum_{i=0}^{2} \beta_i s^i}, \tag{2}$$

which represents the percent noise arising from borderlining relative to $r_{bn} \equiv \sum_{s=-3}^{-1} \sum_{i=0}^{2} \beta_i s^i$, the counterfactual share of students falling within three marks of the threshold.

Our statistical analysis relates the outcomes in block b of neighborhood n to percent noise in the share of students attaining Level 5, the average of which is approximately 19% (see Panel D of Table 1). Table 4 documents the relationship between school quality, q_{bn}, and z_{bn} at Level 5 with the aid of regressions that control for unobserved neighborhood effects, as discussed below.[17] A 3σ difference in z_{bn} yields a $0.0935 \times 3 = 0.28\sigma$ change in the share of students attaining at language Level 5, as seen in column (4) of the table, which is equivalent to an additional 3% over the baseline of 26% (see Table 1). The table also implies effects of $0.0478 \times 3 = 0.14\sigma$ for math Level 5 (1.5% over a baseline of 27%) and of $0.0596 \times 3 = 0.18\sigma$ for the average point score (1% of the mean).

Manipulation is as Good as Random

The extent of score manipulation is as good as random across blocks. First, a simple scatterplot of z_{bn} at Level 5 pictures random variability at all values of the counterfactual share r_{bn}. This result can be seen in Panel A of Figure 5, where the correlation coefficient between z_{bn} and r_{bn} is 3% and is not statistically different from zero. Panel B of the figure replicates the analysis considering values of z_{bn} at Level 4, and yields the same conclusions.

Table 5 shows that the following regression:

$$y_{bnt} = \tau_0(t, n) + \tau_1 z_{bn} + \tau_2 x_{bn} + u_{bnt}, \tag{3}$$

yields precise zeros for the coefficients on z_{bn} at Level 5 considering a long array of variables predetermined with respect to bordelining. In the notation employed, $\tau_0(t, n)$ is shorthand for sets of time and neighborhood effects, x_{bn} is a vector of covariates spanning the dimensions used by the ONS for the computation of their index of deprivation, and y_{bnt} is the block-

are robust to this choice. Appendix B offers an in-depth analysis of this generated regressor, showing that the block-specific effects of borderlining are precisely estimated.

[17]The quantity q_{bn} is the outcome in an equation similar to (3), without adjusting for time effects.

averaged outcome at time t (before 1998, the first year used for computing z_{bn}).[18] In addition, equation (3) controls for a quadratic polynomial in r_{bn}, a quadratic polynomial in the distance from the block to the closest school, the number of schools within the radius considered, the population density and a quadratic polynomial for average enrollment in KS2 grades. Standard errors are clustered on LA, and different years are used depending on the outcome considered (as shown in tables).

The effects of borderlining are unrelated to past prices, demographic composition, housing structure or the volume of sales in the residential block. Controlling for neighborhood unobservables in columns (2), (4) and (6) generally lowers the size of the estimated coefficients. Panel A of Table 5 rules out any relationship between z_{bn} and house prices between 1995 and 1998 (the discussion in Section 2 supports the assumption that 1998 prices are pre-determined with respect to z_{bn}). We find that a 3σ increase in z_{bn} in column (4) yields a 0.2% change in prices, a precisely measured and statistically negligible effect.[19] The same panel considers the share of transactions by house type between 1995 and 1998 from the Land Registry data standardized by year. The data on residential sales are representative of the flow of new homeowners into blocks. To overcome this limitation, Panel B shows the results obtained using house characteristics from the 1991 census. An additional array of variables from the 1991 census is considered in Panel C, and the results corroborate the validity of our placebo regressions.

Finally, our calculations show that the percent noise in the share of students attaining Level 5 is not serially correlated within the same LA. This result can be seen in appendix Figure C.5, which is obtained when estimating equation (1) by year and school district (LA) and reporting the scatterplot of the current and lagged values of (2). A formal test for the significance of the auto-correlation coefficient fails to reject the hypothesis.

[18] The variables x_{bn} consist of the following block-level statistics from the 1991 census: percent of unemployed, percent of college dropouts, percent of managers, percent of professionals, percent of whites and percent of blacks. Following the Standard Occupational Classification (SOC2000) in the census, managers include employees with managerial positions in organizations and businesses, and professionals include employees in professional occupations (e.g., science research, teaching).

[19] The left-hand side of equation (3) employs block-year averages of residuals from a regression of log-transformed 1995-2015 prices on time and LA dummies, LA-specific quadratic trends (meant to capture general the deterioration or beautification of school districts) and house characteristics. All regressions involving outcomes from the Land Registry (e.g., prices) are weighted by the number of residential sales used to compute block-year averages. The same definitions are adopted when prices are considered on the left-hand side of equation (4), below.

6 Results

Manipulation Affects Residential Sales

We consider parametric models that exploit discontinuities in the score distributions arising from borderlining. The following equation is estimated:

$$y_{bnt} = \lambda_0(t, n) + \lambda_1 z_{bn} d_t + \lambda_2 x_{bn} + v_{bnt}, \tag{4}$$

where $\lambda_0(t, n)$ is shorthand for sets of time and neighborhood effects, d_t is a dummy for years post 2007 and z_{bn} is the extent of manipulation at Level 5. All regressions control for the same variables in equation (3) and use data for pre- ($t = 1995, \ldots, 1998$) and post-borderlining ($t = 2008, \ldots, 2015$) years.[20] The interaction term (no effect of z_{bn} on the baseline outcomes) is motivated by the placebo tests above. The thought experiment establishes a comparison of blocks in the catchments of schools with the same counterfactual share of students below Level 5 (i.e., the share without inflation) but different ranks in performance tables because of manipulation. The underlying assumption is that outcomes across blocks would have changed similarly over time as a result of the same manipulation (or with no manipulation at all) between 1998 and 2007.

The noise z_{bn} is capitalized in house prices, as shown in Panel A of Table 6. A 3σ increase in the perception of school quality caused by borderlining yields a house price premium between 1% and 1.5%, depending on the sample considered. The coefficients on z_{bn} are precisely estimated, are larger than in the placebo regressions and survive to the inclusion of the control variables x_{bn}. The figure in column (4) suggests that a move from a block with z_{bn} scoring in the tenth percentile of the sample to a block in the ninetieth percentile would result in a house price increase of £2,560 (the monetary equivalent is computed at the mean of £237,004 in 2015 prices).

How does this estimate compare to other studies on housing valuations of school performance? Using the coefficients in column (4), a juxtaposition with the results in Table 4 suggests that a 3σ increase in z_{bn} results in a 0.28σ change in the share of students attaining a Level 5 in language. Our estimates of the valuation of school quality therefore vary

[20] When considering census outcomes, 1991 and 2001 data are used for pre-borderlining years. The results are robust to the omission of 2001 data or to their inclusion with 2011 data in post-borderlining years. The percent of income claimants is not available for the 1991 census.

between $1/0.28 = 3.58\%$ and $1.5/0.28 = 5.36\%$ for one σ increase in student performance. The range depicted is above the 3% premium documented in past research for England (see Machin, 2011, and Gibbons et al., 2013). However the parameter retrieved here represents the willingness of parents to pay for performance above expectations (at Level 5), as opposed to average performance, and has different external validity compared to results from a boundary discontinuity design.

A mild distance gradient emerges from Table 6, with larger effects for blocks further away from schools - columns (5) and (6). This evidence is consistent with parents hedging to maximize the number of schools to choose from at the time of application. The descriptives in Table 1 show that a block in the 0.6-mile sample belongs, on average, to the catchments of three schools, twice as many as in the 0.3-mile sample. The most cost-effective residential blocks are those in the catchments of multiple schools, arguably because of the widened set of possible choices at the time of application. Consistent with this expectation, we find that the capitalization of z_{bn} in house prices increases when blocks belong to a larger number of school catchments. This can be seen from Panel B of Table 6, where the results are obtained by estimating equation (4) from a stratification of blocks that depends on the number of school catchments (above and below the sample average). A 3σ increase in the perception of school quality caused by borderlining yields a premium of about 3% on house prices, or £7,110, when the number of school catchments is above average. The effect is markedly lower in blocks belonging to fewer school catchments, and indeed not statistically significant in the 0.3-mile and 0.4-mile samples.

These effects are not related to any improvement or beautification of existing dwellings. Variation in the signal of school quality does not affect the Council Tax collected by the LA. This tax represents a fixed amount of the property value, which is assessed and classified into bands and can be re-valued if the property is altered or there are significant changes to the local area, such as a new road being built. Panel A of Table 7 shows that z_{bn} has no effect on the share of properties with low-council tax valuations. There are no detectable changes in the share of townhouses (terraced houses), stand-alone properties (detached houses) or residential dwellings with multiple housing units within the same complex (flats). Using administrative records from the Land Registry, Panel B rules out the possibility of differential changes in the number of residential sales by property type. The possibility of new residential developments is also ruled out.

Manipulation Affects Residential Sorting

The use of distance as a tie-breaker to determine school admission suggests that poorer families might be priced out and would be more likely to live far from high-achieving schools. We should therefore expect changes in the socio-economic indicators of the block, which would eventually be reflected in the changes in the composition of students. This expectation is borne out by Table 8, which reports estimates of equation (4) using a number of indicators of school and area composition collapsed to the block-year level on the left-hand side.

Score manipulation causes households with higher income and education to sort into blocks in the catchment of schools that appear to have better performance. The block-level census outcomes in Panel A show that a 3σ increase in z_{bn} yields a lower share of income benefit claimants ($0.015 - 0.029\sigma$ or $1.2 - 2.3\%$ from the sample mean in column (3) of Table 1, depending on the specifications) and a surge in the number of college graduates ($0.029 - 0.035\sigma$ or $1.5 - 1.8\%$). Although blocks do not seem to have expanded in terms of resident populations, changes to the demographic compositions of blocks are triggered by the increased share of households with children of compulsory school age ($0.024 - 0.027\sigma$ or $1.9 - 2.2\%$). Gentrification associated with the displacement of less-economically affluent residents is reflected in different students attending schools in the area, as shown in Panel B of Table 8. The share of students receiving free school meals shrinks by about 0.1σ (approximately 7.3%) as a result of a 3σ increase in z_{bn}. A similar figure holds for the share of black students, who are displaced by those with English as a first language.

Demographic changes mediated by the increased willingness to pay for student performance need not be homogeneous across neighborhoods. Changes may be more pronounced where the school choice is less constrained, a fact that is consistent with the results in Panel B of Table 6. Another source of heterogeneity may be the parental preference for areas that are less economically well off but that show economic improvement and regeneration. In the latter case, score manipulation may affect sorting and house prices more in up-and-coming neighborhoods than in those that have been traditionally home to more affluent residents and good schools. We look into this issue and model the characteristics of one's block using the same demographics considered by the ONS for the computation of their index of multiple deprivation (IMD). The composition of neighbors is proxied using differences in the IMD between 2004 and 2010 (a time range determined by the availability of official publications for this index) to identify homogeneous blocks in terms of urban development and gentrification. Up-and-coming blocks are defined using an indicator for having IMD above the 2004 median and improvements in IMD between 2004 and 2010 above average. We then estimate equation (4) by stratifying blocks on this indicator.

The effects of manipulation on house prices are larger for blocks of up-and-coming neighborhoods, as shown in the last panel of Table 6. The coefficient in column (4) of Panel C implies a (statistically significant) £5,404 increase from a 3σ change in z_{bn} in up-and-coming blocks, and £2,204 in all remaining blocks. It follows that for the same change in signaled school performance, households are willing to pay additional £3,200 to live in an up-and-coming block. Due to the lower precision, however, this difference is never statistically significant across columns of Panel C. Even leaving precision aside, the between block differences documented here are much smaller than those shown in Panel B of the same table. This suggests the higher willingness to pay for accessing a larger set of schools rather than for moving to a neighborhood facing steady economic improvement.

Manipulation Affects Local Businesses

We know that the geography of household expenditures is markedly local. Consumers visit only a few sites among those available, and a large part of expenditure on food and household services occurs near the place of residence (Agarwal et al., 2017). Changes in shopping patterns are a likely explanation for this behavior: increasingly households top up infrequent large purchases, with frequent visits at local stores (Coibion et al., 2017). We also know that the wealth effects caused by a growth in local housing price may impact negatively on the elasticity of demand of residents. Grocery stores respond to this demand change with higher markups, which trigger local price inflation over and above the effects of changes to retail rents, labor costs and gentrification (Stroebel and Vavra, 2016). Through this channel, the expectation of higher profits may change the economic activity of retailers catering to the surrounding households.

Consistent with this idea, we find that the effects of borderlining spill over to the economic activity of local businesses. A 3σ change in z_{bn} creates approximately two additional business, and increases local employment by 27 units. This can be seen from Table 9, which reports estimates of equation (4) using, on the left-hand side, block-year level indicators of the supply of services offered to local households. For example column (3) shows that a 3σ change in signaled school quality yields $0.5930 \times 3 = 1.78$ additional brick-and-mortar businesses (retailers, restaurants and recreational facilities) in the block, approximately 15% from the sample mean in column (1).[21] The corresponding effect on total employment is $8.9339 \times 3 = 27$ employees (22% from the average), as shown in column (6). The table also considers the category of retailers and restaurants catering mainly to families, which we define by excluding licensed restaurants, bars and pubs. We include in this category grocery stores and neighborhood convenience stores for everyday essentials (like marts, bakeries and butcher shops), food services for customers to take away or to have delivered, unlicensed restaurant and coffeehouses, and other household services like dry cleaning, hairdressing and pharmacies. The breakdown by category reveals significant effects across all economic activities considered.

7 Summary and Directions for Further Work

Score manipulation in the Key Stage 2 exams was used to study how households respond to available information on school quality in England. Quasi-experimental variation arising from the marking process unveils a strong preference of parents for attributes and performance of accessible state-funded primary schools. The willingness to relocate to a good district as a result of increased quality purely induced by manipulation yields a £2,204-7,110 premium (at the mean of 2015 prices) on house prices for additional 3 percentage points of students scoring above expectations in English language over a baseline of 26%. Areas with schools that benefit from a boost in test scores experience lower unemployment and new homeowners from higher socio-economic backgrounds, causing substantial sorting across local neighborhoods. The analysis controls for block-level characteristics using census data, and neighborhood unobservables through the multi-layer geographic hierarchy developed by the Office for National Statistics.

Most research on the effects of school quality on residential sorting and house prices relies

[21] As discussed in Section 3, blocks are akin to squares with sides approximately 0.25-mile long. Marketing research using transactions from credit cards shows that the distance travelled to store is within this limit for 53 percent of consumers in the United Kingdom (see Association of Convenience Stores, 2017).

on boundary discontinuities (see Black and Machin, 2011). In contrast, our strategy uncovers a new identification strategy with improved external validity and has the potential for replication in different institutional settings in the United States and Europe. The estimates suggest that there is a 5% house price premium resulting from a one-standard deviation increase in the share of students attaining above expectations in English, a valuation of school quality that is larger than reported in previous studies for England.

Our findings raise a number of additional questions, including why manipulation that is not motivated by accountability concerns is so prevalent and what can be done to enhance accurate assessments in England and elsewhere. One important policy question is what happens to children of households that are priced out of good areas, a concern motivated by the literature on the long-term effects of exposure to a better neighborhood at a young age (Chetty et al., 2016). It is also worth asking whether score manipulation has long-term effects on student outcomes in England, a question that was already considered for other countries (see, for example, Diamond and Persson, 2016, and Lavy and Megalokonomou, 2017). We hope to address some of these questions in future work.

Table 1: Descriptive statistics

	0.3 mile		0.4 mile		0.6 mile	
	Mean	S.D.	Mean	S.D.	Mean	S.D.
	(1)	(2)	(3)	(4)	(5)	(6)
Panel A. House Characteristics (Land Registry)						
Average prices (2011-15)	236110.50	231213.70	237003.90	224437.60	240755.50	224756.20
Percent detached (2008-15)	0.1434	0.1773	0.1579	0.1857	0.1724	0.1950
Percent semi-detached (2008-15)	0.3009	0.2219	0.3088	0.2190	0.3090	0.2159
Percent flats (2008-15)	0.1907	0.2494	0.1836	0.2407	0.1813	0.2372
Panel B. Area Characteristics (Census)						
Density of households with children aged 5-16 (1991)	4.28	3.45	4.01	3.28	3.81	3.20
Density of households with children aged 5-16 (2011)	3.78	3.03	3.57	2.88	3.40	2.81
Percent of individuals with college degree (1991)	0.0670	0.0698	0.0685	0.0700	0.0704	0.0705
Percent of individuals with college degree (2011)	0.2531	0.1301	0.2559	0.1285	0.2598	0.1280
Percent of lower class households (1991)	0.2001	0.1093	0.1943	0.1088	0.1893	0.1084
Percent of households with no cars (1991)	0.3652	0.1751	0.3509	0.1725	0.3396	0.1718
Percent of households not home-owners (1991)	0.3541	0.2326	0.3382	0.2281	0.3269	0.2254
Percent of households in overcrowded dwellings (1991)	0.0261	0.0278	0.0244	0.0262	0.0233	0.0253
Population density (2011)	54.40	45.49	51.00	42.95	48.53	41.73
Percent income claimants (2011)	0.0349	0.0264	0.0332	0.0259	0.0319	0.0255
Panel C. School Characteristics (NPD)						
Percent attaining at math L5 (1998-2007)	0.2640	0.1080	0.2682	0.1027	0.2721	0.0945
Percent attaining at English L5 (1998-2007)	0.2595	0.1083	0.2642	0.1032	0.2690	0.0956
Percent attaining at science L5 (1998-2007)	0.3580	0.1313	0.3634	0.1246	0.3685	0.1135
Average Point Score (1998-2007)	27.32	1.33	27.36	1.26	27.40	1.15
Percent students on free school meals (2008-2011)	0.2104	0.1558	0.2009	0.1470	0.1917	0.1349
Percent students with English as first language (2008-2011)	0.7041	0.2884	0.7131	0.2812	0.7187	0.2720
Percent black students (2008-2011)	0.0638	0.1194	0.0582	0.1086	0.0548	0.1017
KS2 enrolment (2002-2007)	28.99	8.41	29.41	8.15	29.60	7.70
Number of schools around block	1.45	0.74	1.95	1.18	2.98	1.96
Panel D. Incidence of manipulation (NPD)						
Percent Noise L5	0.1858	0.0845	0.1901	0.0755	0.1950	0.0640
Percent Noise L4	0.1614	0.1122	0.1661	0.1023	0.1725	0.0889
Percent Noise L3	0.0506	0.3225	0.0617	0.4305	0.0714	0.2990
Counterfactual share of students below L5	0.0755	0.0141	0.0749	0.0132	0.0742	0.0118
Counterfactual share of students below L4	0.0460	0.0135	0.0445	0.0124	0.0427	0.0110
Counterfactual share of students below L3	0.0159	0.0096	0.0145	0.0084	0.0129	0.0073
Number of blocks	17,923		23,774		27,250	

Note. This table reports means and standard deviations across blocks (LSOAs) in urban areas of England. Only blocks with at least one school within a 0.3-mile radius are retained in columns (1) to (2). Alternative samples with at least one school within a 0.4-mile and 0.6-mile radius are considered in columns (3) to (4) and (5) to (6), respectively. Numbers in brackets next to variable names represent the period over which means and standard deviations are computed, depending on data availability.

Table 2: Effects of borderlining on scores

	deviations from thresholds:					missing mass
	-3	-2	-1	0	1	
	(1)	(2)	(3)	(4)	(5)	(6)
			Panel A. Real thresholds			
Level 3	-0.0001	-0.0002	-0.0007***	0.0008***	0.0002	0.0010***
	(0.0002)	(0.0002)	(0.0002)	(0.0002)	(0.0001)	(0.0002)
Level 4	-0.0018***	-0.0019***	-0.0037***	0.0062***	0.0011*	0.0074***
	(0.0005)	(0.0005)	(0.0006)	(0.0007)	(0.0006)	(0.0008)
Level 5	-0.0046***	-0.0036***	-0.0066***	0.0136***	0.0012***	0.0148***
	(0.0006)	(0.0005)	(0.0006)	(0.0010)	(0.0004)	(0.0009)
			Panel B. Placebo thresholds			
Level 3	0.0000	-0.0000	0.0000	-0.0000	0.0000	0.0000
	(0.0001)	(0.0001)	(0.0001)	(0.0001)	(0.0001)	(0.0001)
Level 4	0.0000	-0.0000	-0.0000	-0.0000	0.0000	0.0000
	(0.0002)	(0.0001)	(0.0001)	(0.0001)	(0.0002)	(0.0002)
Level 5	-0.0001	-0.0001	-0.0000	0.0000	0.0002	0.0002
	(0.0004)	(0.0003)	(0.0003)	(0.0002)	(0.0002)	(0.0003)

Note. Differences between observed and counterfactual score densities from equation (1). Scores range from -3, three points below threshold in column (1), to 1, one point above threshold in column (5). Column (6) reports estimates of the notch induced by bordelining. Panel A is obtained using thresholds (at Level 3, Level 4 and Level 5) used by external markers. Panel B considers placebo thresholds centered ten points below critical scores. Standard errors are shown in brackets. * p<0.10, ** p<0.05, *** p<0.01.

Table 3: Relationship of borderlining with neighborhood and student characteristics

	deviations from Level 5 threshold:					missing mass
	-3	-2	-1	0	1	
	(1)	(2)	(3)	(4)	(5)	(6)
	Panel A. Student characteristics (NPD)					
Male	-0.0008	-0.0001	-0.0004	0.0006	0.0008	0.0013
	(0.0022)	(0.0018)	(0.0016)	(0.0014)	(0.0016)	(0.0021)
White	0.0005	0.0004	-0.0000	-0.0009*	-0.0001	-0.0009
	(0.0005)	(0.0005)	(0.0005)	(0.0005)	(0.0006)	(0.0007)
On free school meals	-0.0000	-0.0003	0.0008	0.0003	-0.0008	-0.0005
	(0.0011)	(0.0007)	(0.0007)	(0.0006)	(0.0008)	(0.0010)
English speaking	0.0004	-0.0001	0.0002	-0.0005	0.0001	-0.0004
	(0.0003)	(0.0005)	(0.0004)	(0.0003)	(0.0005)	(0.0005)
Attending independent schools	0.0003	-0.0013	0.0013**	0.0000	-0.0004	-0.0003
	(0.0007)	(0.0011)	(0.0006)	(0.0005)	(0.0007)	(0.0009)
	Panel B. Area characteristics (1991 Census)					
Percent of households with children aged 5-16	0.0000	0.0000	-0.0001	0.0000	0.0000	0.0001
	(0.0001)	(0.0001)	(0.0001)	(0.0001)	(0.0001)	(0.0001)
Percent of individuals with college degree	0.0000	-0.0001	0.0001	-0.0000	-0.0001	-0.0001
	(0.0002)	(0.0002)	(0.0002)	(0.0001)	(0.0001)	(0.0002)
Percent of lower class households	-0.0004	0.0001	0.0001	0.0002	-0.0000	0.0002
	(0.0003)	(0.0003)	(0.0003)	(0.0002)	(0.0002)	(0.0003)
Percent of households with no cars	-0.0003	0.0003	-0.0000	0.0001	-0.0000	0.0001
	(0.0006)	(0.0005)	(0.0004)	(0.0003)	(0.0004)	(0.0005)
Percent of households not home-owners	-0.0002	-0.0000	0.0004	-0.0001	-0.0000	-0.0001
	(0.0008)	(0.0005)	(0.0006)	(0.0004)	(0.0005)	(0.0007)
Percent of households in overcrowded dwellings	-0.0000	0.0000	-0.0000	0.0000	0.0000	0.0000
	(0.0001)	(0.0001)	(0.0000)	(0.0000)	(0.0000)	(0.0001)

Note. Differences between observed and counterfactual densities from equation (1) using student and area characteristics (see Section 4 for details) at Level 5 thresholds. Scores range from -3, three points below threshold in column (1), to 1, one point above threshold in column (5). Column (6) reports estimates of the notch induced by bordelining. Panel A considers student characteristics from NPD data; Panel B considers area characteristics from the 1991 census. Standard errors are shown in brackets. * p<0.10, ** p<0.05, *** p<0.01.

Table 4: Effects on measured school quality

	0.3 miles		0.4 miles		0.6 miles	
	(1)	(2)	(3)	(4)	(5)	(6)
Average Point Score	0.0560***	0.0545***	0.0612***	0.0596***	0.0577***	0.0558***
	(0.0082)	(0.0080)	(0.0062)	(0.0062)	(0.0077)	(0.0076)
Percent Level 5 math	0.0380***	0.0364***	0.0498***	0.0478***	0.0514***	0.0491***
	(0.0107)	(0.0105)	(0.0084)	(0.0083)	(0.0089)	(0.0087)
Percent Level 5 language	0.0858***	0.0836***	0.0957***	0.0935***	0.0897***	0.0872***
	(0.0101)	(0.0099)	(0.0084)	(0.0084)	(0.0092)	(0.0091)
Number of blocks	17,923	17,923	23,774	23,774	27,250	27,250
Neighborhood FE	Yes	Yes	Yes	Yes	Yes	Yes
1991 controls		Yes		Yes		Yes

Note. The table shows the effects of score manipulation on indicators of school quality. All variables are standardized to have zero mean and unit variance. Coefficients show the effect of a one-standard deviation change in score manipulation at Level 5. All columns control for a quadratic polynomial in the running variable, a quadratic polynomial in distance from the closest school, number of schools around the block, population density, a quadratic polynomial in average school enrollment, and neighborhood (MSOA) fixed effects. Columns (2), (4) and (6) add 1991 Census controls (percent unemployed, college dropouts, managers, professionals, white and black). Standard errors, shown in brackets, are clustered on Local Authority. *** p<0.01, ** p<0.05, * p<0.1.

Table 5: Placebo regressions

	0.3 miles		0.4 miles		0.6 miles	
	(1)	(2)	(3)	(4)	(5)	(6)
	Panel A. Prices and House Characteristics (1995-98 Land Registry)					
Log prices	0.0003	0.0009	-0.0011	0.0008	-0.0017	0.0023
	(0.0026)	(0.0019)	(0.0023)	(0.0014)	(0.0024)	(0.0017)
Percent detached or semi-detached	0.0029	-0.0008	0.0105	0.0011	0.0032	-0.0116*
	(0.0080)	(0.0080)	(0.0081)	(0.0071)	(0.0079)	(0.0067)
Percent terraced	-0.0175*	0.0024	-0.0157*	0.0020	-0.0129	0.0059
	(0.0095)	(0.0088)	(0.0090)	(0.0075)	(0.0082)	(0.0063)
Percent flats	0.0175**	-0.0018	0.0047	-0.0040	0.0115	0.0089
	(0.0082)	(0.0086)	(0.0097)	(0.0078)	(0.0099)	(0.0094)
Observations	71,388	71,388	94,765	94,765	108,656	108,656
	Panel B. House Characteristics (1991 Census)					
Percent detached or semi-detached	0.0079	0.0021	0.0097	0.0007	0.0066	-0.0083
	(0.0070)	(0.0070)	(0.0074)	(0.0066)	(0.0071)	(0.0063)
Percent terraced	-0.0093	-0.0033	-0.0090	-0.0018	-0.0107	0.0039
	(0.0091)	(0.0097)	(0.0089)	(0.0079)	(0.0084)	(0.0071)
Percent flats	-0.0007	0.0006	-0.0036	0.0009	0.0024	0.0072
	(0.0069)	(0.0078)	(0.0072)	(0.0067)	(0.0075)	(0.0075)
Observations	17,923	17,923	23,774	23,774	27,250	27,250
	Panel C. Area Characteristics (1991 Census)					
Density of households with children aged 5-16	0.0052	0.0061	0.0064*	0.0045	0.0024	0.0002
	(0.0045)	(0.0047)	(0.0038)	(0.0039)	(0.0035)	(0.0035)
Percent of individuals with college degree	0.0039	0.0070*	0.0012	0.0034	0.0010	0.0049
	(0.0030)	(0.0037)	(0.0028)	(0.0032)	(0.0030)	(0.0030)
Percent of lower class households	-0.0070	0.0053	-0.0092*	0.0086	-0.0134**	-0.0020
	(0.0053)	(0.0080)	(0.0051)	(0.0072)	(0.0056)	(0.0078)
Percent of households with no cars	-0.0045	-0.0018	-0.0073	-0.0034	-0.0048	0.0028
	(0.0054)	(0.0058)	(0.0047)	(0.0048)	(0.0050)	(0.0049)
Percent of households not home-owners	-0.0055	-0.0044	-0.0049	-0.0077	-0.0005	-0.0025
	(0.0061)	(0.0063)	(0.0058)	(0.0053)	(0.0061)	(0.0052)
Percent of households in overcrowded dwellings	0.0025	0.0015	0.0036	0.0066	0.0028	0.0056
	(0.0043)	(0.0049)	(0.0042)	(0.0044)	(0.0048)	(0.0038)
Observations	17,923	17,923	23,774	23,774	27,250	27,250
Neighborhood FE		Yes		Yes		Yes

Note. The table shows placebo regressions for price and house characteristics from 1995-98 Land Registry data (Panel A) and for house characteristics (Panel B) and neighborhood characteristics (Panel C) from the 1991 Census. With the exception of house prices (in logs), all other variables are standardized to have zero mean and unit variance. Coefficients show the effect for one standard deviation in score manipulation at Level 5. All columns control for a quadratic polynomial in the running variable, a quadratic polynomial in distance from the closest school, number of schools around the block, population density, a quadratic polynomial in average school enrollment, and 1991 controls (percent unemployed, college dropouts, managers, professionals, white and black). Columns (2), (4) and (6) add neighborhood (MSOA) fixed effects. All regressions for house characteristics (Panel A) are weighted by the number of residential sales in the area and include year fixed effects. Standard errors, shown in brackets, are clustered on Local Authority. *** p<0.01, ** p<0.05, * p<0.1.

Table 6: Effects on house prices

	0.3 miles		0.4 miles		0.6 miles	
	(1)	(2)	(3)	(4)	(5)	(6)
			Panel A. Full sample			
Percent noise at Level 5	0.0039**	0.0031**	0.0050***	0.0045***	0.0063***	0.0055***
	(0.0016)	(0.0015)	(0.0015)	(0.0015)	(0.0017)	(0.0016)
Observations	160,874	160,874	213,476	213,476	244,743	244,743
			Panel B. By number of schools			
Number of school catchments above average	0.0108***	0.0099***	0.0097***	0.0096***	0.0105***	0.0099***
	(0.0029)	(0.0030)	(0.0028)	(0.0028)	(0.0036)	(0.0036)
Observations	53,315	53,315	115,310	115,310	125,234	125,234
Number of school catchments below average	0.0025	0.0019	0.0024	0.0019	0.0042**	0.0037**
	(0.0016)	(0.0015)	(0.0015)	(0.0015)	(0.0017)	(0.0016)
Observations	107,559	107,559	98,166	98,166	119,509	119,509
			Panel C. By deprivation			
Improvement above average	0.0058	0.0053	0.0077**	0.0076**	0.0080**	0.0079**
	(0.0036)	(0.0033)	(0.0037)	(0.0036)	(0.0037)	(0.0038)
Observations	44,391	44,391	56,039	56,039	61,723	61,723
Improvement below average	0.0027*	0.0022	0.0034**	0.0031**	0.0043**	0.0038**
	(0.0016)	(0.0016)	(0.0014)	(0.0014)	(0.0017)	(0.0017)
Observations	112,560	112,560	152,276	152,276	177,088	177,088
Year FE	Yes	Yes	Yes	Yes	Yes	Yes
Neighborhood FE	Yes	Yes	Yes	Yes	Yes	Yes
1991 controls		Yes		Yes		Yes

Note. The table shows the effects of score manipulation on 2011-15 house prices (in logs). The time series of England's housing market in Figure 2 motivates the breakdown by crisis (2008-2010) and post-crisis (2011-2015) periods. The effects on house prices are precisely estimated at zero for 2008-2010 and are available on request. The coefficients here show the effect for a one-standard deviation change in score manipulation at Level 5. Panel A presents results for the full sample. Panel B distinguishes between blocks belonging to a number of school catchments above or below the sample average. Panel C distinguishes between blocks depending on changes to the index of multiple deprivation (IMD) between 2004 and 2010 (see Section 6 for details). Results in Panel B and Panel C follows from regressions stratified by number of school catchments and IMD growth, respectively. All columns control for a quadratic polynomial in the running variable, a quadratic polynomial in distance from the closest school, number of schools around the block, population density, a quadratic polynomial in average school enrollment, and neighborhood (MSOA) fixed effects. Columns (2), (4) and (6) add 1991 Census controls (percent unemployed, college dropouts, managers, professionals, white and black). Standard errors, shown in brackets, are clustered on Local Authority. *** p<0.01, ** p<0.05, * p<0.1.

Table 7: Effects on sales and house characteristics

	0.3 miles		0.4 miles		0.6 miles	
	(1)	(2)	(3)	(4)	(5)	(6)
	Panel A. Census and Neighbourhood Statistics					
Percent detached or semi-detached	-0.0065	-0.0084**	-0.0056	-0.0079**	-0.0105***	-0.0129***
	(0.0043)	(0.0041)	(0.0036)	(0.0036)	(0.0032)	(0.0032)
Percent terraced	0.0008	0.0018	0.0006	0.0028	0.0008	0.0034
	(0.0052)	(0.0051)	(0.0043)	(0.0044)	(0.0037)	(0.0036)
Percent flats	0.0022	0.0037	0.0010	0.0019	0.0055	0.0062
	(0.0043)	(0.0040)	(0.0038)	(0.0039)	(0.0038)	(0.0039)
Density of rooms	0.0019	0.0016	0.0019	0.0017	0.0023	0.0018
	(0.0019)	(0.0019)	(0.0015)	(0.0015)	(0.0014)	(0.0014)
Percent with low council tax valuation˜	-0.0036	0.0001	-0.0036	0.0005	-0.0059*	-0.0002
	(0.0044)	(0.0038)	(0.0033)	(0.0029)	(0.0035)	(0.0028)
Percent of council homes	-0.0090*	-0.0037	-0.0100***	-0.0045	-0.0054	0.0005
	(0.0050)	(0.0039)	(0.0037)	(0.0033)	(0.0036)	(0.0032)
Observations	53,327	53,327	70,741	70,741	81,083	81,083
	Panel B. Land Registry					
Percent detached or semi-detached	-0.0007	-0.0016	-0.0004	-0.0011	-0.0016	-0.0023*
	(0.0017)	(0.0017)	(0.0014)	(0.0015)	(0.0014)	(0.0013)
Percent terraced	-0.0003	0.0003	-0.0015	-0.0008	-0.0008	-0.0000
	(0.0014)	(0.0014)	(0.0012)	(0.0013)	(0.0012)	(0.0011)
Percent flats	0.0010	0.0013	0.0018	0.0019*	0.0024*	0.0024*
	(0.0013)	(0.0013)	(0.0012)	(0.0011)	(0.0014)	(0.0014)
Percent newly built	0.0016	0.0019	0.0007	0.0008	0.0017	0.0018
	(0.0018)	(0.0018)	(0.0016)	(0.0016)	(0.0015)	(0.0015)
Observations	160,874	160,874	213,476	213,476	244,743	244,743
Year FE	Yes	Yes	Yes	Yes	Yes	Yes
Neighborhood FE	Yes	Yes	Yes	Yes	Yes	Yes
1991 controls		Yes		Yes		Yes

Note. The table shows the effects of score manipulation on house characteristics from the 2011 Census (Panel B) and 2011-15 Land Registry (Panel A) data. All variables are standardized to have zero mean and unit variance. The coefficients show the effect for a one-standard deviation change in score manipulation at Level 5. All columns control for a quadratic polynomial in the running variable, a quadratic polynomial in distance from the closest school, number of schools around the block, population density, a quadratic polynomial in average school enrollment, and neighborhood (MSOA) fixed effects. Columns (2), (4) and (6) add 1991 Census controls (percent unemployed, college dropouts, managers, professionals, white and black). Standard errors, shown in brackets, are clustered on Local Authority. *** p<0.01, ** p<0.05, * p<0.1. ˜ No data for 1991.

Table 8: Effects on neighborhood and school characteristics

	0.3 miles		0.4 miles		0.6 miles	
	(1)	(2)	(3)	(4)	(5)	(6)
	Panel A. Census and Neighbourhood Statistics (2011)					
Population density	0.0008	0.0009	0.0026	0.0027	0.0015	0.0017
	(0.0020)	(0.0020)	(0.0018)	(0.0018)	(0.0018)	(0.0018)
Density of households with children aged 5-16	0.0047	0.0056	0.0076***	0.0080***	0.0082***	0.0091***
	(0.0034)	(0.0034)	(0.0028)	(0.0028)	(0.0026)	(0.0026)
Percent with college degree	0.0166***	0.0115***	0.0157***	0.0097***	0.0174***	0.0101***
	(0.0034)	(0.0029)	(0.0033)	(0.0031)	(0.0035)	(0.0032)
Percent income claimants˜	-0.0122**	-0.0071	-0.0153***	-0.0098**	-0.0108***	-0.0050
	(0.0058)	(0.0044)	(0.0042)	(0.0038)	(0.0037)	(0.0035)
Observations	53,327	53,327	70,741	70,741	81,083	81,083
	Panel B. School Composition (2008-11)					
Percent on free school meals	-0.0241***	-0.0229***	-0.0343***	-0.0333***	-0.0367***	-0.0358***
	(0.0090)	(0.0086)	(0.0090)	(0.0089)	(0.0081)	(0.0081)
Percent with English as first langauge	0.0215*	0.0218**	0.0232**	0.0235**	0.0321**	0.0319**
	(0.0110)	(0.0110)	(0.0107)	(0.0107)	(0.0132)	(0.0133)
Percent black	-0.0231***	-0.0227***	-0.0227***	-0.0230***	-0.0237***	-0.0237***
	(0.0084)	(0.0085)	(0.0077)	(0.0078)	(0.0080)	(0.0080)
Observations	32,943	32,943	43,471	43,471	50,979	50,979
Year FE	Yes	Yes	Yes	Yes	Yes	Yes
Neighborhood FE	Yes	Yes	Yes	Yes	Yes	Yes
1991 Controls		Yes		Yes		Yes

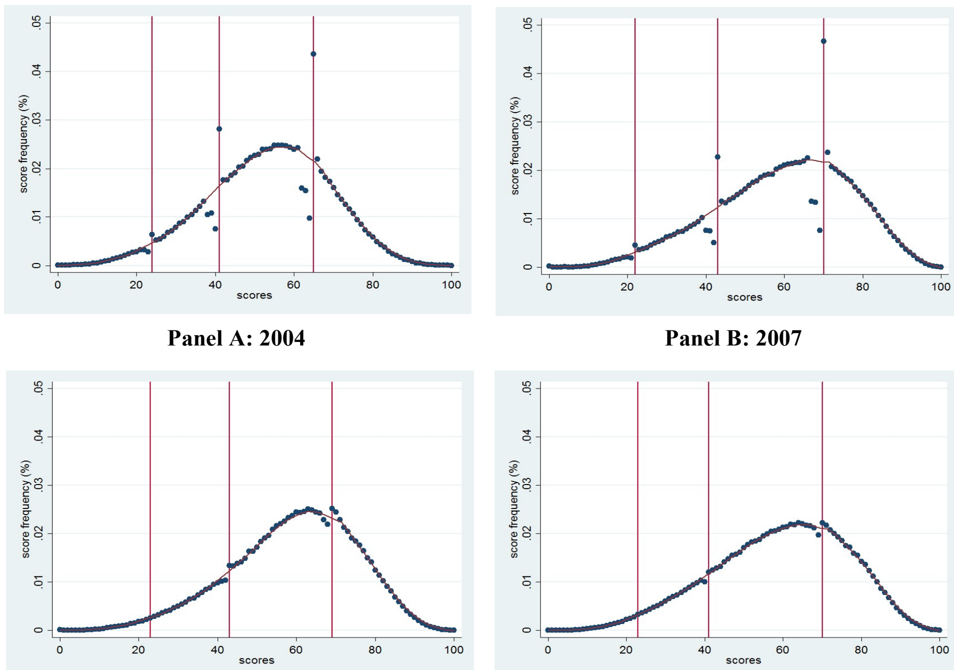
Note. The table shows the effects of score manipulation on neighborhood (Panel A) and school (Panel B) composition for outcomes in the 2011 Census data, 2011 Neighbourhood Statistics from the Office for National Statistics, and 2008-11 National Pupil Database records. All variables are standardized to have zero mean and unit variance. Coefficients show the effect for a one-standard deviation change in score manipulation at Level 5. All columns control for a quadratic polynomial in the running variable, a quadratic polynomial in distance from the closest school, number of schools around the block, population density, a quadratic polynomial in average school enrollment, and neighborhood (MSOA) fixed effects. Columns (2), (4) and (6) add 1991 Census controls (percent unemployed, college dropouts, managers, professionals, white and black). Standard errors, shown in brackets, are clustered on Local Authority. *** p<0.01, ** p<0.05, * p<0.1. ˜ No data for 1991.

Table 9: Effects on local businesses

	Number of businesses			Employment		
	Block Average	Effects		Block Average	Effects	
	(1)	(2)	(3)	(4)	(5)	(6)
All businesses	11.8385	0.5654***	0.5930***	120.0032	8.4671**	8.9339**
		(0.2024)	(0.2030)		(3.3862)	(3.4490)
Family friendly services	9.6810	0.4414***	0.4629***	97.9158	6.6410**	7.0211**
		(0.1593)	(0.1594)		(2.7125)	(2.7564)
Grocery stores	2.7007	0.0722**	0.0788**	41.6645	1.6988*	1.8329*
		(0.0312)	(0.0311)		(0.9420)	(0.9497)
Bars and public houses	1.3227	0.0588***	0.0620***	12.3951	0.8444**	0.8802***
		(0.0221)	(0.0222)		(0.3264)	(0.3309)
Licensed food and beverage services	0.8348	0.0652**	0.0681**	9.6923	0.9817*	1.0325*
		(0.0287)	(0.0291)		(0.5321)	(0.5456)
Unlicensed food and beverage services	1.5432	0.0537**	0.0578**	10.8592	0.4237	0.4618
		(0.0255)	(0.0257)		(0.3230)	(0.3263)
Other household services	5.4370	0.3155***	0.3262***	45.3921	4.5184**	4.7265**
		(0.1082)	(0.1082)		(1.7845)	(1.8161)
Observations		269,468	269,468		269,468	269,468
Year FE		Yes	Yes		Yes	Yes
Neighborhood FE		Yes	Yes		Yes	Yes
1991 Controls			Yes			Yes

Note. The table shows the effects of score manipulation on number of, and total employment in, brick-and-mortar businesses (retailers, restaurants and recreational facilities). Coefficients here show the effect for a one-standard deviation in score manipulation at Level 5 using the 0.4-mile radius sample. Grocery stores consist of small retailers (e.g. bakeries and butchers), convenience stores (including those in a chain) and large supermarkets. Unlicensed food and beverage services consist of take-away shops, restaurants and coffeehouses (including those in a chain). Other household services include dry cleaning, hairdressing, travel agencies, medical and dental practices, pharmacies, veterinaries and clothes shops (the full list is available on request). The category of retailers and restaurants catering to families is defined as (a) grocery stores, (b) unlicensed food and beverage services, and (c) other household services. Columns (1) to (3) consider the number of businesses, while workplace employment is in columns (4) to (6). Columns (1) and (4) show block averages for 2008-2015. All columns control for a quadratic polynomial in the running variable, a quadratic polynomial in distance from the closest school, number of schools around the block, population density, a quadratic polynomial in average school enrolment, year and neighborhood (MSOA) fixed effects. Columns (3) and (6) add 1991 Census controls (percent unemployed, college dropouts, managers, professionals, white and black). Standard errors, shown in brackets, are clustered on Local Authority. *** p<0.01, ** p<0.05, * p<0.1.

Figure 1: Score manipulation over time: English



Panel A: 2004

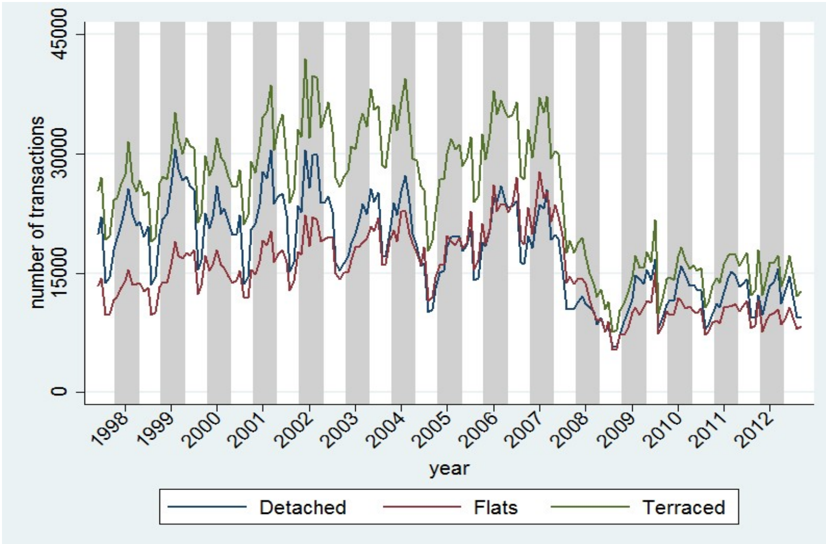
Panel B: 2007

Panel C: 2008

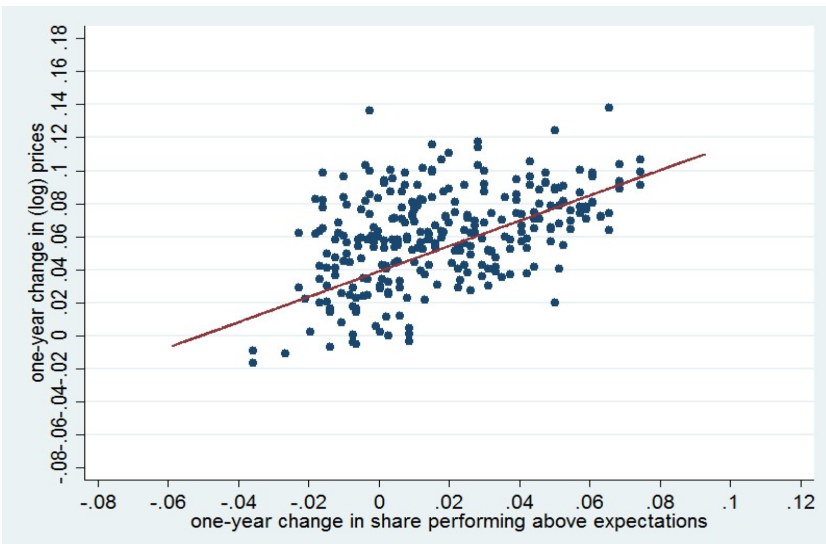
Panel D: 2011

Note. This figure shows Key Stage 2 score distributions for English in selected years before (2004 and 2007) and after (2008 and 2011) the removal of borderlining. In each panel, the vertical lines are critical attainment thresholds set in that year. The continuous line is a local linear regression fit obtained excluding observations in the [-3,2] windows around thresholds.

Figure 2: Price and school quality changes



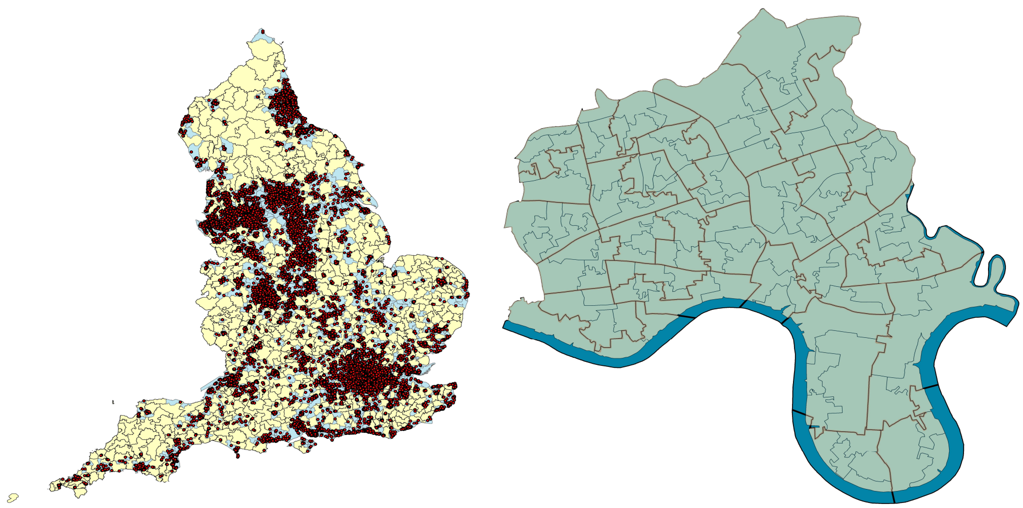
Panel A. Seasonality in House Prices



Panel B. Changes in Price and School Quality in the Short Term

Note. Panel A shows the number of residential property transactions over time. Shaded areas mark the second and third trimester of every year. Panel B plots one-year changes in (log) prices against one-year changes in the share of students performing above expectations, with a linear fit superimposed.

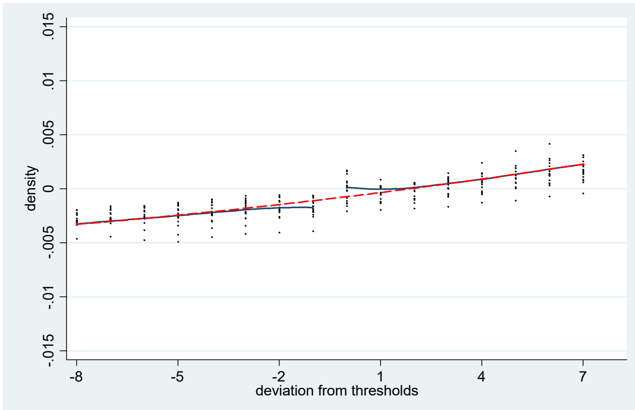
Figure 3: Geography and sample selection



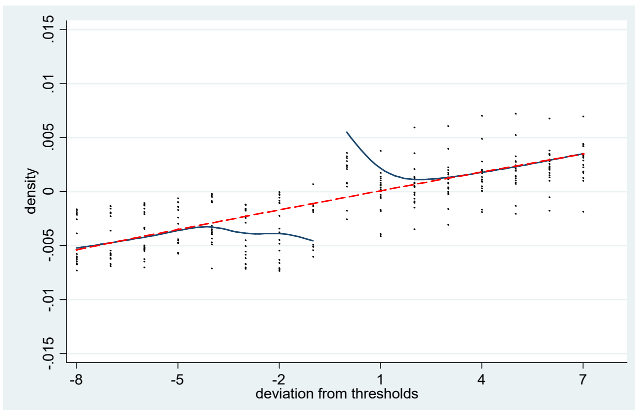
Panel A **Panel B**

Note. Panel A shows urban MSOAs in England (light blue areas) with the corresponding listing of schools (red dots) superimposed. Panel B shows the geographic hierarchy defined by LSOAs and MSOAs for the London borough of Tower Hamlets, comprising 31 MSOAs (brown outline) and 130 LSOAs (grey outline). In the text, LSOAs correspond to blocks and MSOAs correspond to neighborhoods.

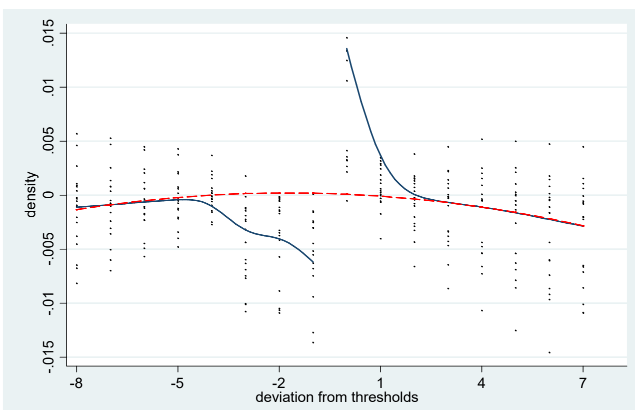
Figure 4: Bunching around achievement thresholds



Panel A. Level 3



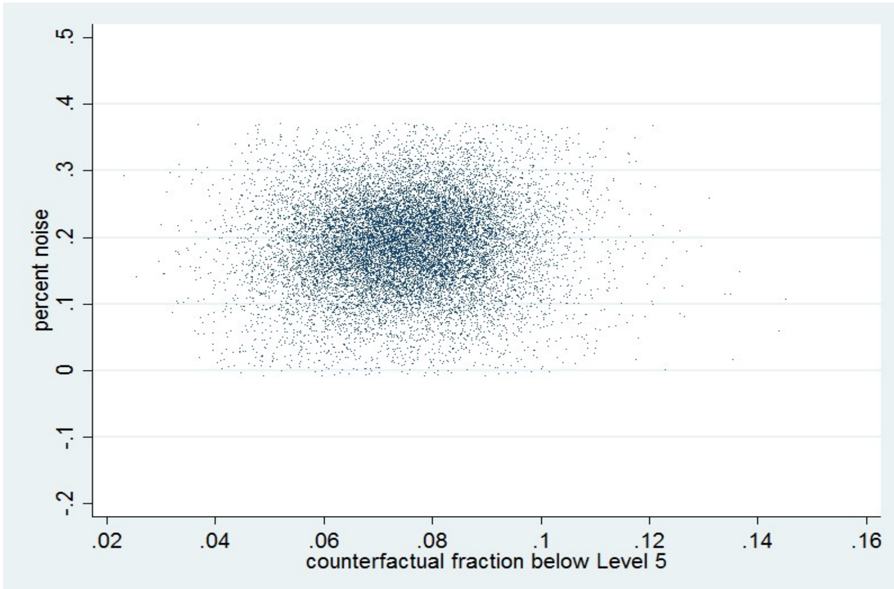
Panel B. Level 4



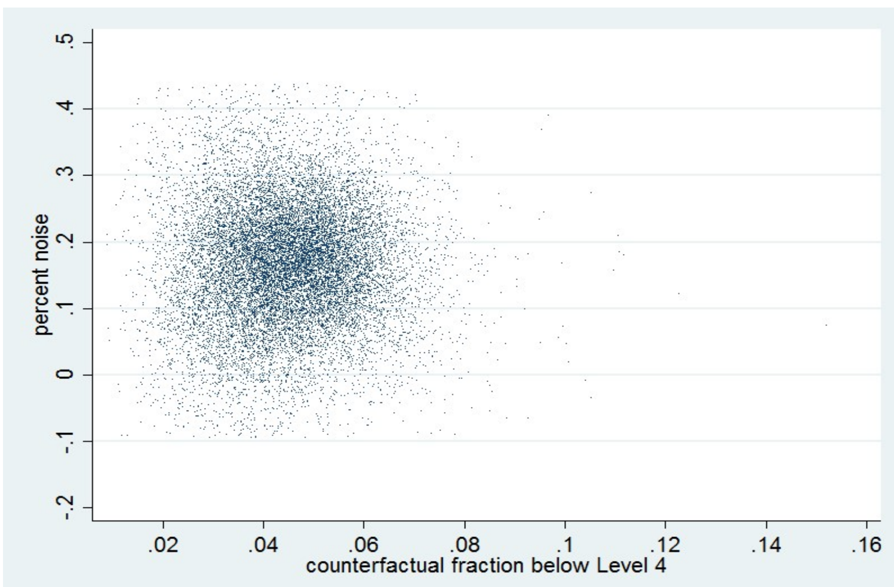
Panel C. Level 5

Note. This figure plots residuals of test score frequencies around Level 3 threshold (Panel A), Level 4 threshold (Panel B) and Level 5 threshold (Panel C). Continuous lines are fitted values obtained with local linear regressions (LLRs). The window around the cutoffs is chosen so that consecutive windows do not overlap. Dashed lines represent the test score distribution that would have been observed in the absence of bunching around achievement thresholds. See text for more details.

Figure 5: Variability in percent noise



Panel A. Level 5



Panel B. Level 4

Note. This figure shows how the percent noise in school quality (vertical axis) varies with the share of students below the critical threshold (horizontal axis). Figures are drawn considering the 0.4-mile radius sample. Panel A refers to Level 5, while Panel B considers Level 4.

References

Agarwal, S., Jensen, J. B., and Monte, F. (2017). The geography of consumption. *NBER Working Paper No. 23616.*

Angrist, J. D., Battistin, E., and Vuri, D. (2017). In a small moment: Class size and moral hazard in the mezzogiorno. *American Economic Journal: Applied Economics*, Forthcoming.

Battistin, E. (2016). How manipulating test scores affects school accountability and student achievement. *IZA World of Labor*, 295.

Bayer, P., Ferreira, F., and McMillan, R. (2007). A unified framework for measuring preferences for schools and neighborhoods. *Journal of Political Economy*, 115(4):588–638.

Bhattacharji, P. and Kingdon, G. (2017). The great Indian exam debacle. *The Observer Research Foundation Brief*, 191.

Black, S. E. (1999). Do better schools matter? Parental valuation of elementary education. *The Quarterly Journal of Economics*, 114(2):577–599.

Black, S. E. and Machin, S. (2011). Housing valuations of school performance. *Handbook of the Economics of Education*, 3:485–519.

Burgess, S. and Greaves, E. (2013). Test scores, subjective assessment, and stereotyping of ethnic minorities. *Journal of Labor Economics*, 31(3):535–576.

Burgess, S., Greaves, E., Vignoles, A., and Wilson, D. (2015). What parents want: school preferences and school choice. *The Economic Journal*, 125(587):1262–1289.

Caetano, G. (2016). Neighborhood sorting and the valuation of public school quality. *Unpublished Manuscript.*

Chetty, R., Friedman, J. N., and Saez, E. (2013). Using differences in knowledge across neighborhoods to uncover the impacts of the eitc on earnings. *American Economic Review*, 103(7):2683–2721.

Chetty, R., Hendren, N., and Katz, L. F. (2016). The effects of exposure to better neighborhoods on children: new evidence from the moving to opportunity experiment. *American Economic Review*, 106(4):855–902.

Coibion, O., Gorodnichenko, Y., and Koustas, D. (2017). Consumption inequality and the frequency of purchases. *NBER Working Paper No. 23357*.

Dee, T. S., Dobbie, W., Jacob, B. A., and Rockoff, J. (2016). The causes and consequences of test score manipulation: evidence from the new york regents examinations. *NBER Working Paper No. 22165*.

Department for Children, Schools and Families (2009). National curriculum assessments at key stage 2 in England. *Statistical First Release*, 32.

Department for Education (2016). Schools, pupils and their characteristics. *Statistical First Release*, 20.

Diamond, R. and Persson, P. (2016). The long-term consequences of teacher discretion in grading of high-stakes tests. *NBER Working Paper No. 22207*.

Duranton, G. and Puga, D. (2015). Urban land use. *Handbook of Regional and Urban Economics*, 5:467–560.

Eyles, A. and Machin, S. (2015). The introduction of academy schools to England's education. *CEP Discussion Paper No. 1368*.

Fack, G. and Grenet, J. (2010). When do better schools raise housing prices? Evidence from Paris public and private schools. *Journal of Public Economics*, 94:59–77.

Figlio, D. N. and Lucas, M. E. (2004). What's in a grade? School report cards and the housing market. *American Economic Review*, pages 591–604.

Gibbons, S. and Machin, S. (2003). Valuing English primary schools. *Journal of Urban Economics*, 53:197–219.

Gibbons, S., Machin, S., and Silva, O. (2013). Valuing school quality using boundary discontinuities. *Journal of Urban Economics*, 75.

Hanna, R. N. and Linden, L. L. (2012). Discrimination in grading. *American Economic Journal: Economic Policy*, 4(4):146–168.

Hastings, J. S. and Weinstein, J. M. (2008). Information, school choice, and academic achievement: evidence from two experiments. *The Quarterly Journal of Economics*, 123(4):1373–1414.

Hussain, I. (2015). Subjective performance evaluation in the public sector: evidence from school inspections. *Journal of Human Resources*, 50(1):189–221.

Hussain, I. (2017). Do consumers respond to short-term innovations in school productivity? Evidence from the housing market and parents' school choices. *Unpublished Manuscript.*

Jacob, B. A. and Levitt, S. D. (2003). Rotten apples: an investigation of the prevalence and predictors of teacher cheating. *The Quarterly Journal of Economics*, pages 843–877.

Kleven, H. J. (2016). Bunching. *Annual Review of Economics*, 8.

Lavy, V. (2008). Do gender stereotypes reduce girls' or boys' human capital outcomes? Evidence from a natural experiment. *Journal of Public Economics*, 92:2083–2105.

Lavy, V. (2009). Performance pay and teachers' effort, productivity, and grading ethics. *American Economic Review*, 99(5):1979–2011.

Lavy, V. and Megalokonomou, R. (2017). Persistency in teachers' grading biases and effect on longer term outcomes: university admission exams and choice of field of study. *Unpublished Manuscript.*

Lavy, V. and Sand, E. (2015). On the origins of gender human capital gaps: short and long term consequences of teachers' stereotypical biases. *NBER Working Paper No. 20909.*

Machin, S. (2011). Houses and schools: valuation of school quality through the housing market. *Labour Economics*, 18:723–729.

Mizala, A. and Urquiola, M. (2013). School markets: the impact of information approximating schools' effectiveness. *Journal of Development Economics*, 103.

Neal, D. (2013). The consequences of using one assessment system to pursue two objectives. *Journal of Economic Education*, 44(4):339–352.

Neal, D. and Schanzenbach, D. W. (2010). Left behind by design: proficiency counts and test-based accountability. *The Review of Economics and Statistics*, 92(2):263–283.

Ries, J. and Somerville, T. (2010). School quality and residential property values: evidence from Vancouver rezoning. *The Review of Economics and Statistics*, 92(4):928–944.

Stroebel, J. and Vavra, J. (2016). House prices, local demand, and retail prices. *Kilts Center for Marketing at Chicago Booth - Nielsen Dataset Paper Series 1-030*.

Tannenbaum, D. I. (2015). Does school quality affect neighborhood development? Evidence from a redistricting reform. *Unpublished Manuscript*.

Terrier, C. (2016). Boys lag behind: how teachers' gender biases affect student achievement. *IZA Discussion Paper No. 10343*.

The Association of Convenience Stores (2017). The local shop report.

Wilson, D., Croxson, B., and Atkinson, A. (2006). What gets measured gets done. Head teachers' responses to the English secondary school performance management system. *Policy Studies*, 27(2):153–71.

Appendix A Sample Selection

Geographic hierarchies and area selection

The analysis is limited to Middle Layer Super Output Areas (MSOAs) located in metropolitan counties and Greater London and urban MSOAs in non-metropolitan counties. MSOAs are a geographic hierarchy developed by the Office for National Statistics (ONS) consisting of 7,194 homogeneous areas (6,781 in England and 413 in Wales) with a minimum population of 5,000 (an average of 7,200) and a minimum resident household of 2,000 (an average of 3,000). Metropolitan counties, non-metropolitan counties, and the region of Greater London are official administrative subdivisions in England. The final sample of areas consists of 6,133 MSOAs in England (90.44% of the total in the country). The listing of MSOAs is obtained according to the following steps.

- There are six metropolitan counties, typically with populations of 1.2 to 2.8 million: Greater Manchester, Merseyside (e.g., Liverpool), South Yorkshire (e.g., Sheffield), Tyne and Wear (e.g., Newcastle), West Midlands (e.g., Birmingham) and West Yorkshire (e.g., Leeds and Bradford). We keep all 1,499 MSOAs in this group of areas (24.44% of the final sample).

- Greater London comprises districts around London, and its structure is similar to that of metropolitan counties. We keep all 983 MSOAs in this region (16.03% of the final sample).

- Non-metropolitan counties consist of areas not in the two points above. The definition of rurality follows from the official classification of Output Areas (OAs) published by the ONS and based on land use. We classify as rural those MSOAs with the majority of OAs (above 50%) falling within the "Small Town and Fringe areas", "Village" or "Hamlet and Isolated Dwelling" categories and areas for which the surrounding areas are sparsely populated. This classification leaves us with 3,651 urban MSOAs in this group of areas (59.53% of the final sample).

Selection of schools and catchment areas

We consider community, faith and foundation public primary schools in the MSOAs identified in the previous section. There are 12,978 such schools in the register of educational establishments provided by the Department for Education in the years 1998 to 2007. This is the period relevant to the construction of the effects of borderlining on school quality. We drop a limited number of special schools (e.g., pupil referral units) that are specifically organized to provide education to children with special needs (excluded, sick, or unable to follow the mainstream curriculum). The number and composition of schools are substantially stable over the period considered, without major mergers or institutional changes (e.g., transformation from community to autonomous or into academies).[22] The left-hand panel of Figure 3 shows the areas of England selected together with the listing of schools considered.

The neighborhood definition used in the analysis is restricted to Lower Layer Super Output Areas (LSOAs) of school catchments. The right-hand side of Figure 3 shows the ONS geography of LSOAs for the borough of Tower Hamlets in East London. We consider houses and amenities located in LSOAs selected by applying the following criteria.

- We compute the distance from the centroids of all LSOAs to the closest school and keep LSOAs with at least one school of the same LA within 0.4 miles (the 50th percentile of the student-school distance in the NPD data). This is the sample described in Section 3. The corresponding sample size gradient implied by the various selection steps is presented in Table A.1.

- The same procedure is replicated using a 0.3-mile and 0.6-mile radius (35th and 65th percentiles from the NPD student-school distance distribution, respectively), and the sensitivity of our findings in Section 5 is investigated considering these samples, which consist of 17,923 and 27,250 LSOAs, respectively.

Our sample cut implies that the same LSOA might belong to the catchment area of multiple schools. This is important for the computation of schools z_{bn}, as discussed in Appendix B.

[22] The Edubase database, which can be accessed at http://www.education.gov.uk/edubase/home.xhtml, provides additional information on these institutional changes. There were approximately 203 academies in England until 2010, mainly at secondary school, and this number grew in the last five years. Currently, 2,075 out of 3,381 secondary schools are academies, while 2,440 of 16,766 primary schools have academy status.

Table A.1: Sample selection criteria

	Census blocks (1)	Schools (2)	Students (3)
In England	32,482	17,961	5,874,230
Drop private schools and school with special status		15,711	5,732,873
Schools with non-missing geographical data		15,607	5,702,658
Census blocks:			
- in metropolitan counties	7,183		
- in the Greater London region	4,765		
- in urban non-metropolitan counties	17,676		
In the three areas above	29,624	12,978	5,241,535
- with at least 5 years in the sample		12,258	5,144,587
With at least one school within:			
- 0.6 miles	27,250	11,724	5,072,684
- 0.4 miles	23,774	11,484	5,009,817
- 0.3 miles	17,923	10,985	4,834,063

Note. This table shows the selection criteria applied to define the working samples. Column (1) shows the number of Census blocks left after each step; column (2) shows the number of schools left; column (3) shows the number of students left.

Appendix B Effects of Borderlining

The variable z_{bn} is defined to proxy the effects of borderlining on test scores and school quality measurements published in the performance tables. This variable is constructed by estimating equation (1) for each LSOA (block) and pooling tests for all schools having that LSOA in their catchment (as explained in Appendix A). To gain precision, we estimate a LSOA's missing mass using the share of students f_{scj} scoring $s \in [-8, 7]$ around cutoff c (Level 3, Level 4 and Level 5) for subject j (English, math and science) by pooling (a) tests from 1998 to 2007 and (b) all schools associated to the LSOA within a certain radius. The following equation is estimated for each LSOA:

$$f_{scj} = \alpha(j) + \sum_{i=0}^{2} \beta_i s^i + \sum_{i=-3}^{1} \gamma_i 1(s = i) + \varepsilon_{scj}, \tag{5}$$

and by attainment cutoff. This equation is a variant of the specification of equation (1) discussed in the main text. The value $\sum_{i=-3}^{-1} \gamma_i$ is then calculated, which represents our estimate of the notch induced by bordelining for schools having the LSOA considered in their catchment. We then obtain block-specific z_{bn} values by iterating this procedure over all LSOAs and using the following definition:

$$\frac{\sum_{i=-3}^{-1} \gamma_i}{\sum_{s=-3}^{-1} \sum_{i=0}^{2} \beta_i s^i}.$$

Panel A of Table B.1 reports deciles of the distribution of $\sum_{i=-3}^{-1} \gamma_i$, the missing mass, across LSOAs by attainment threshold - see columns (1), (3) and (5). We also present in columns (2), (4) and (6) the average p-values after grouping estimates by decile. The missing mass in the score distributions is precisely estimated. For example, the value 0.0087 in column (6) represents the average p-value associated to estimates of the Level 5 missing mass falling in the 5th decile. Using the same format, Panel B of Table B.1 reports estimates of $\sum_{s=-3}^{-1} \sum_{i=0}^{2} \beta_i s^i$, the counterfactual share of students scoring below thresholds. The resulting distribution across LSOAs of z_{bn} at the Level 5 threshold, which is the quantity considered for the causal effects in the main analysis, is the green histogram in Figure B.1. Figure B.2 shows how the variable z_{bn}, computed using scores from 1998 to 2007 around Level 5, is used in the analysis. The outcomes are computed from 2008 onwards, while placebo tests as in Table 5 are carried out using data until 1997 (see Figure B.2 for a visual representation of the timeline considered).

The effect of borderlining should be zero, by design, at any point in the distribution three marks away from the achievement thresholds. It follows that z_{bn} should be centered at zero when estimated from (5) away from critical points. We replicate the procedure above using the share of students f_{scj} scoring $s \in [-8, 7]$ in a window centered at scores away from relevant thresholds and present the summary statistics in Table B.2 using the same format discussed above. The placebo thresholds are located 10 marks below math thresholds and 9 marks below the English and science thresholds to avoid the inclusion of the true thresholds in the $s \in [-8, 7]$ interval. A comparison with Table B.1 reveals the expected pattern, with the notches estimated below fictitious cutoffs being centered at zero and precisely estimated.

Figure B.1 also reports the distribution across LSOAs of Level 5 z_{bn} values generated at placebo thresholds. This histogram is centered at zero, as expected, with the upper tail markedly below the bulk of the distribution corresponding to real thresholds.

Table B.1: Summary statistics for bunching estimates

	Level 3		Level 4		Level 5	
	(1)	(2)	(3)	(4)	(5)	(6)
			Panel A. Missing Mass			
Decile	Estimate	P-Value	Estimate	P-Value	Estimate	P-Value
1	-0.0031	0.1777	-0.0002	0.6297	0.0037	0.3120
2	-0.0012	0.3682	0.0029	0.2699	0.0082	0.0475
3	-0.0005	0.6458	0.0044	0.1283	0.0102	0.0236
4	0.0000	0.9014	0.0057	0.0668	0.0118	0.0136
5	0.0004	0.6650	0.0067	0.0423	0.0132	0.0087
6	0.0009	0.4326	0.0078	0.0322	0.0146	0.0061
7	0.0014	0.2825	0.0089	0.0223	0.0162	0.0044
8	0.0021	0.1724	0.0103	0.0160	0.0179	0.0033
9	0.0031	0.0985	0.0122	0.0131	0.0203	0.0029
10	0.0058	0.0505	0.0167	0.0087	0.0258	0.0024
			Panel B. Counterfactual Share Below Threshold			
	Estimate	P-Value	Estimate	P-Value	Estimate	P-Value
1	0.0172	0.0143	0.0383	0.0000	0.0656	0.0001
2	0.0127	0.0136	0.0381	0.0000	0.0678	0.0000
3	0.0122	0.0116	0.0414	0.0000	0.0695	0.0000
4	0.0118	0.0075	0.0414	0.0000	0.0712	0.0000
5	0.0120	0.0051	0.0430	0.0000	0.0726	0.0000
6	0.0129	0.0054	0.0453	0.0000	0.0745	0.0000
7	0.0135	0.0054	0.0465	0.0000	0.0758	0.0000
8	0.0149	0.0085	0.0481	0.0000	0.0779	0.0000
9	0.0165	0.0026	0.0505	0.0000	0.0815	0.0000
10	0.0223	0.0045	0.0557	0.0000	0.0860	0.0000
Observations	23,774		23,774		23,774	

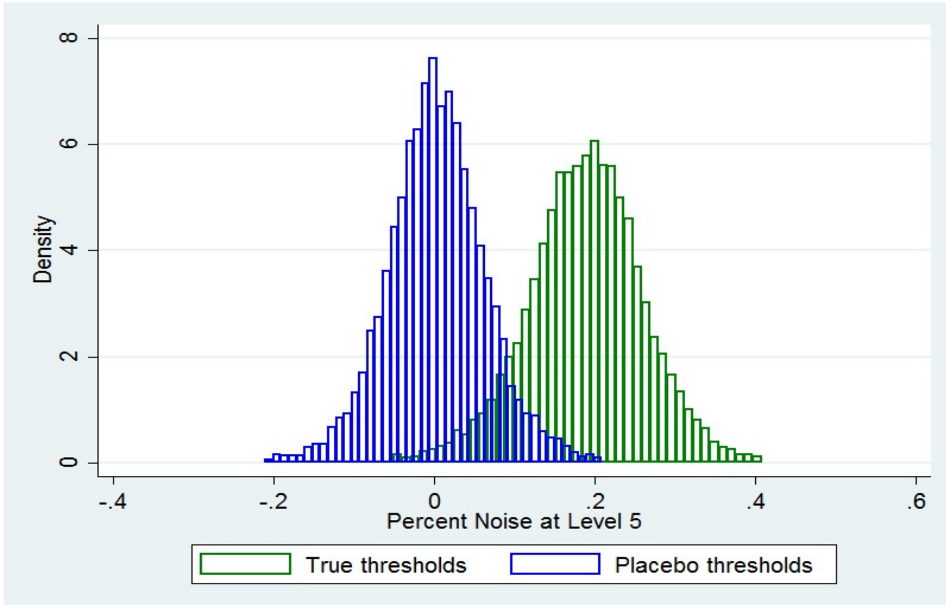
Note. This table shows estimates for the missing mass induced by the borderlining (Panel A) and the share of students below thresholds (Panel B). Deciles of the distribution of estimates across blocks (LSOAs) are reported in columns (1), (3) and (5). Columns (2), (4) and (6) present the average p-value of estimates by decile. All estimates are computed considering schools within 0.4 miles of a block's centroid.

Table B.2: Summary statistics for bunching estimates with placebo thresholds

	Level 3		Level 4		Level 5	
	(1)	(2)	(3)	(4)	(5)	(6)
			Panel A. Missing Mass			
Decile	Estimate	P-Value	Estimate	P-Value	Estimate	P-Value
1	-0.0033	0.1629	-0.0049	0.0878	-0.0071	0.0941
2	-0.0015	0.2338	-0.0025	0.2217	-0.0037	0.2371
3	-0.0010	0.3264	-0.0015	0.3865	-0.0022	0.4271
4	-0.0006	0.4716	-0.0008	0.6059	-0.0012	0.6580
5	-0.0003	0.6603	-0.0002	0.8722	-0.0002	0.9177
6	0.0000	0.8929	0.0003	0.8326	0.0007	0.7964
7	0.0003	0.7097	0.0010	0.5595	0.0017	0.5454
8	0.0006	0.4440	0.0017	0.3505	0.0028	0.3434
9	0.0012	0.2714	0.0028	0.2083	0.0043	0.1997
10	0.0034	0.1405	0.0055	0.0971	0.0081	0.0814
			Panel B. Counterfactual Share Below Threshold			
	Estimate	P-Value	Estimate	P-Value	Estimate	P-Value
1	0.0146	0.1177	0.0352	0.0005	0.0711	0.0000
2	0.0074	0.1079	0.0316	0.0002	0.0695	0.0000
3	0.0063	0.0839	0.0305	0.0009	0.0693	0.0000
4	0.0058	0.0776	0.0293	0.0004	0.0688	0.0000
5	0.0055	0.0745	0.0291	0.0007	0.0690	0.0000
6	0.0068	0.0508	0.0294	0.0005	0.0688	0.0000
7	0.0061	0.0645	0.0297	0.0000	0.0693	0.0000
8	0.0070	0.0502	0.0312	0.0002	0.0690	0.0000
9	0.0091	0.0468	0.0327	0.0001	0.0696	0.0000
10	0.0184	0.0315	0.0378	0.0006	0.0719	0.0000
Mean	23,774		23,774		23,774	

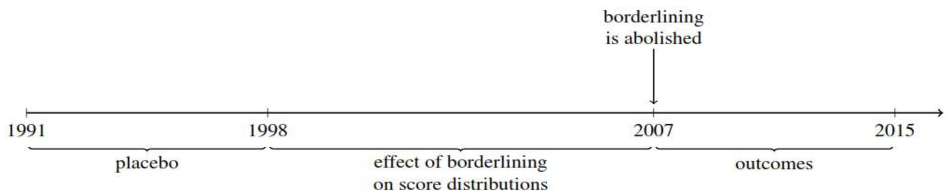
Note. This table shows estimates for the missing mass computed at placebo thresholds (Panel A) and the share of students below thresholds (Panel B). Placebo thresholds are set 10 scores below level thresholds for math and 9 scores below level thresholds for English and science. Deciles of the distribution of estimates across blocks (LSOAs) are reported in columns (1), (3) and (5). Columns (2), (4) and (6) present the average p-values of the estimates by decile. All estimates are computed considering schools within 0.4 miles of a block's centroid.

Figure B.1: Percent noise at Level 5 (real and placebo thresholds)



Note. This figure shows the distribution of the estimated effects of borderlining across blocks (LSOAs) using real thresholds or placebo thresholds. Estimates are obtained from equation (5) using the 0.4-mile radius sample.

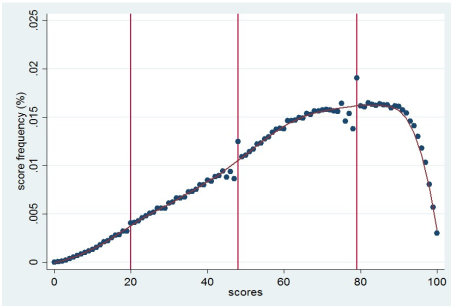
Figure B.2: Timeline



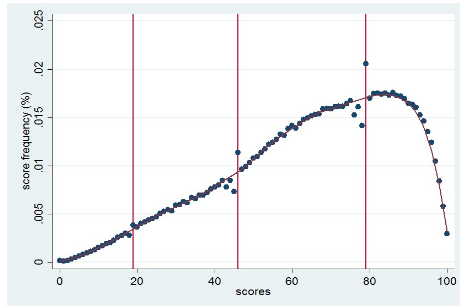
Note. This figure shows the timeline of events. Placebo regressions are computed using variables from 1991 to 1998. The incidence of score manipulation induced by borderlining is computed by pooling data from 1998 to 2007. Effects on outcomes are computed for years after the abolition of borderlining, from 2008.

Appendix C Additional Figures

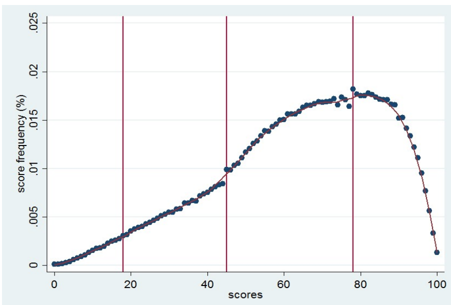
Figure C.1: Score manipulation over time: math



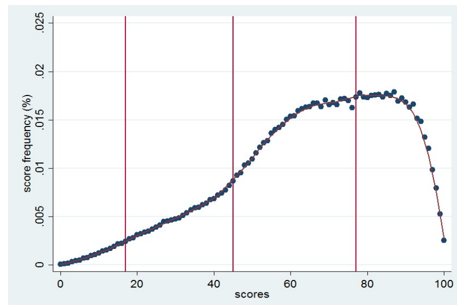
Panel A: 2004



Panel B: 2007



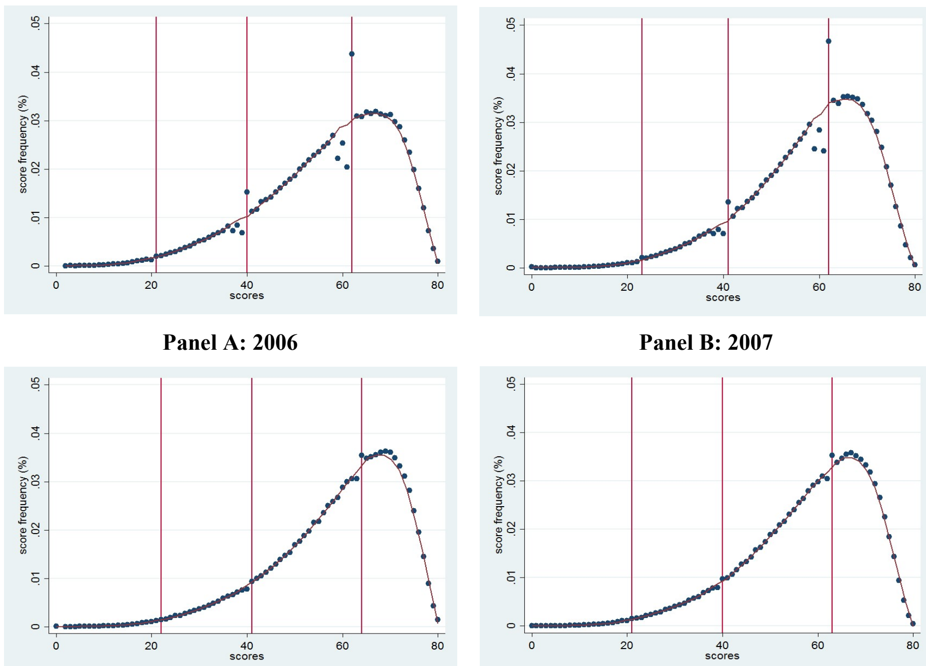
Panel C: 2008



Panel D: 2011

Note. This figure shows Key Stage 2 score distributions for math in selected years before (2004 and 2007) and after (2008 and 2011) the removal of borderlining. In each panel, the vertical lines are critical attainment thresholds set in that year. The continuous line is a local linear regression fit obtained excluding observations in the [-3,2] windows around thresholds.

Figure C.2: Score manipulation over time: science

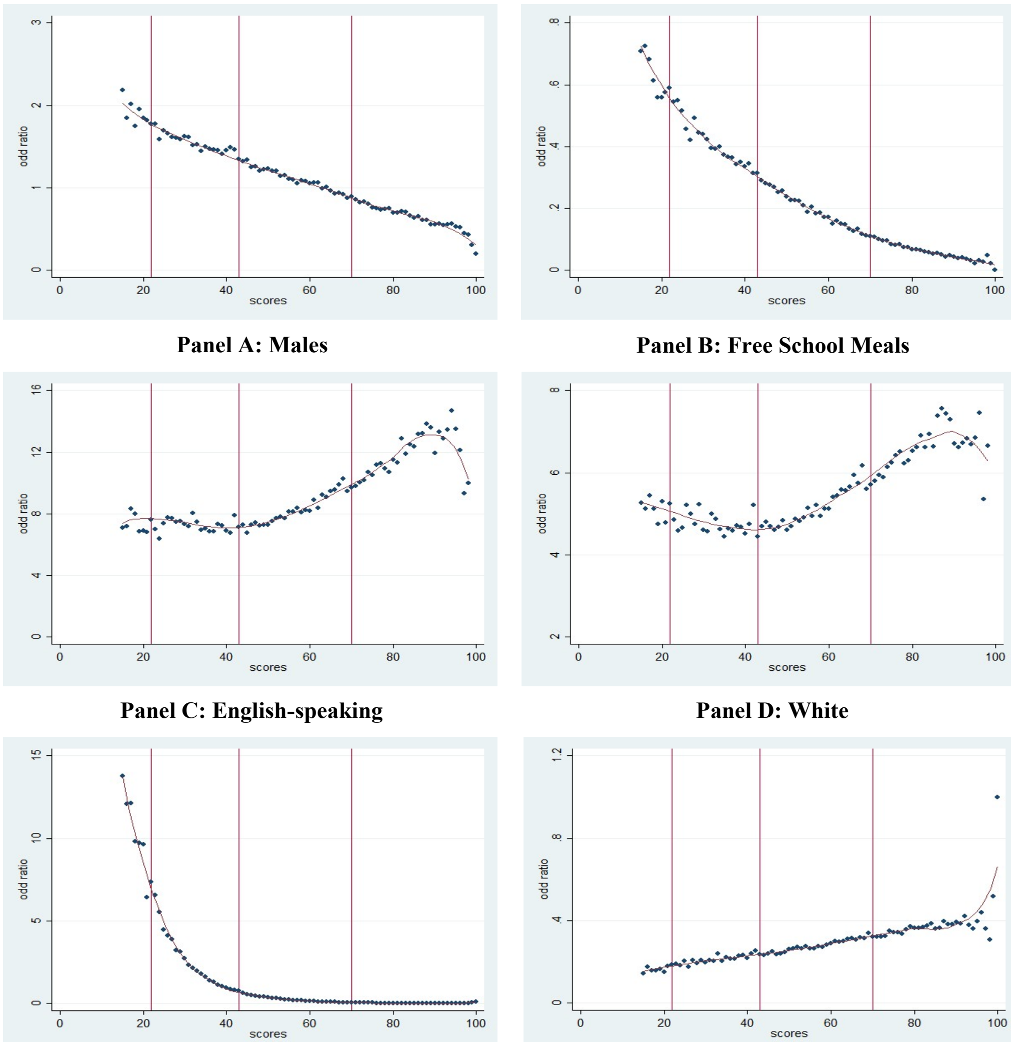


Panel A: 2006 **Panel B: 2007**

Panel C: 2008 **Panel D: 2009**

Note. This figure shows Key Stage 2 score distributions for science in selected years before (2006 and 2007) and after (2008 and 2009) the removal of borderlining. In each panel, the vertical lines are critical attainment thresholds set in that year. The science test was dismissed in 2010. The continuous line is a local linear regression fit obtained excluding observations in the [-3,2] windows around thresholds.

Figure C.3: Anatomy of manipulation



Panel A: Males

Panel B: Free School Meals

Panel C: English-speaking

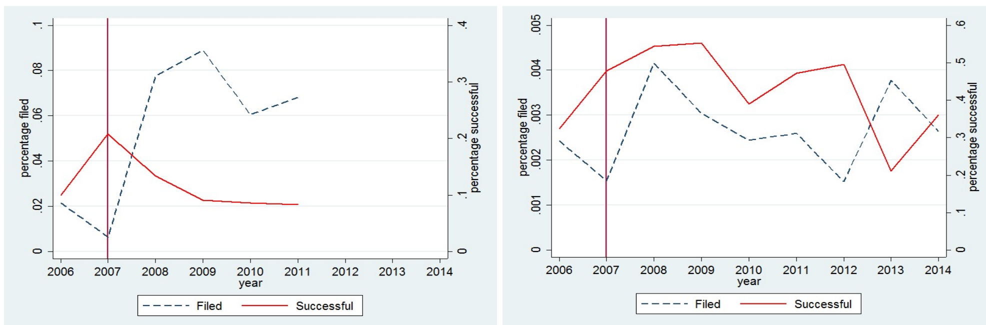
Panel D: White

Panel E: Special Education Needs

Panel F: Autonomous schools

Note. This figure shows odds ratios for selected school and student characteristics by values of the English score in 2007. Odds ratios are constructed as the probability of an event (e.g., the student is male) divided by one minus this probability. The following variables are considered: gender of student (Panel A), student is on free school meals (Panel B), student speaks English (Panel C), ethnicity of student (Panel D), student with special education needs (Panel E) and school is autonomous (Panel F). In each panel, the vertical lines are critical attainment thresholds set in 2007. The continuous line is a local linear regression fit obtained excluding observations in the [-3,2] windows around thresholds.

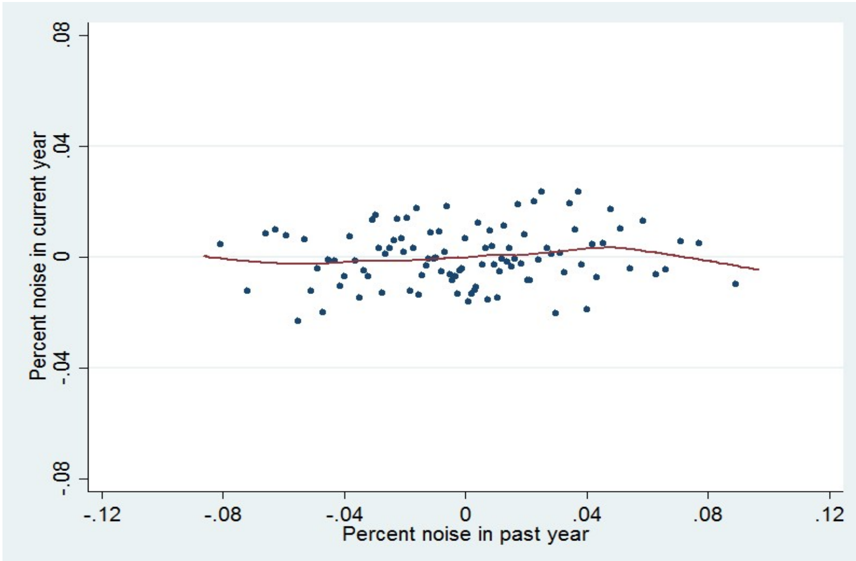
Figure C.4: Administrative data on school appeals



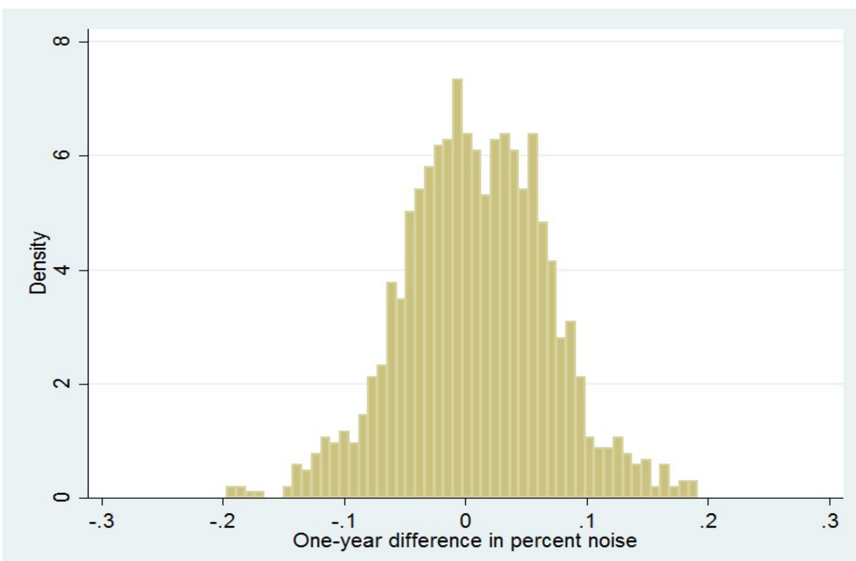
Panel A: Language **Panel B: Math**

Note. The two panels show appeals filed by school as the fraction of the number of scripts marked (dashed line) and successful appeals as the fraction of the number of appeals (solid line). Panel A refers to language scripts, whereas Panel B refers to math scripts. The language test was dismissed in 2012. Vertical line denotes the last year before abolition of borderlining (2007).

Figure C.5: Serial correlation of percent noise



Panel A. Autocorrelation



Panel B. Percent Noise Difference

Note. Panel A of this figure shows the first-order auto-correlation in the extent of score manipulation and plots the percent noise at Level 5 in the current year (on the vertical axis) against the percent noise lagged by one year (on the horizontal axis). The unit considered is the Local Authority (LA), and superimposed is a locally weighted fit. Panel B plots the distribution of the one-year differences in percent noise at Level 5, again using the LA as the

Figure C.6: Example of open-ended question in reading tests

13. Why do you think many people admire Evelyn Glennie?

13

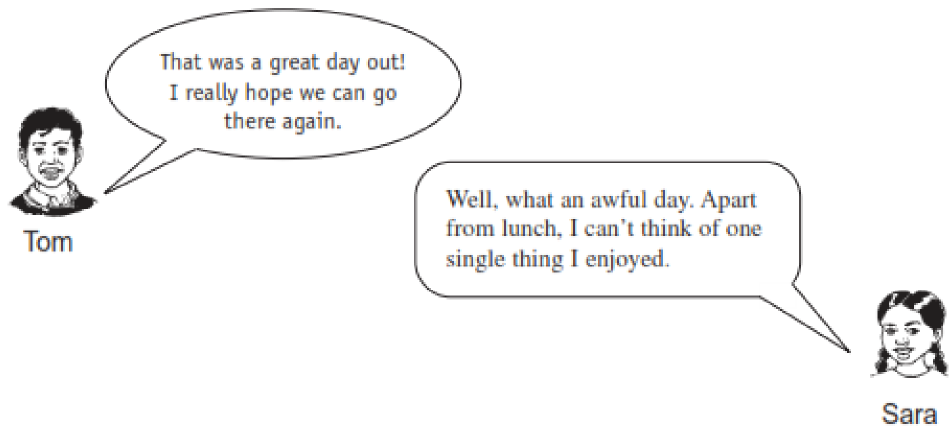
3 marks

Figure C.7: Example of open-ended question in writing tests

Dear Diary ...

A brother and sister went on a day out with their family.

Tom really enjoyed the outing, but Sara did not.



When they returned home, Tom and Sara wrote about the day in their diaries.

Your task is to write Tom and Sara's diary entries.

Use your imagination to decide what Tom and Sara would write in their diaries.

Figure C.8: Example of open-ended question in math tests

15

Here is a number chart.

Every third number in the chart has a circle on it.

1	2	3	4	5
6	7	8	9	10
11	12	13	14	15
16	17	18	19	20
21	22	...		

The chart continues in the same way.

Here is another row in the chart.

Draw the missing circles.

71	72	73	74	75

15a

1 mark

Will the number **1003** have a circle on it?
Circle **Yes** or **No**.

Yes / No

Explain how you know.

15b

1 mark