Oxford Economic Papers, 2019, 1–24
doi: 10.1093/oep/gpz028

OXFORD

Do selective schooling systems increase inequality?

By Simon Burgess[a], Matt Dickson[b], and Lindsey Macmillan[c]

[a]Department of Economics, University of Bristol, 2 Priory Rd, Bristol BS8 1TX; e-mail: simon.burgess@bristol.ac.uk

[b]Institute for Policy Research, University of Bath, BA2 7AY; e-mail: m.dickson@bath.ac.uk

[c]Department of Social Science, UCL Institute of Education, 20 Bedford Way, London, WC1H 0AL; e-mail: l.macmillan@ucl.ac.uk

Abstract

We investigate the impact on earnings inequality of a selective education system in which school assignment is based on initial test scores. We use a large, representative household panel survey to compare adult earnings inequality of those growing up under a selective education system with those educated under a comprehensive system in England. Controlling for a range of background characteristics and the current location, the wage distribution for individuals who grew up in selective schooling areas is substantially and significantly more unequal. The total effect sizes are large: 24% of the raw 90–10 earnings gap and 19% of the conditional 90–10 earnings gap can be explained by differences across schooling systems.

JEL classifications: I24, J31.

1. Introduction

One of the key roles of any education system is to define the mechanism that assigns children to schools. The choice of mechanism is likely to affect the level and distribution of schooling outcomes and therefore later life outcomes. One such mechanism is to assign students based on test scores: those with high scores attend one school, those with lower scores go elsewhere. This is like tracking, but across schools rather than within school. In England this is known as the grammar school system, and was used to allocate children to schools from the time of a unified system of education in the 1940s through the 1980s; now only a few areas still use this as the main system (see Appendix A for full details of institutional background). Nevertheless, grammar schools continue to be a prominent policy issue in England, with the UK government consulting in 2016 on the possibility of reintroducing or expanding selection by ability and £50million being allocated to the expansion of new grammar schools in May 2018. There is a parallel debate in the US about elite or exam schools, reviewed below.

Much of the research on grammar schools has focused on two important questions: who gets into grammar schools (is access 'fair'?), and what is the impact of attending a grammar school (is there a causal gain in attainment?). There is much less evidence focusing on the system as a whole, namely comparing the outcomes of students assigned by one mechanism (by ability) compared to another (choice, or 'by house prices'). That is the contribution of this article: we examine the impact of a grammar school *system* on earnings inequality.

We use data from a large and representative household panel data set for England and compare the spread of the earnings distribution in middle age. We match existing selective systems to very similar comprehensive systems, based on a wide range of area characteristics such as political control, average wages and proportion of highly educated individuals living in the area. The richness of our individual-level data means that we know where an individual grew up and can map this back to the nature of the education system in that place at that time. This allows us to categorize respondents as growing up in a selective system or not, and to compare the earnings distribution they experience as adults, separating the effect of grammars from the local labour market effect using variation in those who move across local education authorities (LEAs). We can also control for the parental background of the individual, as well as their current geographic location.

We show that individuals who grew up in areas operating a selective schooling system have a more unequal wage distribution in later life. Those growing up in selective systems who make it to the top of the earnings distribution are significantly better off compared to their similar non-selective counterparts. For those at the bottom of the earnings distribution, those growing up in a selective system earn significantly *less* than their matched non-selective counterparts. These differences remain after controlling for a range of background characteristics and current local area. In summary, there are both winners and losers from the grammar system: the additional earnings differential between the 90th and 10th percentiles in selective systems accounts for 24% of the total raw 90–10 earnings gap and 19% of the conditional 90–10 earnings gap. This article establishes this new fact for policy discussion. We do not attempt to unpack this into different component mechanisms here (though obviously, the impact on educational attainment is likely to be key).

In the next section we review the related literature on the impact of selective systems on later outcomes before describing the framework for our analysis in Section 3. Our empirical approach and the data used are outlined in Section 4 and our results are presented in Section 5. We end with some brief conclusions.

2. Related literature

Much of the previous literature on selective schools focuses on the benefit to the marginal student of attending a grammar school compared to not attending. In Great Britain, Clark (2010) uses access data from East Riding (a local government district in the north of England) to estimate the causal impact of attending a grammar school on attainment at 16, the types of course taken and university enrolment. He finds small effects of grammar schools on test scores at 16 but larger effects on longer-run outcomes such as taking more academic courses—which allow access to A levels and university enrolment. Similarly, Clark and Del Bono (2014) implement a regression discontinuity design to assess the impact of attending a grammar school for a cohort of young people born in Aberdeen, Scotland, in the 1950s. They find large effects on educational attainment, and for women there are longer-term impacts on labour market outcomes and reduced fertility. For men there were no long-term impacts identified.

Abdulkadiroglu *et al.* (2012) and Dobbie and Fryer (2011) assessed the impact of attending exam schools in Boston and New York on attainment and test scores. Both studies found limited impacts on student achievements from attending these selective schools, though Dobbie and Fryer (2011) found that these students were more likely to choose more academically rigorous subjects. Dustmann *et al.* (2014) similarly found little impact of the marginal student attending a more advanced track on their longer-term outcomes.

Sullivan and Heath (2002) and Galindo-Rueda and Vignoles (2005) used the National Child Development Study (NCDS) data from the UK to compare the outcomes of those attending grammar schools to comprehensive schools and secondary moderns. Both use a value-added approach alongside school-level controls to assess the impact of the different schools on educational attainment. In addition, Galindo-Rueda and Vignoles (2005) also instrument school type with the political power of the LEA at the time, arguing that the political power of the LEA at the time of reform affected the speed at which the systems were switched from selective to mixed schooling. Both studies find significant positive effects on attainment of grammar education compared to comprehensives although Manning and Pischke (2006) use a falsification test of value-added from age 7 to 11 to show that these studies are still affected by selection bias.

A slightly different question is addressed by Guyon *et al.* (2012), who use data from Northern Ireland and exploit a policy change that compelled grammar schools to increase the number of children admitted each year. The change induced a discontinuous increase in the proportion of the school year group going to grammar schools, and this is used to identify the effect of school segregation by ability on the average performance in examinations taken at age 16, at age 18 and on university entrance rates. Rather than the impact on the marginal students who are shifted into the grammar school by the policy change, the estimates provide an assessment of the impact on the whole distribution. Guyon *et al.* (2012) find substantial positive impacts of the increased grammar attendance on average examination results and university entrance. However, as we might anticipate, disaggregating this into the impact on the grammar school results and the impact on the non-grammar school results, reveals a *negative* impact on the average results in the non-grammar schools as a consequence of the change in student composition induced by the policy.

While each of these approaches have clear strengths, and Guyon *et al.* in particular look at the distribution of results not just the effect on the marginal student, these studies say little about differences across selective and non-selective *systems*. Closer to our study are those of Atkinson *et al.* (2006) and Jesson (2000), who use data from the National Pupil Database (NPD) for England and Wales, to compare LEAs that are still selective now to non-selective LEAs. These studies are therefore more in line with our research, comparing the outcomes of pupils in systems as a whole rather than the outcomes of the marginal pupil who makes it into a grammar school. Both Jesson (2000) and Atkinson *et al.* (2006) use NPD data to compare value added attainment across selective and non-selective LEAs. While Jesson (2000) is open to the critique of Manning and Pischke (2006) that value-added alone does not remove selection bias, Atkinson *et al.* (2006) match LEAs to attempt to control for this. They show that prior attainment when comparing selective LEAs to the comprehensive population as a whole is much higher in the selective LEAs but when comparing prior attainment in the matched LEAs, this is very similar. While neither study finds evidence of higher attainment across selective and non-selective systems as a whole, Atkinson *et al.* (2006) find that grammar-educated children in selective LEAs outperform similar children in non-selective LEAs on average while non-grammar-educated children in selective LEAs underperform compared to similar children in non-selective LEAs. This is in line with our findings of greater inequality in earnings later in life for those from selective LEAs.

3. Framework

A selective school system, assigning individuals to schools based on their performance on a test, is one way of assigning students to schools. In England, the grammar school system assigns students to schools based on their performance on a test at age 11, commonly referred to as the '11+' test. Typically in LEAs that operate a grammar system, students who achieve above a certain threshold are entitled to a place at a grammar school while students below the threshold are entitled to a place at a secondary modern school.

We compare the outcome of this system to the main alternative in England, namely school choice. In England, this involves families stating their preferred schools. However, given that the better schools quickly become over-subscribed and the criterion for assigning students in this case becomes proximity of the student's home to the school, school choice quickly reduces down to neighbourhood schooling or 'selection by house prices'. We therefore consider the differences in outcomes between two systems where, in their simplest form, one allocates pupils to schools based on ability[1] and one allocates pupils to schools based on proximity.

We present a very simple framework for thinking about the earnings inequality implied by each system.

Think of a population, where students have ability, a, and parental resources, r. These have distributions with variances σ^2_{a} and σ^2_{r}; they are positively correlated with covariance σ_{ar}.

The schooling outcome, s, for student i depends on ability, school quality, q, and peer group ability $\bar{a}$:

$$s_i = s(a_i,\ \bar{a}_i,\ q_i).$$

Later adult earnings depend on both the ability of the student and her schooling outcome:

$$y_i = a_i + \gamma . s_i$$

where γ is the relative weight on schooling.

To determine the relative impacts of the alternative schooling systems on earnings inequality, we must evaluate how each system translates ability into outcomes and therefore what each system implies for $\bar{a}(a)$ and $q(a)$—that is, how each system relates student ability to peer group ability and to school (teacher) quality.

The school assignment mechanism is different in the two systems. In a grammar school system, each student is assigned to the grammar school if a potentially noisy function of her ability is above some threshold (determined by the number of places in the grammar schools relative to the population). In a choice-based comprehensive system, admission

1 Of course, there are issues concerning whether the tests used actually measure ability. Given the role of 'tutoring to the test', they are more likely to be measuring some mix of ability and attainment although this is not central to our analysis here.

depends on preferences and on priority. We could either assume random preferences or that all have preferences for high quality schooling; in either case, the driving force is priority. The most common priority rule in England is proximity: students living closest to the school are admitted. Under standard assumptions, the operation of the housing market means that houses nearby high performing schools are valued more highly[2] and so the likelihood of admission to the higher performing schools depends on family resources, r.

3.1 Grammar system—assignment through selection on ability

By definition, grammar school systems sort pupils based on their ability: so $\overline{a}(a)$ will be positive and very strong. Schools with high ability pupils are attractive to high ability teachers, hence we assume grammar schools attract and retain high quality teaching staff, hence $q(a)$ will be positive and strong.

$$s_i = s(a_i,\ \overline{a}_i(a_i),\ q_i(a_i)) = s_g(a_i) \text{ and earnings will be: } y_i = a_i + \gamma.s_g(a_i)$$

3.2 Comprehensive system—assignment through residential proximity to school

We assume that the high-quality schools are randomly distributed around an area. However, because of the proximity rule, the quality of the school attended depends on parental resources: $q(r)$. As a covariance exists between r and a, we can write this as $q(r(a))$. This also induces variation in peer groups, so $\overline{a}(a)$ again, but only through r. Therefore, there is also a positive association between peer groups and ability and teaching quality and ability in this system, although these work through the correlation between r and a rather than directly as in the grammar system.

$$s_i = s(a_i, \overline{a}_i(r(a_i)), q_i(r(a_i))) = s_c(a_i) \text{ and earnings will be: } y_i = a_i + \gamma.s_c(a_i).$$

Using these, we can express the variance of earnings in each system as:

$$var_k(y(a_i)) = \left\{\left(1 + \gamma s_k{}'(\mu_a)\right)\right\}^2 \sigma_a^2$$

where $k = g$ (grammar) or c (comprehensive). Consequently, $var_g(y) <$ or $> var_c(y)$ depending on whether $\frac{dS_g(a)}{da} <$ or $> \frac{dS_c(a)}{da}$.

Therefore, how the schooling system creates more equal or unequal wage distributions depends, among other things, on how the two systems translate individual ability into schooling outcomes. As we have seen, this will depend on how individual ability is related to peer group ability and how individual ability is related to school (teacher) quality in each system, both directly and indirectly via parental resources. These are empirical questions that we bring to the data.

4. Empirical analysis

To perfectly estimate the impact of selective systems compared to non-selective systems we would need to be living in an ideal world. This is the thought experiment: imagine two communities of identical families, growing up separately. One community has a grammar school system; the other has a comprehensive system (allocation by proximity). Following

2 This has repeatedly been shown to be the case for England, for a recent contribution see Gibbons *et al.* (2013).

their education, both sets of individuals go on to work in the same labour market. A comparison of the distribution of wages amongst those who grew up in the selective system with the distribution for those who grew up in the non-selective system, would tell us something about the impact of selective schooling on the whole distribution of wages.

Unfortunately, this thought experiment cannot be run in practice and we therefore have to use empirical methods to get as close to this ideal world as possible. In order to empirically test our model, we need to be able to compare the distribution of wages for individuals who grew up in LEAs[3] operating a selective mechanism for allocating students to schools, with the distribution amongst individuals who grew up in areas that were very similar along a number of relevant dimensions but that were operating the comprehensive system. This should ensure that we are not incorrectly attributing the effects of other area characteristics on later wages to the effect of growing up in a selective school area.

We use *Understanding Society* for our empirical analysis. This is a large longitudinal panel study following approximately 40,000 households in the UK, beginning in 2009. Information is collected from all individuals in the household aged 16 and over, on a wide range of topics, including parental background, labour market status and earnings. We make use of the special license release of the data, which includes the individual's age, current local authority of residence and crucially for our purposes, the local authority district where the individual was born. Each wave is collected over 24 months: the first was collected between January 2009 and January 2011, the second between January 2010 and January 2012—we make use of both of these waves in our analysis. Given our sample requirements and matching process, our final analysis focuses on 2033 individuals who were born in 33 selective or similar non-selective LEAs from 1961–1983 (from these 33 LEAs we have observations from 111 LEA*years that were selective, and observations from 172 LEA*years that were non-selective). Appendix C Fig. C1 illustrates how these LEAs are distributed between those always selective (14), those always non-selective (15) and those that are selective in some years, non-selective in others (4). The birth year range means that the individuals in our sample are aged between 25 and 50 and so are of prime working age.

4.1 Defining selectivity

We begin by defining LEAs of birth as selective or non-selective. Selectivity of an area is calculated using school level data from the Annual Schools Census: schools are allocated to their LEA then the aggregated LEA data is used to calculate the percentage of children aged 13[4] in the LEA who had a place allocated by the selective system (grammar or secondary modern places).[5] The time-series of data runs from 1967 to 1983, however post-1983 there

3 There are currently 152 LEAs, or local authorities as they are now known, in the UK (Department for Education, http://www.education.gov.uk/, last accessed 10 July 2018). The average population in a Local Authority in 1998 was 140,000 individuals ranging from 25,000 to just over 1,000,000 individuals.

4 The proportions were measured at age 13 rather than 11 or 12 because in some secondary schools (upper secondaries) children didn't start in the school until they were 13.

5 We are extremely grateful to Damon Clark for providing this data. The figures for each LEA in each year are gender specific as there were/are a non-trivial proportion of single-sex schools in selective areas. For our purposes, we average the male and female figures to give us an average measure of selectivity for an LEA in a year. For the LEAs in our sample, the difference between the male and female figures is very small or zero (the mean is 0.65 percentage points and median is 0.19 percentage points).

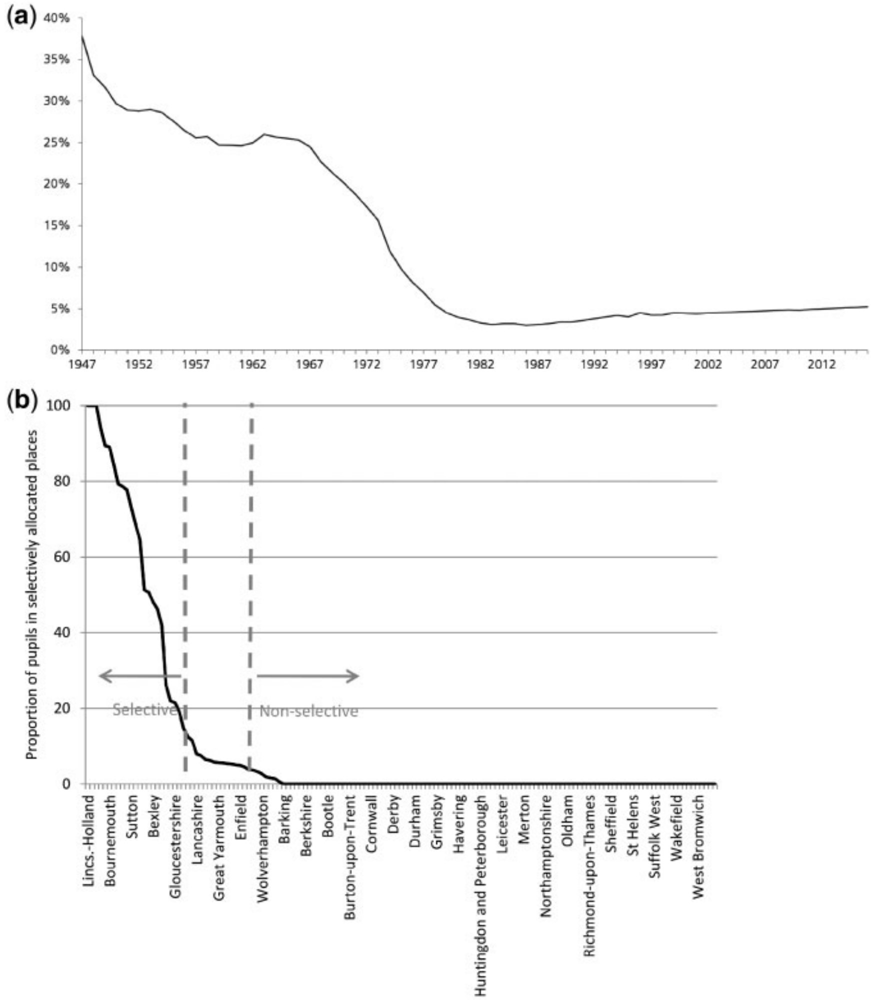

Fig. 1. (a) Proportion of pupils in state-funded grammar schools in England 1947–2016. *Source*: House of Commons Library Briefing Paper 1398 (March 2017), Fig. 2; see Bolton (2017). (b) Distribution of selectivity across LEAs in 1983.

has been very little further comprehensivization (see Crook, 2013) and so we make the assumption that the proportion of selective school places within an LEA has remained similar to the 1983 level henceforth. The upper panel of Fig. 1 shows that between 1983 and 1996 there is only a very slight change in the proportion of grammar school pupils across England as a whole.[6] We do not model the process by which LEAs retained or abandoned selective schools. It is likely to have been influenced by fixed factors such as the size and

6 Despite the total number of grammar school places increasing as grammar schools have expanded, this phenomenon has also been witnessed to a similar extent in other schools, leaving the overall proportion of pupils in grammar schools increasing only ‘very gradually’ over the past 25 years (see Bolton, 2017).

Table 1. Distribution of selectivity in all 145 LEAs for which we have data across all time-periods 1974–1996

	Selective %	Selective % conditional on >0
N	3,335	1,350
Mean	15.13	37.4
SD	29.25	35.8
10^{th}	0.0	2.4
25^{th}	0.0	5.6
50^{th}	0.0	21.7
75^{th}	11.3	70.0
90^{th}	73.0	99.7

Source: Authors' calculations using Understanding Society waves 1–2.

geography of the area (population density and the like) as well as local political control. Our assumption is that the matching of LEAs, discussed below, takes account of most of the statistical force of these factors, and within the matched set, the retention of selection is as good as random.

We define an LEA as selective if more than 20% of children in the LEA were assigned their school place by selection. We define non-selective LEAs as those where less than 5% of 13-year-old children were assigned by selection. As illustrated in the lower panel of Fig. 1, given the distribution of levels of selectivity, these thresholds mark a clear delineation between what were selective and non-selective areas. Table 1 illustrates the distribution of selectivity in 145 LEAs across the time-period considered (i.e. 1974 to 1996). Sixty per cent of LEA*time observations were 100% non-selective. Of those with any selectivity, 52% had greater than 20% selective schools within the LEA and 43% had greater than 30% selective schools. We consider whether our results are sensitive to these cut-offs at the end of the results section.

4.2 Matching

Having defined selectivity, we proceed by matching selective and non-selective LEAs by LEA*year on the basis of local political, socio-demographic, labour market and school market characteristics: the political party in control of the local government,[7] the proportion of the area's residents holding a degree, the proportion in social classes 1 or 2 (professional and managerial), the proportion of people economically active,[8] the local unemployment rate,[9] the local male hourly wage rate,[10] and the proportion of children

7 This information comes from the data from the Elections Centre (www.electionscentre.co.uk, last accessed 10 July 2018) recorded in 1981. We are grateful to Colin Rallings for providing this data.

8 These three variables are from the 1981 census.

9 Taken from the Employment Gazette, 1979 to 1998, county-level tables. Unemployment rates are matched to LEAs within counties with two LEAs in the same county taking the same unemployment rate.

10 Taken from the New Earnings Survey, 1974 to 1996, region and sub-region tables. The specific earnings variable used to match is the average hourly earnings excluding the effect of overtime for full-time male workers over the age of 21 whose pay for the survey pay-period was not affected by absence.

who attend private schools in the area.[11] Individuals turned 13 in different years in our data and hence the matching of LEAs is done separately for each year of our period of interest from 1974 to 1996. We select the three nearest neighbour non-selective LEAs for each selective LEA with replacement and retain only matches that share common support. Following the matching, we retain individuals who grew up in one of the selective or matched non-selective LEAs.

The key to success of our empirical strategy is to ensure that the selective and non-selective areas are matched with respect to characteristics that could impact upon the later earnings of individuals who grew up in these areas. Our matching variables capture essential features of the labour market, private school market, socio-demographics, and the political environment of the local area. These could all influence the propensity for an area to retain a selective system and also impact upon the later labour market outcomes of those growing up in the area. Matching on this broad set of covariates means that we are able to consider the impact of the school system on the earnings distributions of individuals who grew up in areas featuring the same local political control, the same socio-demographic mix and the same labour market conditions but differing in the selectivity of the school system. The ideal experiment would compare identical areas along all relevant dimensions bar the school system, but given data limitations, we believe that this combination of characteristics does a job of capturing the relevant dimensions that could affect later earnings.

Appendix B contains summary statistics for the quality of the covariate balancing after the matching process in each year. These show that for every year and for each of the variables used in the matching, *t*-tests unanimously and comprehensively fail to reject the hypothesis that there is no difference in the means of the variables between the selective and non-selective areas (the average of the absolute value of the *t*-stat across all variables and years is 0.37, median 0.32). Moreover, Kolmogorov-Smirnov tests of the equality of the distributions of the matching covariates between the selective and matched non-selective areas fail to find a significant difference in the distributions in all but 13 of 161 tests (average *p*-value over the 161 tests was 0.462).[12] Hence the post-match test statistics suggest that selective and non-selective LEAs are very well matched with respect to the levels and distribution of their labour market and school market characteristics.

4.3 Data and methodological issues

Ideally the characteristics that we match on would all be measured at exactly the time that the individuals attended secondary school and for much of our data this is the case. However, due to the non-availability of some of this information—in part due to the restructuring of local authority organization during the 1970s—there is some limit to the time-variation in some of the matching variables. For the local unemployment data, 15 of the 23 years of data are unaffected, but for eight of the years that we include in our analysis we have to assign the value for the nearest available year (which is a maximum of five years

11 Compiled using the National Pupil Database 2002. Results are robust to the exclusion of private schools from the matching process, see the Appendix Figs B2a and B3a and Tables B2a and B3a.

12 This is 8% of the tests, which is slightly above the type I error rate we would expect when conducting multiple hypothesis tests, however testing equality of distributions with small numbers ($n < 15$) in each distribution makes the test vulnerable to outliers skewing one distribution and leading to a rejection of the null of equality. All individual test statistics available from the authors.

distance and in the majority of cases three or fewer).[13] Our results are robust to the exclusion of years in which the unemployment information has to be mapped from a nearby year (see Appendix Figs B2b and B3b, Tables B2b and B3b).

Information on the proportion of children attending private/independent schools is only available at the local authority level from 2002 and so there is no time-variation in this variable. However, given that the proportion of full-time pupils in private/independent schools in England and the proportion of English schools that are private/independent has changed very little between the time we have our measure of private school density (2002) and the relevant period for our data (1974 to 1996),[14] it is reasonable to assume that the local private school density has not changed too dramatically and thus our measure is relevant for matching.[15]

Similarly, the information on the political party in control of local government comes from a single measure in 1981, and the information on the proportion of the area's residents holding a degree, the proportion in social classes 1 or 2, and the proportion of people economically active comes from the 1981 census. While time-variation in these variables would be preferable, to the extent that these measures capture persistent, slowly changing, characteristics of a local area, they will improve the matching mechanism. All of our results are robust to the exclusion of non-time-varying matching variables, see Appendix B.

It may be a concern that areas that had a selective schooling system were more diverse than those with a non-selective system and thus differences in prime-age earnings later in life reflect existing inequality rather than being a result of the differing systems. However, to the extent that the diversity of an area is captured by the unemployment rate, average wages, independent school density, local political control, the proportion of the area's residents holding a degree, the proportion in social classes 1 or 2 (professional and managerial), and the proportion of people economically active, we will be comparing similarly diverse areas. The post-matching balancing tests suggest that for each of these variables we achieve a close match, hence the selective and non-selective areas we compare are on average as diverse as each other. Moreover, since our wage regressions include fixed effects for the *current* local labour market area, the estimated effects are identified by those who move area between the time of their schooling and their later wage observations, therefore the estimates *cannot* simply reflect existing inequality in earnings between selective and non-selective areas, a point we return to below.

An obvious concern with our data is that we observe the LEA at birth rather than the LEA that the individual is enrolled into in secondary school. This raises two issues: children may attend a school across the LEA 'border' and so be educated under a different system; or families may move areas between the birth of the child and the start of secondary school. Regarding the first issue, we investigate the extent to which pupils cross borders in the NPD. On average around 11% of pupils attend a school in a different LEA from their LEA of residence. This is most likely to occur in London (over 20% cross-borders on average) where LEAs are small and close together, however given our preferred set of matching

13 In practice this means that for the years 1974 to 1978 each LEA has their 1979 level of unemployment and for the years 1994 to 1996 each LEA has their 1993 level of unemployment.

14 See Ryanand Sibieta (2010).

15 As a robustness test we perform the analysis when matching without private school proportion as a matching variable. Our results are robust to this test, see Appendix Figs B2a, B3a and Tables B2a and B3a.

covariates, London does not feature in the analysis. In robustness checks in which the matching covariates are changed and London does become a matched LEA, the results remain consistent and this is the case whether London is included or not in these checks. We argue that as our main results do not include London the potential of border crossing influencing the findings is much reduced. Moreover, if our alternative specification results are robust to the inclusion/exclusion of London where border crossing is most relevant, then our results are not likely to be driven by border crossing elsewhere which will be much less prevalent.[16]

We also argue that border crossing is likely to understate our findings to the extent to which border crossing across systems is made by: (i) those that are the most able in non-selective systems crossing borders to attend grammar schools; and (ii) those who do not make it into grammars in the selective areas crossing borders to attend comprehensives rather than secondary moderns. In the first case, these individuals will push up the top end of the non-selective earnings distribution if grammars increase earnings relative to comprehensives, and in the second case, these individuals will push up the bottom end of the selective earnings distribution if comprehensives increase earnings relative to secondary moderns. Both of these effects would lead us to *understate* the effects of the selective system at the top and bottom of the earnings distribution.

To consider the issue that families may move areas, we use data from two birth cohort studies, the British Cohort Study (BCS) following children born in 1970, the Millennium Cohort Study (MCS) following children born in 2000, and the NPD to investigate the extent to which we can observe families moving from birth to starting secondary school. The birth cohort studies provide information on movements from birth to age 10 in the BCS and from birth to age seven in the MCS, both at Government Office Region (GOR) level. The NPD provides information on moves from age five to 11 at the postcode level and Travel to Work Area (TTWA) level. As can be seen from Table 2, the vast majority of families do not move during childhood with 10% moving to a different postcode in the NPD data and 1% moving to a different TTWA. The data from the cohort studies suggests that while more families move before children start school, the numbers moving are still small with 7.3% in the BCS and 5.4% in the MCS moving before the cohort member is five.[17] Similar arguments to those regarding border crossing to attend a different school also apply to the potential for families to move between a child's birth and schooling and that this movement might be driven by parents' desire for their child to be educated under a particular system. However, it is again the case that this would lead us to understate the effect of the selective system. If ambitious parents who believe their child will get into the grammar school move from a non-selective to a selective area between a child's birth and schooling, we will include them in the non-selective distribution and if they attend the grammar school and it increases their earnings, this will be recorded as increasing the earnings at the top end of the non-selective distribution, biasing the selective effect downwards. Similarly if parents living in a selective area fear that their young child will fail to make the grade for the grammar school and move to a comprehensive area, this would be picked up as an

16 These additional robustness checks are available from the authors on request and are also contained in an earlier version of the article, see Burgess *et al*. (2014).

17 Families that move are typically slightly more affluent than families that do not move in both cohort studies. For example, in the BCS, families that moved between birth and age 10 had an average family income of £1462 a month compared to an average of £1304 for families who did not move.

Table 2. Proportion moving across different geographical areas during primary school

	Stay	Move
Postcode		
NPD 5–11	90.0	10.0
Travel to Work Area		
NPD 5–11	99.0	1.0
Government Office Region		
BCS		
0–5	92.7	7.3
5–10	95.5	4.5
010	89.4	10.6
MCS		
0–3	96.6	3.4
3–5	98.0	2.0
5–7	98.5	1.5
7–11	98.2	1.8
0–11	92.6	7.4

Source: Authors' calculations using British Cohort Study 1970, Millennium Cohort Study; NPD figures from Allen *et al*. (2010).

increase in earnings at the lower part of the *selective* distribution, again leading us to understate any negative effect of the selective system on the bottom of the distribution.

A final concern with our data is that we need individuals to move between school and when they are observed in the labour market as an adult in order to be able to separate out the effect of the schooling system from that of the local labour market. If everyone stayed where they went to school, our findings could be driven by the characteristics of the LEA that are related to labour market earnings and selection of the schooling system. For example, if selective LEAs were typically more diverse, thus had a more unequal earnings distribution, and individuals from selective LEAs stayed where they were from as adults, we would attribute the spurious association, or indeed reverse causation of inequality in selective areas, to selective areas causing inequality. Fortunately in our data, over 50% of the sample move LEAs between birth and adulthood. As illustrated in Table 3, looking at the largest LEAs operating each type of system the extent of movement is almost identical: 53.3% for those growing up in selective LEAs, 53.8% for those growing up in non-selective LEAs.[18] We therefore argue that we have enough variation in our data to be able to separate the effect of the school system from the effect of the LEAs' labour market characteristics.

4.4 Measuring earnings inequality

To compare earnings distributions in adulthood, we use hourly wages calculated from the recorded usual gross monthly pay including overtime, usual weekly hours and overtime hours, deflated to year 2000 £s. Zero earnings are included for individuals who are

18 Note that selective schooling systems can affect adult wages in a number of ways including potentially increasing the likelihood of students moving away from home after school to attend university and enter the labour market. This is all part of the effect that we are trying to measure.

Table 3. Proportion of people who move between birth and adulthood from the five largest selective and non-selective LEAs

Selective		Non-Selective	
LEA	Proportion move	LEA	Proportion move
Kent	55.6%	Essex	65.1%
Lancashire	44.3%	Hertfordshire	58.3%
Gloucestershire	41.5%	Bedfordshire	51.2%
Buckinghamshire	66.2%	Leicestershire	29.0%
Warwickshire	64.9%	Northamptonshire	33.8%
Weighted average	53.3%	Weighted average	53.8%

Source: Authors' calculations using Understanding Society waves 1–2.

unemployed or long-term sick or disabled at the time of the survey[19] as these are viewed as valid labour market outcomes. Given two waves of data, each individual has either one or two observations. Rather than discarding information, where we have two wage observations for an individual we average them and include that individual as a single observation. This averaging moves us towards a more permanent rather than transitory measure of individuals' earnings. Sixty-five per cent of the main estimation sample (1,325 of 2,033 individuals) have two wage observations. Prior to the averaging, an initial regression is run to remove any year of survey effects from wages.

We begin by estimating an OLS wage regression (1) where y_{ir} is the average hourly wage of individual i in LEA r, $selective_{ir}$, is a dichotomous variable equal to 1 if the individual was born in a selective LEA and 0 if they were born in a matched non-selective LEA and the regression includes a gender (g) specific quadratic in age (a). This ensures that in our baseline specification we are comparing the earnings of similarly aged males and similarly aged females.

$$y_{ir} = \alpha + \beta selective_{ir} + \gamma a_i * g_i + \delta a_i^2 * g_i + u_{ir} \quad (1)$$

In addition to the effects of age and gender, there are other factors—unrelated to schooling—that may affect current wages. In our second specification (2), the conditional model, we control for personal characteristics (gender, ethnicity, plus the quadratic in age interacted with gender), controls for the individual's parental background (parental occupational class and parental education measured when the individual was 14-years-old[20]), dummies for the current local labour market (county), and dummies for the year of the survey (2009–2012).[21] Given this range of pre-determined control variables we are able to compare the earnings of observably similar individuals from similar family

19 Results are robust to the exclusion of the long-term sick and disabled, see Appendix Figs B2b and B3b and Tables B2b and B3b.

20 Notes to Appendix Table B4 detail the parental occupational categories which are dummies for the 10 major groups of the standard occupational classification (SOC 2000).

21 Appendix Table B4 contains the coefficient estimates for the main estimation sample conditional specification. Robust standard errors are obtained in all regressions, clustering at the individual level in cases where more than one observation is used per person.

backgrounds but who experienced different education systems in areas that were otherwise very similar.

$$y_{ir} = \alpha + \beta selective_{ir} + \gamma a_i * g_i + \delta a_i^2 * g_i + \zeta' X_{ir} + u_{ir} \tag{2}$$

In both specifications, we recover the residuals from our wage regressions and compare the distribution of earnings for those growing up in selective and non-selective systems. As we are interested in the *relative* distributions rather than the effects on the average, we remove the global mean from the residual before calculating the deciles of the distribution—allowing us to compare individuals at different points in each distribution with the overall average.[22] We use unconditional simultaneous quantile regressions (eq. 3), regressing adjusted earnings on the dichotomous selection variable to estimate whether growing up in a selective system has a significant effect on earnings at each decile (d) of the distribution of earnings.[23]

$$Q_d(\hat{y}_{ir}) = \alpha + \beta_d selective_{ir} \text{ where } d = [1, 2, \ldots, 9] \tag{3}$$

Finally, we perform tests on linear combinations at the 90th and 10th percentiles and 75th and 25th percentiles to test whether there are significant differences in the effect of selective systems on earnings inequality.

5. Results

Table 4 shows the raw mean and variance statistics for the selective versus non-selective areas: overall, average hourly earnings 2009–2012 are very similar across the two groups although slightly (insignificantly) higher amongst those from the non-selective areas (£8.76 versus £8.75).[24] The variance of earnings is considerably higher for selective areas (£35.49 versus £26.79). Figure 2 illustrates the impact of selective schooling across the whole distribution, plotting the deciles of age*gender adjusted hourly earnings for each system. As can be seen in this figure, the impact of the selective system has a positive effect on earnings at the top of the distribution and a negative effect on earnings at the lower end of the distribution. For those at the top of the earnings distribution, individuals who grew up in selective schooling areas earn more than their non-selective counterparts. At the bottom of the earnings distribution, this is reversed.

Panel A of Table 5 presents the simultaneous quantile regression estimates corresponding to Fig. 2. These estimates show that the differences between the distributions are statistically significant at the 10th percentile, the 20th percentile, the 30th percentile, the 75th percentile, the 80th percentile and the 90th percentile.

Figure 3, and Panel B of Table 5, present the results using the conditional earnings residuals. The qualitative nature of the results remains largely unchanged: at the lower end of the distribution, individuals born in a selective schooling area earn less than those from the matched non-selective areas, while this reverses for the top deciles. The distributions are significantly different at the 10th percentile, the 20th percentile, the 40th percentile and the

22 As we are removing a constant the results hold for non-mean-adjusted earnings. Note the average earnings are not significantly different across groups indicating a good match.

23 We implement the sqreg command in Stata, which provides bootstrapped standard errors.

24 Appendix Table B5 contains summary statistics for the individuals in the main estimation sample and shows that exogenous characteristics are balanced between the two groups.

Table 4. Raw earnings distribution by schooling system type

	Selective	Non-Selective
Hourly wage: mean	8.75	8.76
variance	35.49	26.79
N	962	1,071

Note: hourly earnings in year 2000 £s.
Source: Authors' calculations using Understanding Society waves 1–2.

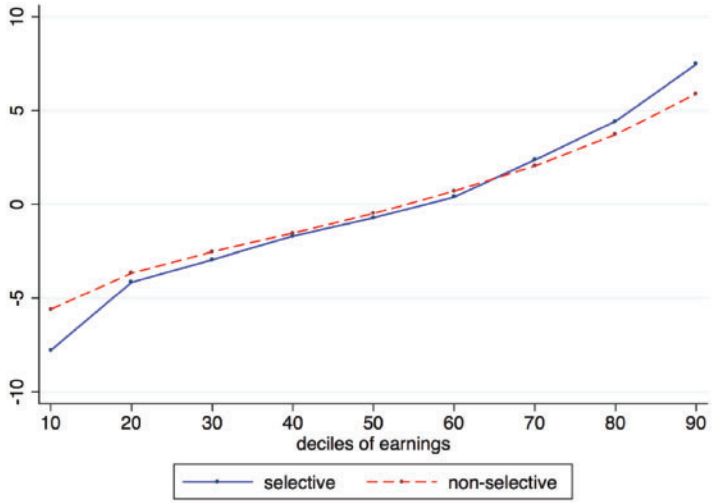

Fig. 2. Deciles of the **raw** earnings distribution by schooling system type.
Notes: residuals from a regression of wage on a gender specific quadratic in age and a selective schooling area dummy, with the global mean of the residual removed. Before averaging wages for individuals with two wage observations, year of survey effects are removed via a regression.

50^{th} percentile. At the top of the distribution there is a statistically significant positive effect of selective schooling at the 90^{th} percentile.

These results are robust to including all observations (i.e. not averaging where an individual has two observations), to including just a single observation per individual, to including only the observations of individuals with two observations, and to altering the definition of selective and non-selective areas—in each case the pattern and levels of significance remain essentially unchanged.[25]

Table 6 presents estimates of the difference in the effect sizes found at the 90^{th} and 10^{th} percentile and 75^{th} and 25^{th} percentiles for both the unconditional (Panel A) and conditional (Panel B) models. Focusing first on Panel A, the 90–10 earnings gap of individuals growing up in a selective LEA is £3.77/hour larger than the 90–10 earnings gap of individuals from a non-selective system. This accounts for 24.1% of the overall 90–10 earnings gap in our sample, and the test against zero has a p-value below 0.001. Focusing on the 75^{th}–25^{th}

25 See Appendix Figs B2a, B2b, B3a and B3b plus Tables B2a, B2b, B3a and B3b.

Table 5. Quantile regression estimates of selective schooling effect on wages

	A: Without controls				B: With controls		
	coeff.	*std. error*	*t*		*coeff.*	*std. error*	*t*
10	−2.195	0.559	−3.93***	10	−1.019	0.399	−2.56**
20	−0.465	0.268	−1.73*	20	−0.592	0.335	−1.77*
25	−0.308	0.233	−1.32	25	−0.284	0.259	−1.09
30	−0.415	0.216	−1.92*	30	−0.300	0.247	−1.21
40	−0.173	0.204	−0.85	40	−0.482	0.258	−1.87*
50	−0.227	0.151	−1.50	50	−0.494	0.259	−1.91*
60	−0.317	0.287	−1.11	60	−0.153	0.308	−0.50
70	0.325	0.369	0.88	70	0.052	0.310	0.17
75	0.699	0.394	1.77*	75	0.103	0.361	0.29
80	0.705	0.427	1.65*	80	0.400	0.355	1.13
90	1.574	0.637	2.47**	90	1.312	0.559	2.35**
	N=2,033				N=2,033		

Note: residuals from a regression of wage on a gender specific quadratic in age and a selective schooling area dummy (Panel A); and residuals from a regression of wage on a gender specific quadratic in age, a selective schooling area dummy, gender, ethnicity, parental occupational class when the individual was 14, parental education and current county of residence (Panel B). Global means of the residual removed. Before averaging wages for individuals with two wage observations the year of survey effects are removed via a regression. ***,**, * indicate significance at the 1%, 5%, and 10% level, respectively.
Source: Authors' estimates using Understanding Society waves 1–2.

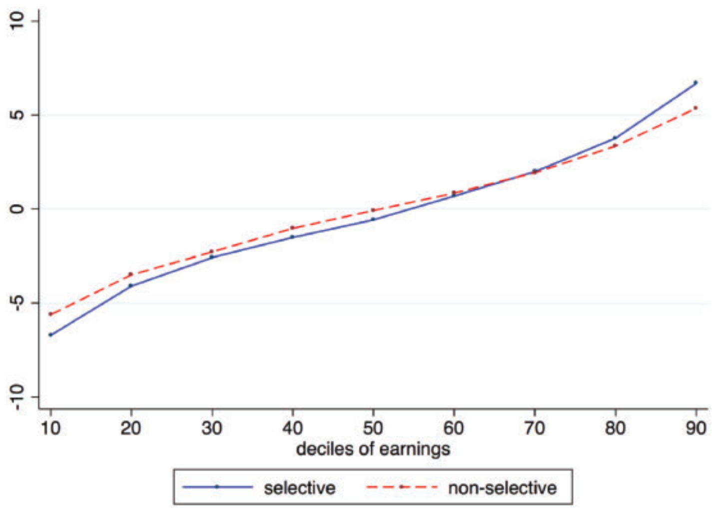

Fig. 3. Deciles of the **conditional** earnings distribution by schooling system type.
Notes: residuals from a regression of wage on a gender specific quadratic in age, gender, ethnicity, parental occupational class when the individual was 14, parental education, current county of residence and a selective schooling area dummy with the global mean of the residual removed. Before averaging wages for individuals with two wage observations, year of survey effects are removed via a regression.

Table 6. Estimated effects sizes

	Sample wage gap	A: Without controls			Sample wage gap	B: With controls		
		coeff.	*std. error*	*Effect size*		*coeff.*	*std. error*	*Effect size*
90–10	15.66	3.768	0.808***	24.06	12.07	2.331	0.615***	19.31
75–25	6.45	1.007	0.390***	15.61	5.77	0.387	0.354	6.71

Note: earnings differentials estimated by testing the linear combination from the simultaneous quantile regressions. The effect size is calculated as the estimated difference divided by the total earnings differential in the sample. ***,**, * indicate significance at the 1%, 5%, and 10% level, respectively.
Source: Authors' estimates using Understanding Society waves 1–2.

percentile earnings difference, the gap is 15.6% of the total raw gap, with a *p*-value of 0.010 for the test against zero.

Panel B shows that in the conditional model, there is a quantitatively and statistically significant difference in the 90–10 earnings gap between the two education systems. This is £2.33/hour, or 19.3% of the total conditional 90–10 gap in the sample, with a *p*-value of below 0.001. However, the difference at the 75th–25th percentiles is smaller and no longer significantly different.

5.1 Differences by gender

While there is no *a priori* reason to think that schooling systems will have differential effects on inequality by gender according to our descriptive framework, it is interesting to consider this question for males and females separately. Tables 7 and 8 and Figs 4 and 5 present the results by gender (showing the conditional model only, for each gender the unconditional model results follow the same pattern and significance of the pooled results). Table 8 shows that the differences in inequality for both males and females in the conditional model also mirror those seen in the pooled sample though somewhat smaller (18.3% of total 90–10 gap for males and 10.6% of total 90–10 gap for females). However, the detail in Table 7 and the figures show a slightly more complex picture: for males, the difference is concentrated at the top of the distribution, whereas for females, the gap is particularly evident at the bottom of the distribution in the conditional specification. It may well be that this is because there was a significant gender difference in school assignment in selective areas. That is, the grammar school era was a time when boys typically outperformed girls at school, and being in a selective area meant that female students disproportionately went to secondary modern schools and male students disproportionately went to grammar schools.

5.2 Robustness

To test whether our results are robust to changes in the definition of selective and non-selective areas we redefine selective LEAs as those assigning more than 30% of places by selection whilst retaining the definition of non-selective as those that assign less than 5% by this method. Appendix Tables B2a and B3a shows the quantile regressions for the models with and without controls. The results are qualitatively and quantitatively similar to the corresponding figures in Table 5 (the 90–10 gap in the conditional results is £2.50 which is 19.4% of the conditional 90–10 difference). Figure 6 illustrates the results of the model with controls and comparison with Fig. 3 provides visual confirmation of the robustness of

Table 7. Quantile regression estimates of selective schooling effect on wages, by gender

	A: Males				B: Females		
	coeff.	*std. error*	*t*		*coeff.*	*std. error*	*t*
10	−0.057	0.568	−0.10	10	−1.069	0.552	−1.94*
20	−0.212	0.555	−0.38	20	−1.309	0.339	−3.86***
25	0.126	0.460	0.27	25	−1.192	0.315	−3.78***
30	0.311	0.410	0.76	30	−1.190	0.267	−4.46***
40	0.493	0.402	1.23	40	−0.952	0.305	−3.12***
50	0.706	0.382	1.85*	50	−0.994	0.218	−4.56***
60	0.797	0.399	2.00**	60	−1.059	0.289	−3.67***
70	1.189	0.401	2.97***	70	−1.066	0.32	−3.34***
75	1.131	0.581	1.95*	75	−1.032	0.391	−2.64***
80	1.594	0.476	3.35***	80	−0.608	0.478	−1.27
90	2.386	0.917	2.60***	90	0.111	0.524	0.210
	912				1,121		

Note: residuals from a regression of wage on a gender specific quadratic in age, a selective schooling area dummy, ethnicity, parental occupational class when the individual was 14, parental education and current county of residence. Men only (Panel A) and Women only (Panel B). Global means of the residual removed. Before averaging wages for individuals with two wage observations the year of survey effects are removed via a regression. ***,**, * indicate significance at the 1%, 5%, and 10% level, respectively.
Source: Authors' estimates using Understanding Society waves 1–2.

Table 8. Estimated effects sizes for conditional specification by gender

	Sample wage gap	A: Males			Sample wage gap	B: Females		
		coeff.	*std. error*	*Effect size*		*coeff.*	*std. error*	*Effect size*
90-10	13.37	2.442	1.030**	18.27	11.09	1.180	0.722*	10.64
75-25	6.02	1.005	0.566*	16.74	5.14	0.160	0.441	3.10

Note: earnings differentials estimated by testing the linear combination from the simultaneous quantile regressions. The effect size is calculated as the estimated difference divided by the total earnings differential in the sample. ***,**, * indicate significance at the 1%, 5%, and 10% level, respectively.
Source: Authors' estimates using Understanding Society waves 1–2.

the results. Further robustness tests are illustrated in Figs B2a and B2b (both for the conditional model) and B3a and B3b (raw model), in which the selectivity definition, the inclusion/exclusion of private school percentage and other time-invariant variables from the matching, the treatment of multiple observations for an individual, the treatment of long-term sick/disabled and the restriction of sample years are all tested. The figures all illustrate the same pattern of results and the corresponding quantile regression results in Tables B2a, B2b, B3a and B3b confirm the robustness of our results.

As one further robustness test, we take our main estimation sample and re-assign at random the indicator for growing up in an area operating the selective system, though maintaining the same proportion of the sample in each system. We then carry out the same exercise in comparing the distribution on earnings for individuals 'growing up' in each

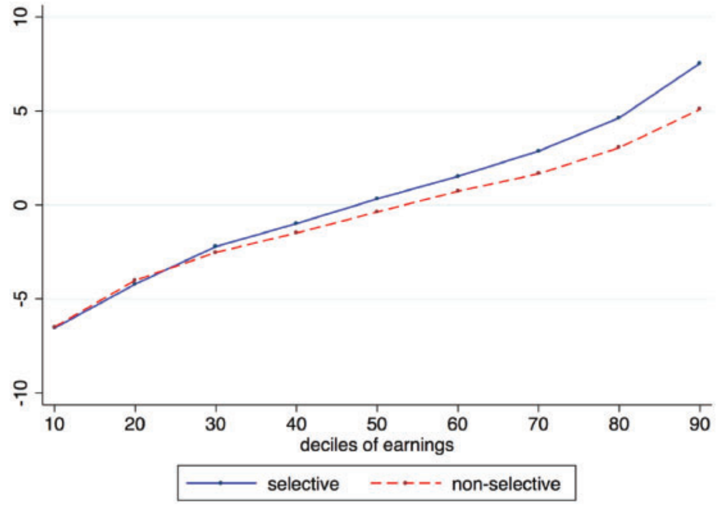

Fig. 4. Deciles of the **conditional** earnings distribution by schooling system type, males.
Notes: residuals from a regression of wage on a gender specific quadratic in age, ethnicity, parental occupational class when the individual was 14, parental education, current county of residence and a selective schooling area dummy with the global mean of the residual removed. Before averaging wages for individuals with two wage observations, year of survey effects are removed via a regression.

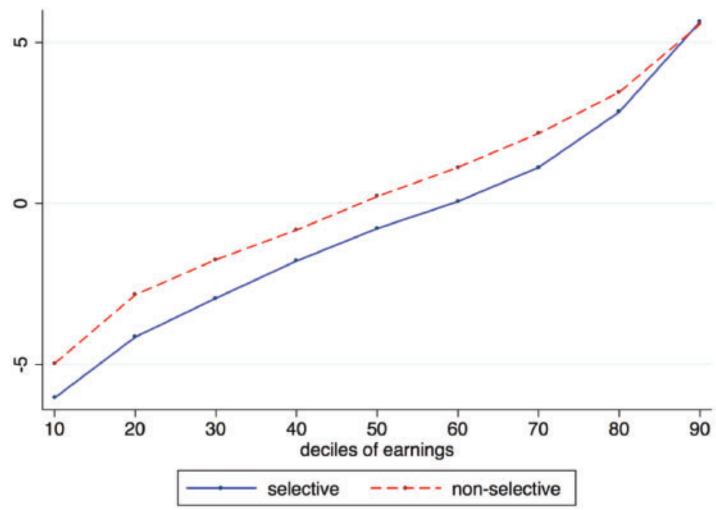

Fig. 5. Deciles of the **conditional** earnings distribution by schooling system type, females.
Notes: residuals from a regression of wage on a gender specific quadratic in age, ethnicity, parental occupational class when the individual was 14, parental education, current county of residence and a selective schooling area dummy with the global mean of the residual removed. Before averaging wages for individuals with two wage observations, year of survey effects are removed via a regression.

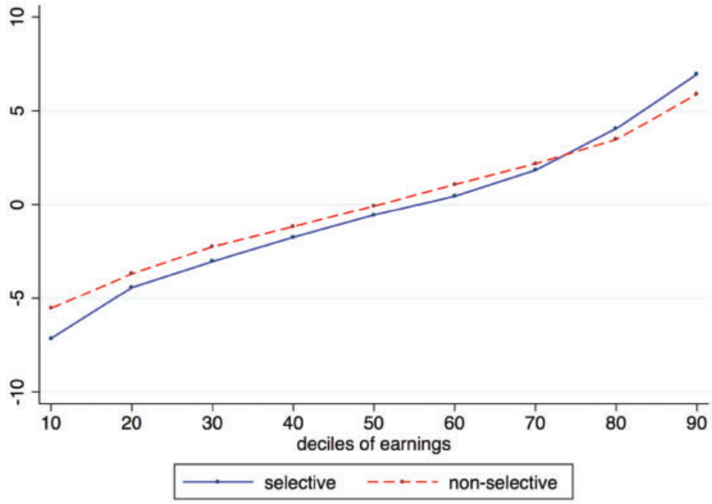

Fig. 6. Deciles of the **conditional** earnings distribution by schooling system type, selective defined as >30% assigned by selection, non-selective <5% assigned by selection.

Notes: residuals from a regression of wage on a gender specific quadratic in age, gender, ethnicity, parental occupational class when the individual was 14, parental education, current county of residence and a selective schooling area dummy with the global mean of the residual removed. Before averaging wages for individuals with two wage observations, year of survey effects are removed via a regression.

system. As can be seen visually in Appendix B Fig. B1, whether the conditional or raw models are used, there is now no significant difference between the distributions at any point—in contrast to our main estimates, which are reproduced to allow comparison. Similarly, Appendix B Table B1 shows the corresponding coefficient estimates, none of which are statistically significant.

6. Conclusion

In this article, we have investigated the impact on earnings inequality of a selective education system in which school assignment is based on initial test scores. In England, this was the system in place until the 1970s, when the comprehensive system started to become the norm. Despite this ever-receding historical background, the role of grammar schools continues to be a lively and contentious issue in the current education policy debate, with the UK government recently consulting on reintroducing or expanding existing grammar schools. This finds a parallel in the US where a similar literature concerns the merits of exam schools.

We have used a large and representative household panel survey with information on each respondent's childhood to compare adult earnings inequality of those growing up under a selective education system with those educated under a comprehensive system. Given the inevitable time-lag between schooling and prime-age earnings, assembling the data required to match local areas on their characteristics at the time of schooling is a non-trivial task. Nevertheless, we believe that we have constructed the best possible data set available to address this question for England and the robustness of the results to different sets of matching covariates and the covariate balance we achieve between the selective and non-selective areas gives confidence that we are making a valid comparison.

Our data allows us to add new information into the selective schooling debate in Britain and beyond. Controlling for a range of background characteristics and their current labour market context, we find that the wage distribution for individuals who grew up in selective schooling areas is substantially more unequal, with this difference being statistically significant. The total effect sizes are large: 24% of the raw 90–10 earnings gap and 19% of the conditional 90–10 earnings gap can be explained by schooling system. These results are robust to a wide range of specification checks. In terms of magnitude, we can compare the proportion of the 90–10 earnings gap that is explained by the schooling system, to the general increase in wage inequality in the first part of this century. Between 1997 and 2009, the 90–10 earnings gap increased by almost 60%, so our results are equivalent to around five years' worth of change.[26] Given this context, it is clear that the schooling system is responsible for a sizeable share of wage inequality.

These results are consistent with previous findings by Atkinson *et al.* (2006) who find that selective systems lead to wider disparities in educational attainment compared to comprehensive systems: grammar-educated children in selective LEAs outperform similar children in non-selective LEAs on average while non-grammar-educated children in selective LEAs *underperform* compared to similar children in non-selective LEAs. Our findings resonate with this, suggesting that these exam performance disparities carry through into the labour market and are reflected in subsequent earnings. This is one potential mechanism for our findings. Our modelling framework highlighted the role of other potential mechanisms including peer groups and school (teacher) quality in magnifying inequality in ability in a selective education system. The evidence on peer effects is mixed, whereas the UK evidence on the wide variation of teacher effectiveness mirrors that in the US (Slater *et al.*, 2012). It seems likely therefore that the main mechanism generating greater inequality is the sorting of the more effective teachers to the highest ability students. Unfortunately, there is no historical data available to test this, and a comparison of the few contemporary grammar schools in England may not be that relevant to this study.

We have shown that cohorts of students growing up in areas with a selective education system experience greater earnings inequality once in the labour market. If higher earnings inequality is coupled with socially graded access to grammar schools then it seems likely that selective systems will also reinforce inequalities across generations. Previous analysis of the impact of the grammar system as a whole, by Atkinson *et al.* (2006), suggested that this is indeed the case. They found that access to grammar schools is heavily skewed towards more affluent families even after taking into consideration prior attainment. High ability children from low-income families are approximately half as likely to go to grammar school as children of comparable ability from better off families (32% versus 60%). This finding has been replicated more recently by Cribb *et al.* (2014), who find 66% of high achieving children from better-off families attend grammar schools, while only 40% of similarly able but lower income children do. Importantly for the debate around the merits of grammar schools for promoting social mobility, these findings, taken with our findings, imply that

26 Authors' calculations using the Annual Survey of Hours and Earnings, 2013. Percentile values of hourly earnings for 1997–2013 available in Table 5 of 'ASHE 1997 to 2014 selected estimates (Excel sheet 408Kb)' see: http://www.ons.gov.uk/ons/rel/ashe/annual-survey-of-hours-and-earnings/2013-revised-results/index.html, last accessed 10 July 2018.

while children from poorer families who make it to the grammar school do well—in exam and earnings terms—those who grow up in selective areas but *do not attend* grammar schools do worse than they would have done, had they not been schooled in a selective area. Given the disproportionate numbers of low-income families whose children *do not* make it to the grammar school, the implication is that the selective system is in fact regressive with respect to social mobility.

There may, in theory, be positive effects of earnings inequality that may off-set some of the negative effects identified here—a social welfare function may accept higher earnings inequality as a cost worth paying for generally higher earnings, for example. However, looking at the mean earnings of those growing up in selective versus non-selective areas (which are almost identical) there is no evidence of higher inequality being accompanied by increased average earnings in this case. Moreover, though it is mostly only for the lower percentiles of the distribution that the negative effect of the selective schooling system is statistically significant, the point estimates are generally negative throughout the distribution up to the 60th or 70th percentile. Thus these findings suggest that the benefits of earnings inequality are narrowly focused on those at the very top of the distribution. Setting up a formal model to weigh the positive and negative effects of earnings inequality is beyond the ambition of this article. Our contribution is to add a new fact to the debate on grammar schools: selective schooling systems do appear to increase inequality.

Supplementary material

Supplementary material is available on the OUP website. The supplementary material comprises the replication files and the Online Appendix. The Appendix includes information on the institutional background of the schooling system in the UK along with additional robustness tests, sample statistics and matching balancing tests.

The main data used in the analysis, from Understanding Society, is not provided as we used a special licence extract: Understanding Society: Waves 1–2, 2009–2011: Special Licence Access, study number 6931. This data can be applied for via the UK Data Service, for further details of the access procedure see https://beta.ukdataservice.ac.uk/datacata logue/studies/study? id=6931#!/access.

The selectivity of LEAs was defined using Annual Schools Census data 1974–1983, aggregated from school level to LEA level. The Annual Schools Census data is freely available here: http://discovery.nationalarchives.gov.uk/details/r/C7066.

These ASC files contain information on the number of pupils in each school, the type of school, and the LEA of the school. This allows creation of the LEA level indicator of the percentage of children assigned their school place by selection in each LEA in each year. The authors would be available to assist further with this process should this be required. A link file to match LEA code numbers in the ASC files to the area of birth codes in the Understanding Society data is contained in the supplementary materials. Once created, the LEA level selectivity indicator can be merged into the Understanding Society data using the code contained in the replication file.

Bespoke data on political control of counties was provided to the authors by the elections centre: http://www.electionscentre.co.uk. The data used is uploaded in the supplementary materials.

All other data sets used in the analysis are provided in the supplementary materials.

Funding

This work was supported by the Economic and Social Research Council [grant number: ES/H005331/1].

Acknowledgements

The authors gratefully acknowledge the helpful comments of the editor and two anonymous referees whose insights were invaluable in revising the article. Many thanks to Damon Clark for compiling the data on selectivity of local education authorities, and to Colin Rallings for providing data on area-level political control. Thanks also to Stephen Jenkins, John Hills and John Micklewright for their comments and to seminar participants at the Bristol-Bath applied micro-workshop, the IOE Department of Quantitative Social Science, the University of Sydney, the CASE Research Workshop and the EALE annual conference 2014 (Ljubljana), Social Policy Association Annual Conference 2017 (Durham).

References

Abdulkadiroglu, A., Angrist, J., and Pathak, P. (2012) The elite illusion: achievement effects at Boston and New York exam schools. IZA Discussion Paper, No. 6790, Bonn.

Allen, R., Burgess, S., and Key, T. (2010) Choosing secondary schools by moving house: school quality and the formation of neighbourhoods, University of Bristol, CMPO Working Paper, No. 10/238.

Atkinson, A., Gregg, P., and McConnell, B. (2006) The result of 11Plus selection: an investigation into equity and efficiency of outcomes for pupils in selective LEAs, University of Bristol, CMPO Working Paper, No. 06/150.

Bolton, P. (2017) Grammar school statistics, House of Commons Library, Briefing Paper, No. 1398, London.

Burgess, S., Dickson, M., and Macmillan, L. (2014) Selective schooling systems increase inequality, IZA Discussion Paper, No. 8505, Bonn.

Clark, D. (2010) Selective schools and academic achievement. *B.E. Journal of Economic Analysis and Policy*, **10**, 1935–682.

Clark, D. and Del Bono, E. (2014) The long-run effects of attending and elite school: evidence from the UK, ISER, University of Essex, Working Paper, No. 2014–05.

Cribb, J., Sibieta, L., and Vignoles, A. (2014) Entry into grammar schools in England, in J. Cribb, D. Jesson, L. Sibieta, A. Skipp, and A. Vignoles (eds) *Poor Grammar: Entry into Grammar Schools for Disadvantage Pupils in England*, Sutton Trust Report, London.

Crook, D. (2013) Politics, politicians and English comprehensive schools, *History of Education: Journal of the History of Education Society*, **42**, 365–80.

Dobbie, W. and Fryer, R. (2011) Exam high schools and academic achievement: evidence from New York City, National Bureau of Economic Research, Working Paper, No. 17286, Cambridge, MA.

Dustmann, C., Puhani, P., and Schonberg, U. (2014) The long-term effects of early track choice. IZA Discussion Paper, No. 7897, Bonn.

Galindo-Rueda, F. and Vignoles, A. (2005) The heterogeneous effect of selection in secondary schools: understanding the changing role of ability, Centre for Economics of Education, London School of Economics, Working Paper, No. 0052.

Gibbons, S., Machin, S., and Silva, O. (2013) Valuing school quality using boundary discontinuities. *Journal of Urban Economics*, **75**, 15–28.

Guyon, N., Maurin, E., and McNally, S. (2012) The effect of tracking students by ability into different schools—a natural experiment. *Journal of Human Resources*, **47**, 684–721.

Jesson, D. (2000) The comparative evaluation of GCSE value-added performance by type of school and LEA, Economics Dept., University of York, Discussion Paper, No. 2000/52.

Manning, A. and Pischke, J.S. (2006) Comprehensive versus selective schooling in England and Wales: what do we know?, National Bureau of Economic Research, Working Paper, No. 12176, Cambridge, MA.

Ryan, C. and Sibieta, L. (2010) Private schooling in the UK and Australia, Briefing Note BN 106, Institute for Fiscal Studies, London.

Slater, H., Davies, N., and Burgess, S. (2012) Do teachers matter? Measuring the variation in teacher effectiveness in England. *Oxford Bulletin of Economics and Statistics*, **74**, 629–45.

Sullivan, A. and Heath, A. (2002) State and private schools in England and Wales, University of Oxford, Sociology Working Paper, No. 2002–02.