serie ad

WP-AD 2017-05

Can Talented Pupils with Low Socio-economic Status Shine? Evidence from a Boarding School

Francesca Foliano, Francis Green and Marcello Sartarelli

Ivie

Working papers

Working papers

Working papers

Versión: octubre 2017 / Version: October 2017

Edita / Published by:
Instituto Valenciano de Investigaciones Económicas, S.A.
C/ Guardia Civil, 22 esc. 2 1º - 46020 Valencia (Spain)

DOI: http://dx.medra.org/10.12842/WPAD-2017-05

WP-AD 2017-05

Can Talented Pupils with Low Socio-economic Status Shine? Evidence from a Boarding School*

Francesca Foliano, Francis Green and Marcello Sartarelli**

Abstract

We study whether substituting family with school inputs for high ability pupils with low socio-economic status (SES) has an impact on achievement in the compulsory school final exams using administrative data on England. By considering a selective secondary, well-resourced boarding school admitting an unusually high share of talented pupils with low SES, we estimate the effect with propensity score matching to obtain comparable control groups in selective day schools. Our main finding is that the probability of being in the top decile of achievement in the exams increases by about 17 percentage points compared to 59% for controls, i.e. by 29%.

Keywords: ability, achievement gap, boarding, education, grammar school, GCSE, private school, socio-economic status, SES.

JEL classification numbers: I21, I24, I32, J62.

* We thank Pedro Albarran, Marianna Battaglia, Richard Dorsett, Amanda Gosling, Iñigo Iturbe, Steve Machin, Giuseppe Migali, Olmo Silva, Tanya Wilson and participants at the Universidad de Alicante Applied Seminar, University of Kent Research Seminar, AEDE 2014, SIEP 2016 and RES 2017 for their useful comments. We also thank the management of Christ Hospital School for kindly providing us with information about the school's admissions criteria and procedures. The responsibility for the findings in this paper are entirely our own. Financial support from the ESRC (Grant No. ESJ0191351), Spanish Ministerio de Economía y Competitividad (ECO2016-77200-P, ECO2015-65820-P and ECO2013-43119) and Universidad de Alicante (GRE 13-04) is gratefully acknowledged.

** F. Foliano: University of Kent and NIESR. F. Green: UCL Institute of Education. M. Sartarelli: Universidad de Alicante, e-mail: marcellosartarelli@gmail.com (corresponding autor).

1 Introduction

Achievement gaps by socio-economic status (SES) are a policy challenge worldwide (see for a review Sirin, 2005; Reardon, 2011). In England they are observed as early as primary education and are either constant or tend to increase with age (Dearden *et al.*, 2011). While gaps by SES tend to be concentrated among pupils who are initially low achievers, they have also been recently observed among high achievers (Crawford *et al.*, 2014; Jerrim, 2017). This may have potentially high opportunity costs if pupils who seem to have the potential to perform well at school, which is a good predictor of success in the labour market in adulthood, are held back or slowed down by low SES.

Most of the policies that have been designed to counteract the low SES negative influence on pupils' achievement compensate poor family inputs with better school inputs. Their impact on children's achievement is studied theoretically and empirically by letting the family and the school be inputs in the child's educational production function (Todd and Wolpin, 2003, 2007). However, the effect of a school policy may be confounded by family inputs, that are not easily observed. Family inputs might reinforce the role of school policies if parents respond by putting more effort into their children's development and vice versa dilute it if they put less.[1]

Boarding schools offer, instead, a suitable example of substantial substitution of family inputs for school inputs, i.e. they reduce the role of family inputs for all pupils, since they offer education during the day and lodging at night in weekdays. However, obtaining clean estimates of the effect of attending a boarding school on, for example, pupils' achievement is an empirical challenge as it may be confounded by a selection effect if boarding schools pupils and pupils in other schools differ substantially in ability or in SES.

Lotteries granting random admission to oversubscribed boarding schools have been used to obtain clean estimates of the effect of attending a boarding school. Randomly admitted pupils obtain substantially higher tests scores than non-admitted ones in boarding schools in poor neighbourhoods in the US (Curto and Fryer Jr,

[1] Recent examples of school policies are the introduction of sponsored Academy schools in disadvantaged areas in the UK (Eyles *et al.*, 2016) and of Charter schools in the US (see a review in Epple *et al.*, 2016), as well as more narrowly targeted interventions in urban schools in the UK, such as Excellence in the Cities (Machin *et al.*, 2004). Exceptionally, the Assisted Places Scheme, which ran in the UK from 1981 to 1997 with mixed success, involved some boarding schools, though only a minority of scheme participants were from disadvantaged backgrounds (Whitty and Edwards, 1998).

2014). Related research also exploiting random admission in an elite school in France obtains similar results (Behaghel *et al.*, 2017). Lotteries offer the advantage that admitted pupils have similar characteristics to those not admitted. However, the estimated effect may be biased upward since oversubscribed schools, the only ones in which lotteries are run, may be on higher demand than other schools because they are of higher quality (Eyles and Machin, 2015). In addition, the low number of observations limits the possibility to study the impact of boarding for high achievers with low SES.

This paper is the first to test the hypothesis that offering high achievers with low SES admission at a truly selective boarding school, whose name is Christ Hospital (CH hereafter), leads to higher achievement (H_1). We use rich administrative data on England and measure achievement at age 7, 11 (Key Stage 1 and 2) and in the compulsory school final exams at age 16 (General Certificate of Secondary Education, GCSE hereafter) for five consecutive cohorts of pupils. The aim of our research design is to find for each pupil at CH a pupil in a selective day school who is as similar as possible in observable characteristics by using propensity score matching. We use the following two measures of SES: the income deprivation affecting children index (IDACI), measuring the share of children in low income households by local area, and, in addition, whether pupils' parents obtain income support from the government, proxied by whether pupils are eligible for Free School Meals (FSM).

We also test whether there is heterogeneity in the impact of CH on achievement as a related hypothesis (H_2). We study whether it differs for pupils with very low SES, proxied by an IDACI greater than the median value, to assess if they are the ones who benefit most from CH as their learning environment is the one improving the most upon admission. In addition, we examine the effects of CH separately for girls and boys given that the home environment's contribution to academic achievement may be gendered and in light of the widely documented achievement gap in favour of girls in England (Machin and McNally, 2005).

CH is an outlier in English private education, in that it educates a very high share of high achievers with low SES. Its especially wealthy foundation enables it to admit the majority of its pupils with low or no fees through means-tested bursaries. Since it is selective and boarding, we compare its pupils with those in the following two types of selective day schools, i.e. our *control groups*. The first are pupils in grammar schools, that are highly selective state schools with substantially fewer resources than CH. The second group are pupils in independent schools, that are also well-resourced like CH, although they tend to be less selective based on academic merit. In addition, in our control groups we only selected pupils from primary schools located in the same local authority as those attended by pupils who then went to CH. This ensures that the school and non-school environment that a pupil at CH and a very similar pupil in a selective day school experienced before secondary school is also comparable.

We find that pupils' achievement at CH is significantly higher than for control pupils (H_1). The probability of at least five GCSE exams (GCSEs hereafter) at A-A*, i.e. of being in the top decile in the distribution of the number of GCSEs at A-A*, is significantly higher at CH by 17.4 and 12.6 percentage points, i.e. a 20-29% increase with respect to 59-64% for matched pupils in grammar and independent day schools respectively. When we assess whether there is heterogeneity in our main results (H_2), we find that they tend to be driven by higher point estimates for pupils from poorer areas, proxied by an IDACI level above the median, and for girls. Crucially, predetermined characteristics for pupils at CH and for controls in grammar and independent schools are balanced and, in addition, our main results are robust to a set of sensitivity checks.

To the best of our knowledge, our results are the first in showing that a boarding school admitting high ability and low SES pupils significantly improves their achievement (H_1). They contribute to the school choice literature by suggesting that currently available alternatives to standard schooling options may play a relevant role in reversing the achievement gap for these pupils. Finding that the impact is higher for poor pupils (H_2) also contributes to studies on children's educational production function by suggesting that the substitutability between the inputs that these pupils obtain from parents and from the school is not substantially diluted by economic, cognitive or psychological disadvantages associated with low SES. Our related finding on the greater effect for girls is also relevant for an understanding of the gender achievement gap for high ability pupils in England. Finally, our results contribute to the recent policy debate over the use of boarding schools for disadvantaged children in England (Department for Education, 2014, 2016) that so far lacks a thorough quantitative economic analysis.

The rest of the paper is structured as follows. Section 2 reviews the relevant literature. Section 3 describes the institutional setting of compulsory education in England and the data that we use in the empirical analysis. Section 4 outlines the econometric strategy. Section 5 describes the main results, section 6 reports the results of a sensitivity analysis and, finally, section 7 discusses them and concludes.

2 Literature review

In this section, we consider what existing studies around the world reveal about the effects of boarding on academic and non-academic outcomes. In the US, in recent years, boarding secondary schools for bright pupils with low SES have been introduced. SEED schools in Washington and Baltimore are the only urban public schools that combine the charter school model with a 5-day-a-week boarding program in poor neighbourhoods. Curto and Fryer Jr (2014) estimate the impact of attending a SEED school in Washington on achievement by exploiting lottery-driven admission that is used when a school is oversubscribed. By comparing achievement of students admitted and of those turned down by the lottery, they find that SEED increases achievement by about 20% of a standard deviation in reading and in math, with results being mainly driven by females.

Similarly, in France public 'boarding schools of excellence' for poor and high achieving pupils have been opened in deprived suburbs of large French cities. Behaghel *et al.* (2017) exploit an admission lottery to study the effect of attending one such school in the suburbs of Paris. They find that by the end of the first year, achievement in French and in maths is lower (60-270%) although differences are not significant. A subjective measure of well-being, obtained by way of a survey, is also weakly significantly lower (183%) and it is driven by frictions in adapting to the boarding environment. In contrast, after the second year, maths scores are significantly higher (ten times) while they are lower in French (five times) although the difference is not significant. Well-being is also significantly higher than in the first year (18%), driven by significantly higher scores in the question on whether children feel at home. As for the improvement in maths, it is driven by those students who were in the top three deciles of the distribution of maths scores when they enrolled.

Curto and Fryer Jr (2014) and Behaghel *et al.* (2017) are the only studies that, to the best of our knowledge, quantify the effect of boarding school on achievement for low SES pupils by exploiting admission lotteries. While it is argued that clean estimates are obtained by using this quasi-experimental setting, the main drawbacks of lotteries are at least two. The first is that since oversubscribed schools are in higher demand than others that are not oversubscribed, their quality is higher, e.g. they may have more resources and more motivated or more qualified teachers, and since quality tends to be unobserved, estimates of the boarding school effect obtained by exploiting lotteries may be upward biased. The second is that the number of observations tends to be small, e.g. about 400 in total in each of these studies, as only very few schools are oversubscribed. This number may not be big enough, for example, to estimate a parameter of interest separately for a subgroup of pupils by ability and SES.

Excellence in the Cities is an example of a policy intervention to improve school inputs in urban day schools in poor neighbourhoods in England, that targets talented pupils as part of its third core strand. Its main effect is an increase in maths achievement (2.5-5% relative to the mean value for children in control schools), with this result being driven by children in disadvantaged schools and particularly by those with ability above average in these schools (Machin *et al.*, 2004).[2]

Related policies have focused on giving schools greater autonomy over, for example, hiring teachers and using those teaching methods that are most suitable to children's learning needs in neighbourhoods with different socio-demographic characteristics. In the early 1990's Charter Schools were introduced in the US 'as laboratories for educational innovation' (Epple *et al.*, 2016). They led to a highly significant increase mainly in math (on average 5-10% relative to children in other schools or 25-40% of a standard deviation, s.d. hereafter) while the improvement in English (3-6% or 20% s.d.) is smaller, less precise and not always significant (see for a review Epple *et al.*, 2016). Similarly, in the early 2000's a new type of school, called Academy, was introduced in the UK to improve standards of low performing schools, which are more frequently located in poorer neighbourhoods. Overall, these schools led to a significant increase in achievement in the compulsory school final exams (7% s.d. in GCSE points score) and in the probability of degree completion (10% relative to mean value), driven by children with low SES (see for a review Eyles and Machin, 2015; Eyles *et al.*, 2016).

The main assumption behind these interventions for children with low SES is that the potentially negative impact of the home environment can be offset either by offering children better school inputs in the case of school-based policies or by substituting family inputs with a better learning environment, in the case of boarding

[2] Kirabo Jackson (2010) exploits exogenous variation arising from secondary school choice conditional on merit-based rules assigning pupils to schools in Trinidad and Tobago to study the effect of attending better schools, i.e. those with a pool of higher ability pupils and with better resources. Results show that attending a better school significantly increases exams performance at the end of secondary school, with the effect being larger for girls.

schools, such as CH. If cognitive outcomes are the results of an education production function as the one described in Todd and Wolpin (2003, 2007), then introducing a better school input, such as a more learning-oriented environment, would result in more desirable outcomes for children. However, the effect of better school inputs may confounded by parental responses to it as simultaneous changes in family inputs are not fully accounted for.

Boarding schools have also been studied in psychological research with a focus on their consequences on pupils' well-being. Results obtained either using simple differences in means or OLS regressions are mixed. Lester *et al.* (2015) show using data on 150 boarders in the US that they experience significantly higher bullying when their boarding experience starts. Wires *et al.* (1994) show using data on 197 boarders in the US that the development of their identity improves with age although it is lower for those with adolescent behavioural problems. Fisher *et al.* (1986) show using data on 115 boarders in the UK that their initially high level of homesickness decreases over time.

In partial contrast, Martin *et al.* (2014) find no significant differences in boarders' subjective well-being or in their academic achievement using data on 2,002 school children in Australia, 30% of whom are boarders. In related work, Hodges *et al.* (2016) show using survey data on 415 boarders in Australia that they perceive the boarding environment more negatively than the school staff does, in line with research on children in day schools.[3]

Recent studies on boarding in rural China, whose institutions differ markedly from the Anglosaxon ones in which boarding schools were first created, also show mixed results. Shu and Tong (2015) use about 2,000 observations from the 2010 wave of the survey of Chinese households and, by using propensity score matching and school fixed effects regressions, find that boarders do significantly better in tests although they also have higher depression scores. In related research Wang *et al.* (2016) use data from surveys they conducted on about 5,000 children in the period 2008-2013 to study the effect of a boarding program introduced in 2001 and find that boarders' academic achievement is, instead, significantly lower.

Overall, the evidence summarised in this section from studies in different disciplines in social science shows that pupils in boarding schools tend to have higher

[3] Schaverien (2004, 2011) present qualitative evidence on the interplay between emotional deprivation and social success. However, it is based on 2-3 case studies per study, i.e very few observations to draw any conclusion based on them.

achievement and that this seems to be driven by higher motivation and study effort. Although this evidence expands our knowledge on the role played by boarding schools, to the best of our knowledge no study has tested whether a conducive boarding environment can compensate, through substitution of family and school inputs, for high ability children with a low SES.

3 Institutions and data

In this section we describe the institutional setting of compulsory education in England, along with the main characteristics of CH, grammar and independent schools in section 3.1. We then describe the data we use in the empirical analysis in section 3.2.

3.1 Institutional setting

The state school system in England entails 11 years of compulsory education divided into two phases, primary and secondary, and 4 Key Stages, summarised in Table 1. Primary school starts with Key Stage 1 (age 5 to 7) and it is followed by Key Stage 2 (age 7-11), whereas secondary school starts with Key Stage 3 (age 11 to 14) followed by Key Stage 4 (age 15-16). All Key Stages end with a national standardised assessment, either carried out by teachers or externally and the Department for Education set a level or target that pupils are expected to achieve in compulsory tests in English, Math and Science. Pupils' tests are marked using an integer score from 0 to 100. Targets are cutoff values in test scores that are set out to help pupils, parents and schools interpreting progress throughout compulsory education.

Table 1: Compulsory education in England

Phase	Age	School year	Key Stage	Assessment	Expected achievement level
Primary School	5-7	1-2	1	Teachers (state schools)	2
	7-11	3-6	2	External (state schools)	4
Secondary School	11-14	7-9	3	Teachers (state schools)	5 or 6
	15-16	10-11	4	External (GCSE) (all schools)	5 GCSEs at A*-C

CH is an independent selective and boarding-only mixed school that funds over

80% of the costs of pupils' education. It is a Christian institution dedicated to providing every year a stable background and boarding education of high standard to 830 boys and girls, having regard especially to children of those families in social, financial or other particular need, as stated in its mission statement. It is located in West Sussex, in South-East England, and anecdotal evidence suggests that it relies mainly on word of mouth by its alumni for publicity.

Applicants to CH have to meet its academic standards and also be judged suitable to board. They are expected to be working towards level 5 at Key Stage 2 in English, Maths and Science. After a first selection based on school reports, successful applicants are invited in for an initial assessment in English and Maths. Those who pass it will be asked in for a second assessment stage consisting in additional English and Maths tests a few months later and also to stay in the school overnight: this will help the school to assess their suitability to board. Calculations from CH show that each assessment stage screens approximately 50% of all applicants. Overall, achievement at Key Stage 2, SES and suitability to board are CH admission criteria.

Mixed grammar schools, our first control group, are highly selective, academically-oriented for historical reasons and include different school types. In our data, about 38% of pupils are in Academy Converters grammars, that are independent of control from local authorities, 28% are Foundation and 17% are Voluntary Aided or Voluntary Controlled, that also enjoy some degree of independence from local authorities although a lower one relative to Academies. The remaining 13% are Community grammars, that are not independent of control from local authorities.

Our second control group are independent schools, which are fee-paying private schools that are attended by about 7% of pupils either as day pupils or as boarders. The schools for pupils up to ages 11 or 13 are typically referred to as preparatory schools and from age 11 or 13 they can attend senior or high school. Some independent schools cover the full age range from age 3 to 18. Since CH is Christian, we restrict our attention to Christian independent day schools to study the boarding effect. Independent schools set their own examinations at the end of each year and the only national assessment their pupils sit during compulsory schooling is GCSE.[4]

[4]The percentage of pupils attending independent schools varies between about 5% for pupils aged 5-10, 8% for those aged 11 to 15, and 18% for those aged 16 to 18. About 13.5% of pupils are boarders in independent schools and only 1% of all independent schools has only boarding pupils. The average termly boarding fee is 8,780 pounds while the average termly day fee is 3,903 pounds. Bursaries, scholarships and discounts are available: around 8% of pupils have received means-tested bursaries and 1% of all pupils paid no fees at all (Independent Schools Council, 2014).

Table 2: Resources in different types of schools

	CH	Grammar	Independent	State
Pupil/teacher ratio	8.800	16.444	7.911	14.652
Full-time qualified teachers per-pupil	0.101	0.053	0.147	0.072
Part-time qualified teacher per-pupil	0.020	0.016	0.119	0.026
Total n. qualified teachers per-pupil	0.121	0.069	0.266	0.099

Grammar schools are funded by the government as they are state schools. Independent schools, instead, currently receive no direct government funding, though about 80% of them are constituted as charities (Independent Schools Council, 2014) and therefore receive important tax exemptions. They receive most of their income in the form of fees. Table 2 shows proxies of schools' teaching resources separately for CH, for our control groups and for state schools. To quantify potential differences, we matched pupil-level data on achievement and school-type to school-level data provided by the Department for Education (School Workforce Census for the school year 2006/07) with information on schools' resources, using the school identifier to match them. The teaching resources of CH are quite close to those of independent schools. Relative to grammars, however, CH has a substantially lower pupil/teacher ratio and a higher number of qualified teachers per pupil. Relative to other state schools, both CH and independent schools have higher resources while grammars have similar ones to other state schools. Finally, independent schools also devote very substantially more than state schools to non-teaching resources which may have spillover benefits for academic outcomes (Davies and Davies, 2014).

3.2 Data

Our analysis is based on individual-level administrative data on England, the National Pupil Database (NPD) matched with the Pupil Level Annual School Census, that contain information on attainment and a set of pupil characteristics for all those enrolled in state schools since primary education. The final dataset, with about 2 million pupils, contains information on five cohorts who attended primary state schools and sat their Key Stage 2 tests in years 2002-2006 and accordingly sat GCSEs at the end of Key Stage 4 in years 2007-2011.

Out of all pupils in the data, 429 went to CH after completing primary education in state schools, or about 86 pupils on average each year. About 70,000 went to secondary grammar schools and about 80,000 to independent ones. These pupils are our control groups and, in the remaining of this section, we will describe in detail the criteria to select the subsample that we will use in the empirical analysis.[5]

Figure 1: Number of CH pupils by local authority where they went to primary school and grammar (g) and independent schools (i) chosen by their peers

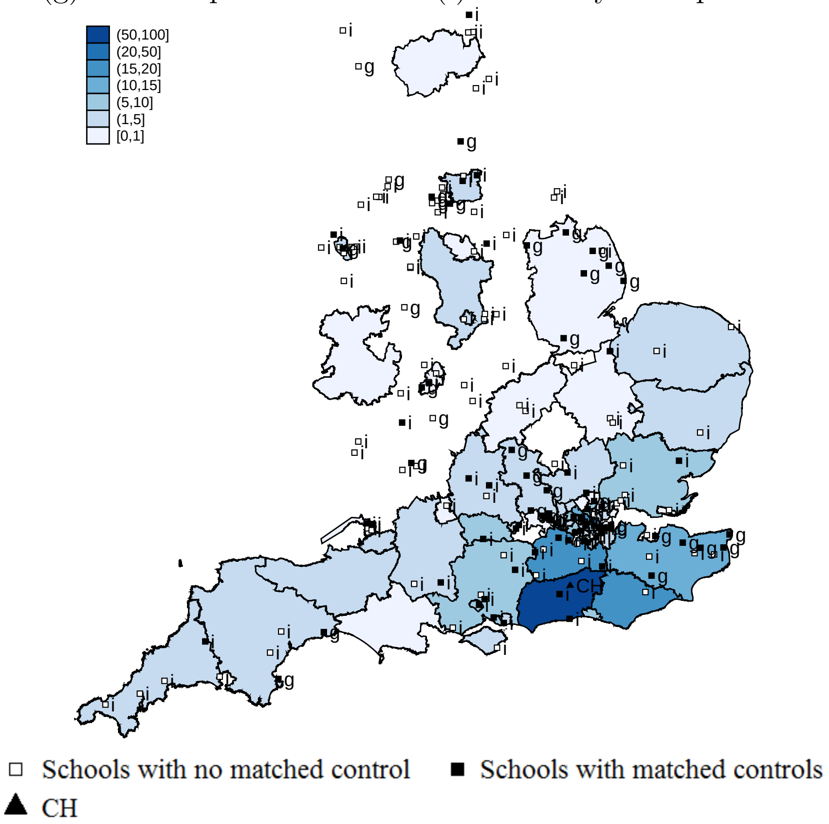


Figure 1 shows, by using a map of England, the number of CH pupils by local authority where they went to primary school as well as CH location marked by a black triangle towards the bottom end of the map. The contour of local authorities from which no pupil goes to CH after completing primary school is instead not shown. Overall, the figure shows that the majority of CH pupils come from local authorities not very far away from CH, indicated using darker colours, and, in addition, that a small number of pupils who went to primary school further away attends CH.

Two additional pieces of information are shown in Figure 1 to illustrate the reason for choosing selective day schools, i.e. grammar and independent schools, as our control groups in estimating the effect of boarding at CH. The first is the set of all grammar and independent secondary schools attended by pupils who were

[5]The name of Christ Hospital School is used in the empirical analysis in our paper in compliance with guidelines on disclosure control that can be found in point 9.5 in the National Pupil Database Agreement for the supply of data and after obtaining written approval by the Department for Education.

in a primary school located in the same local authority as those attended by CH pupils, marked using squares and with 'g' and 'i' respectively in the figure. The fact that these schools are located either in the same local authority or in adjacent ones shows that pupils may not choose the closest secondary school in the local authority in which they attended primary school. The second and related information is the set of schools actually chosen by pupils similar to those at CH, i.e. matched using the propensity score that will be defined in section 4. They are marked as black squares while white square indicate schools attended by pupils not similar enough to CH ones, according to the propensity score, and show heterogeneity in pupils' characteristics even between selective day schools.

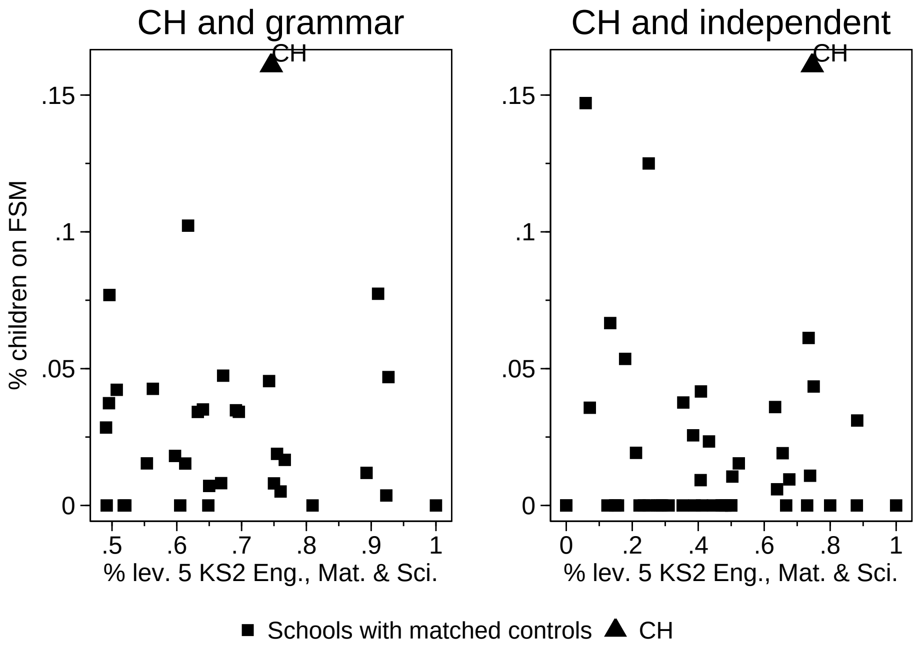
Pupils who attended grammar secondary schools are the first control group that we use to obtain an estimate of the effect of attending CH as, like CH, are academically selective while they differ through not offering boarding and through having substantially fewer resources. The second control group consists of pupils who attended independent Christian day schools after state primary as, like CH, are academically selective to varying degrees and deploy far more resources than state schools (Green *et al.*, 2012). Like CH, 90% of independent schools in our data have a Christian foundation, either Church of England or Roman Catholic, and the remaining are not religious.

In addition, we restrict our analysis to pupils in grammar and in independent schools who went to a primary school in the same local authorities as those attended by pupils at CH, who are approximately 10% of all pupils in grammar and in independent schools. This restriction ensures that CH pupils and those in the control group face the same choice set of secondary schools and, in addition, they live in the same geographical area and are ruled by the same local government.

Figure 2 shows scatterplots of the percentage of pupils on FSM, measured on the vertical axis, and the percentage of pupils that obtained the top level in Key Stage 2 tests, i.e. 5, in all three subjects, measured on the horizontal axis, by using school-level data. CH is marked in the figure using a black triangle while the black squares show the grammar and independent schools chosen by pupils who are very similar to those at CH, i.e. matched using the propensity score.

The grammar and independent schools shown in Figure 2 have been selected as some of their pupils are very similar, i.e. matched, to those at CH and we also notice that a high percentage of their pupils obtains level 5 in all Key Stage 2, particularly for grammar schools. However, CH stands out with a percentage of admitted pupils

Figure 2: Achievement at Key Stage 2 (KS2) and % FSM by secondary school



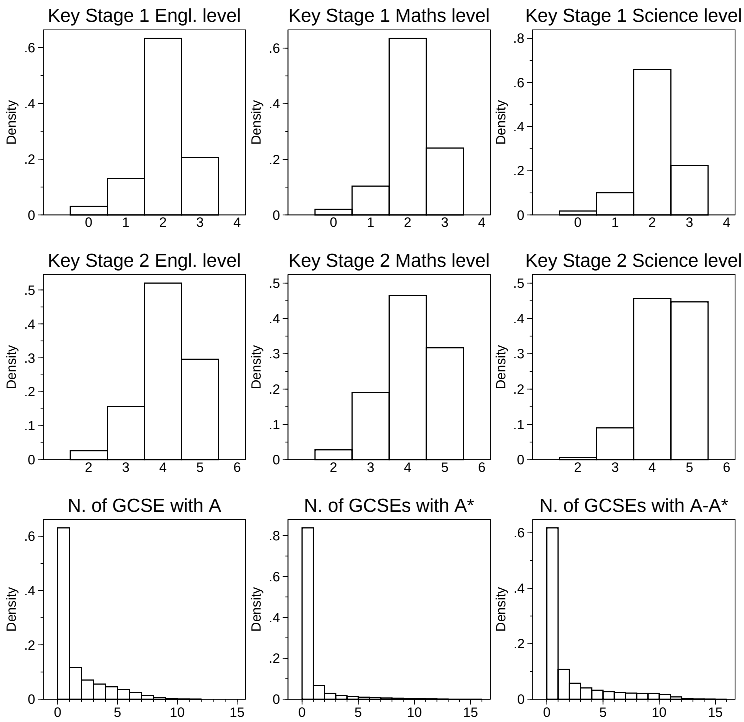
on FSM twice or three times higher than in the grammar and independent schools shown in the figure.[6]

Figure 3 shows histograms of achievement at Key Stage 1, 2 and 4 using our full dataset with 2 million pupils to help us defining meaningful achievement measures for the pupils in selective schools in our empirical analysis. The central panel shows that the percentage obtaining in individual tests a level greater than 4, that pupils are expected to achieve and coincides with the modal frequency, varies between 30 and 40%. The percentage obtaining 5 in all tests, that was used in Figure 2 as a measure of high achievement at Key Stage 2 is, instead, typically lower. Figure 3 also shows in the top panel histograms of achievement levels at Key Stage 1, in which the expected level is 2 and it coincides with the modal frequency. Hence, in the empirical analysis we will use as predetermined measures of achievement a dummy equal to 1 if a pupil obtains a level greater than 2 by subject at Key Stage 1 and a dummy equal to 1 if a level greater than 4 by subject is obtained at Key Stage 2, in addition to using Key Stage 2 tests scores.

We chose three binary variables, i.e. dummies, as outcomes. The first one is equal to 1 if a pupil obtains at least one GCSE at A and 0 otherwise. The histogram on the left-hand side in the bottom panel in Figure 3 shows that only pupils who are approximately in the top quartile of the distribution of the number of GCSEs at

[6]Additional information about how pupils are matched is found in section 4.

Figure 3: Histograms of achievement at Key Stage 1 and 2 and at GCSE



A obtain this qualification. The second one is equal to 1 if at least one GCSE is at A* and 0 otherwise, with only 10-15% of pupils obtaining this qualification, as shown by the histogram of the number of GCSEs at A* in the centre in the bottom panel in Figure 3. Finally, the third one is equal to 1 if 5 or more GCSEs are at A or A*, with only pupils in the top decile achieving this, as shown by the histogram of the number of GCSEs at A-A* on the right-hand side in the bottom panel in Figure 3. These outcomes are typically good predictors of the decision to enrol in post-compulsory education (Chowdry *et al.*, 2013).[7]

We try to match pupils at CH with as similar as possible ones in grammar or independent schools according to additional predetermined characteristics. They include socio-demographics, such as gender, ethnicity, quarter of birth and two proxies for SES. The first is the income deprivation affecting children index (IDACI), measuring the share of children in low income households in an area of about 40 households

[7] In choosing our outcomes of interest we focused on the highest grades in GCSEs, i.e. A or A*, since all secondary schools we consider are selective and its pupils tend to achieve towards the high end of the distribution of grades in GCSEs. We did not choose, instead, the probability of achieving five or more GCSEs at A*-C, a lower grade, as it is about 98% in selective schools and, similarly, the mean number of GCSEs taken by pupils in these schools is 10 and shows little variation across schools. Achievement in English and Maths at GCSE are not used as outcomes as this information is not available for CH and for independent schools in NPD data.

and 100 persons called super output areas, that is the smallest unit used for census purposes. The second is a dummy equal to 1 if a pupil obtains a Free School Meal (FSM) as her/his parents receive some form of income support. Extra dummies are used to measure if a pupil takes English as an additional language (EAL) where it is not his/her native language and whether a pupil has special education needs (SEN), both being assessed case-by-case by educational specialists in schools.

A preview of the descriptive statistics of pupils' relevant predetermined characteristics, that are shown in Figure 5 in section 4 separately for pupils at CH and in control schools shows that differences are small and not significant.

4 Econometric strategy

We estimate the effect of going to CH on achievement in the compulsory school leaving exams and whether it differs by SES and gender using propensity score (pscore) matching, an econometric strategy based on selection on observables. This is possible thanks to the unique admission criteria based jointly on merit and on SES and to the rich set of pupils' observable characteristics in the administrative data.

$$\Delta^{ATT} = E[A(1) - A(0) \mid D = 1] \tag{1}$$

Let D be a dummy indicating whether pupils go to CH, with $D = 1$ for pupils at CH (treatment) and $D = 0$ for those in a selective day school (controls). Let also $A(1)$ and $A(0)$ be the potential outcome, i.e. achievement, for treated and for controls. Finally, let X be a set of predetermined observable characteristics for pupils. Our parameter of interest is the average treatment on the treated (ATT), which we denote Δ^{ATT} and define in our setting as the mean effect of attending CH, i.e. the treatment group, rather than a selective day school, i.e. the control group, as shown in equation (1).

To recover via the law of iterated expectations the unobservable term $E[E[A(0) \mid D = 0] \mid D = 1]$ in equation (1) we rely on the assumption that admission to CH depends only on unobservables, also known as selection on observables or Conditional Independence Assumption (CIA). Under this assumption assignment to the treatment or to the control group is independent on the treatment D conditional on the set of observables X, formally $A(1), A(0) \perp D \mid X$. However, when the number of observable characteristics in the vector X is high, it may not be possible to find for some pupils at CH pupils in control schools with the same observables X, unless the number of observations in the data is very high. This problem, known as curse of dimensionality, is solved by using instead the probability of going to CH given observable characteristics X or pscore, i.e. $P(D = 1 \mid X)$.

In addition, we ensure that for each pupil at CH there is one or more with very similar observables in the control group by imposing the common support (CS) condition, i.e. $0 < P(D = 1 \mid X) < 1$. Finally, after estimating the pscore with a logit model, we match treated pupils with very similar pupils from the control group by using nearest neighbour matching method. By using as two different control groups grammar and independent schools, we obtain two sets of estimates. While in our preferred specification we use the nearest neighbour method to match pupils, we also assess the sensitivity of our results to using different matching methods based on the pscore.[8]

In our analysis so far, the assumption we made was that the choice faced by talented pupils was binary: either CH or another type of selective school, for example an independent school. However, at the end of primary school a talented pupil may have been granted admission to CH, as well as to a grammar and an independent secondary school. This can be accounted for by extending the binary propensity score matching framework to the case of multiple treatments thanks to the matching estimator proposed in Lechner (2002). By allowing multiple treatments, the treatment variable D in our setup is no longer binary and can, instead, take multiple values. In our setup of secondary school choice, D is equal to 0 if a pupil chooses an independent school, which we set as the baseline although this choice does not affect results, to 1 if the choice is a grammar one and to 2 for CH.

Firstly, a multinomial logit model of school choice is estimated using as covariates the set of observables X. Secondly, we compute the predicted probabilities $\hat{P}^j(X) = \hat{P}(D = j \mid X)$ of attending an independent school ($j = 0$), a grammar school ($j = 1$) or CH ($j = 2$). To estimate the effect of attending CH relative to, for example, an independent school we compute the conditional probability $\hat{P}^{2|2,0}(X) = \frac{\hat{P}^2(X)}{\hat{P}^2(X)+\hat{P}^0(X)}$. Finally, the estimated conditional probability is used in Lechner (2002) as balancing score in a matching estimator setting with multiple treatments to estimate the unobserved term $E[E[A(0) \mid D = 0, P^{2|2,0}] \mid D = 2]$, i.e.

[8] ATT estimation with binary treatment was conducted using the software routines described in Becker and Ichino (2002); Leuven and Sianesi (2015).

to match pupils at CH ($D = 2$) and pupils in independent schools ($D = 0$) with very similar values of the conditional probability $P^{2|2,0}$. Analogously, the procedure is repeated to estimate the effect of attending CH relative to a grammar school.[9]

5 Results

In this section, we firstly show estimates of the pscore and also means of predetermined characteristics separately for pupils at CH and for those in the control groups in subsection 5.1. Then, we show propensity score matching estimates of the effect of going to CH on achievement in the compulsory school final exam in subsection 5.2. A sensitivity analysis is then presented in section 6.

5.1 Propensity score and balance of predetermined characteristics

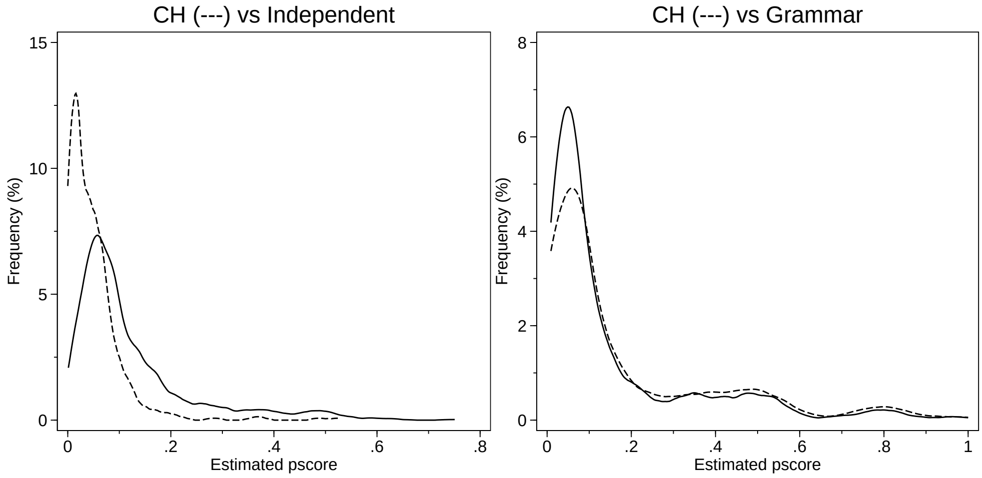
We estimate the propensity score by using a logit model and the following socio-demographic characteristics: gender, ethnicity dummies, a dummy equal to 1 if primary school was a faith school, a dummy for FSM, a dummy equal to 1 if IDACI is above the median or a dummy equal to 1 if it is in the top quartile of its distribution. We also use scores in Key Stage 2 tests by subject and dummies for whether the level obtained by the pupil was greater than the expected level 4. The advantage of using five cohorts of data in the empirical analysis is that larger samples improve the quality of the matching between CH pupils and pupils with very similar observable characteristics in the control groups.

Figure 4 shows the estimated propensity score distribution for CH pupils and for matched pupils in each of the two control groups separately. The *common support*, measured on the horizontal axis, is the interval of propensity score values over which the probability of observing pupils, measured on the vertical axis, is positive both for the control and for the treatment group and it varies from 0 to about 0.6 and to 0.9 for grammar and independent schools respectively.

By following Black and Smith (2004), we also use a more conservative definition of support, called *thick support*, that consists in using only data on pupils in the 'thick' region of the pscore distribution for treated and for controls, and is a subset of the common support. Guided by the pscore empirical distribution in Figure

[9] ATT estimation with multiple treatments was conducted by implementing the algorithms proposed in Gerfin and Lechner (2002); Lechner (2002); Frölich *et al.* (2004).

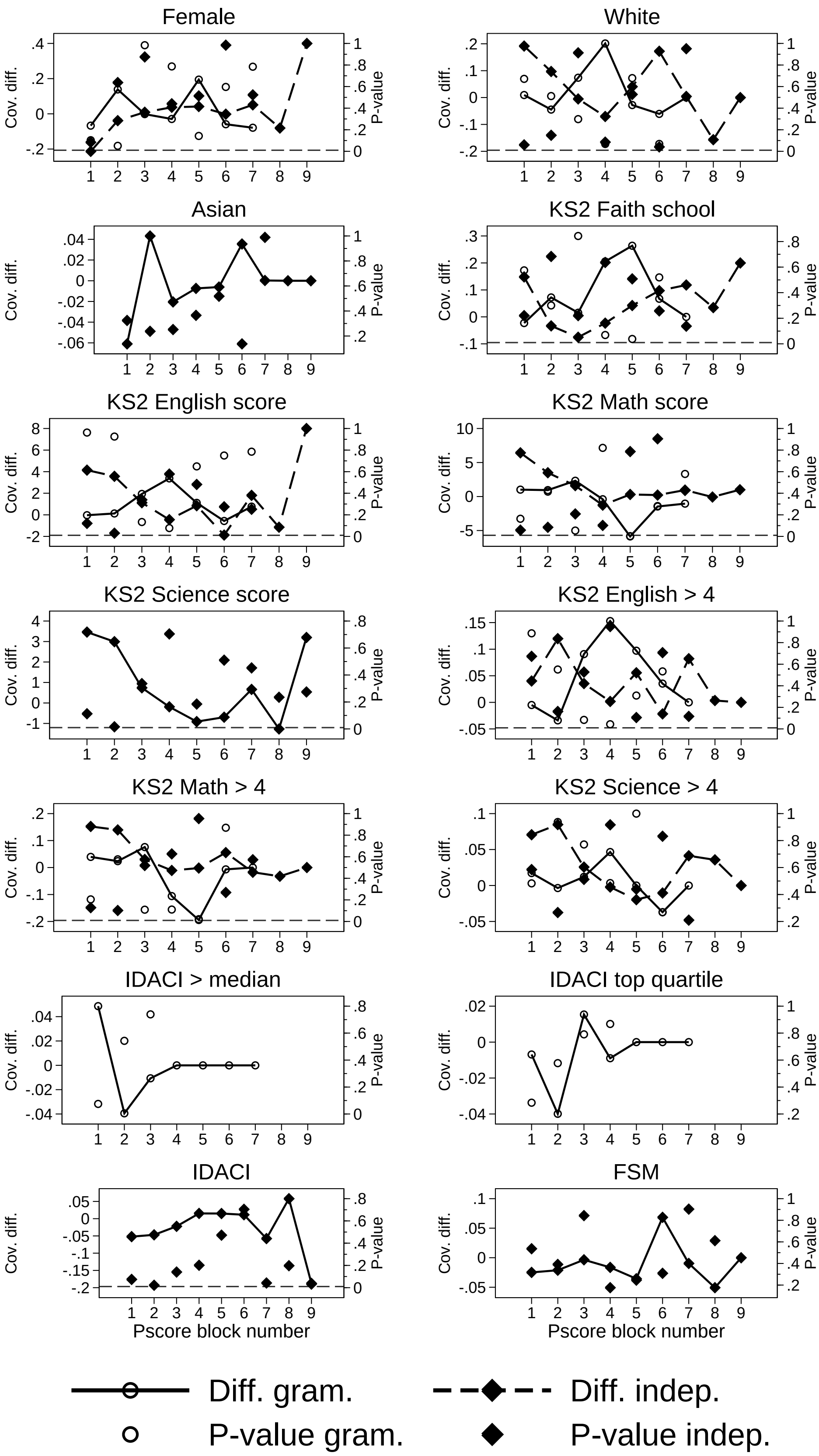
Figure 4: Kernel density estimate of the propensity score



4, we chose as thick support the interval between 0 and 0.2 and an even smaller interval, 0-0.1, i.e. we drop observations for pupils with pscore in the right tail of the distribution. Estimates obtained after excluding pupils outside the thick support region are helpful to assess whether those obtained under the common support are potentially biased due to self-selection into or out of CH, since it is more likely for pupils in the tails rather than in the middle of the pscore distribution.

Descriptive statistics of pupils' predetermined characteristics by the time they started secondary education, that are used to assess the balancing property after estimating the propensity score, are shown in Figure 5. The vertical axis on the left-hand side measures the difference between pupils at CH and controls in, for example, the relative frequency of females in the top left of the figure, separately for pscore blocks measured along the horizontal axis. After estimating the pscore, the blocks were defined along the pscore support to ensure that predetermined characteristics are balanced. Pscore estimation using pupils in grammar schools as controls required splitting the data sample into 7 different blocks according to pupils' estimated pscore while 9 blocks were used when the control group were pupils in independent schools. In addition, the vertical axis on the right-hand side measures p-values of t-tests of the null hypothesis of no difference in the mean value between treated and controls by block.

Figure 5: Covariates differences for CH relative to grammar and independent by pscore block



The plot on the top left in Figure 5 shows that the difference in the relative frequency of females by pscore block in CH relative to grammar schools, reported using a continuous line marked by diamonds, is either slightly positive or zero and p-values, reported using a scatterplot of diamonds, are greater than the 5% conventional level. The difference in the frequency of females by pscore block at CH relative to independent schools is reported, instead, using a dotted line marked by circles and its p-values, reported using a scatterplot of circles, are also greater than 5%. Overall, Figure 5 shows that predetermined characteristics are balanced, except for p-values close to 5% for some pscore blocks of Key Stage 2 scores. This suggests that the propensity score was helpful to choose in grammar and in independent schools those pupils who are most similar to pupils at CH in terms of observables.[10]

5.2 The effect of CH on achievement

In this section, we report ATT estimates of the impact on achievement in the compulsory school final exam of attending CH, rather than a day grammar or independent school, to test our first hypothesis that offering a better learning and non-school environment to high ability pupils with low SES increases their achievement (H_1). Overall, the positive and significant ATT estimates in Table 3 offer support to our hypothesis.

Table 3: Effect of attending CH on results in school-leaving exams

	Grammar schools			Independent schools		
	ATT	Matched	All	ATT	Matched	All
1+ GCSE with A	0.044**	0.888	0.868	0.100***	0.832	0.754
S.e.	0.020			0.024		
1+ GCSE with A*	0.170***	0.653	0.570	0.084**	0.739	0.527
S.e.	0.031			0.031		
5+ GCSE with A-A*	0.174***	0.593	0.513	0.126***	0.641	0.422
S.e.	0.033			0.034		
N		494	7,075		369	8,118

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

To match controls to treated we used as our preferred method nearest neighbour matching with replacement and set to 0.01 the maximum distance in pscore that is

[10] Since in Figure 5 the difference in the relative frequency of pupils with an IDACI above the median or an IDACI in the top quartile is zero in some blocks, corresponding p-values are not reported.

allowed to perform a match. The estimates in Table 3, obtained using the common support, show that the probability of obtaining at least 1 (1+ hereafter) GCSEs with A is 4.4 and 10 percentage points (ppt hereafter) higher relative to grammar and independent schools respectively, with ATT estimates being significant. This is 5% and 12% higher relative to the value for matched controls, that is also shown in the table. Differences in the probability of obtaining 1+ GCSEs with A* are also significant and show that the point estimate is 17 and 8.4 ppt higher or about 26% and 11% relative to grammar and independent schools. Finally, the probability of obtaining 5+ GCSEs with A-A* is 17.4 and 12.6 ppt higher or about 29% and 20% relative to the control groups respectively. Overall, the point estimates are higher when using as outcome the dummy equal to 1 if pupils obtain 1+ GCSEs with A* or 5+ GCSEs with A-A*, who are approximately in the top decile of the distribution of achievement in GCSE exams among all pupils in the administrative data, as shown in Figure 3.[11]

In addition to ATT estimates and mean values of outcomes for matched controls, Table 3 also shows mean values for all controls to compare our ATT estimates with naive estimates obtained as the difference in mean achievement between CH pupils and all pupils in grammar and in independent schools respectively. Naive estimates have the same sign as our ATT estimates. However, the point estimates are greater since the mean value of the outcomes for all controls is smaller than for matched controls. Under our untestable identifying assumption of selection on observables naive estimates are then biased upwards relative to our ATT estimates. This comparison also suggests that had pupils at CH instead gone to grammar or independent day schools, they would have obtained higher scores than the average in those schools.

Finally, Table 3 shows that the ATT for the probability of obtaining 1+ GCSEs at A, i.e. of being a moderately high achiever at GCSE, is higher at CH when controls are pupils from independent schools while the probability of obtaining 1+ GCSEs at A* or 5+ GCSEs at A-A*, i.e. of being a very high achiever, tends to be higher when controls are from grammar schools. However, the significance of

[11] Results not reported but available upon request show that ATT estimates change little if additional predetermined characteristics are used, such as achievement in all tests at Key Stage 1, the type of school at Key Stage 2 and the distance to the closest secondary schools. However, since we used as criterion to choose the predetermined characteristics that are used as covariates in estimating the pscore the results of covariates balancing analysis that is described in section 5.1, we did not use these additional covariates as they were slightly unbalanced.

the difference is not testable with our econometric strategy based on selection on observables without making additional assumptions.

Table 4: Effect of attending CH for pupils in the pscore thick support

	Grammar schools			Independent schools		
	ATT	Matched	All	ATT	Matched	All
Pscore thick support 0-0.2						
1+ GCSE with A	0.065***	0.868	0.868	0.072***	0.856	0.754
S.e	0.021			0.026		
1+ GCSE with A*	0.192***	0.623	0.570	0.084**	0.730	0.527
S.e	0.032			0.034		
5+ GCSE with A-A*	0.179***	0.576	0.513	0.126**	0.631	0.422
S.e	0.034			0.037		
N		450	7,075		306	8,118
Pscore thick support 0-0.1						
1+ GCSE with A	0.079***	0.844	0.868	0.081***	0.856	0.754
S.e	0.026			0.030		
1+ GCSE with A*	0.231***	0.576	0.570	0.077*	0.703	0.527
S.e	0.038			0.043		
5+ GCSE with A-A*	0.218***	0.523	0.513	0.126***	0.581	0.422
S.e	0.040			0.046		
S.e		338	7,075		214	8,118

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table 4 reports additional ATT estimates of the impact of attending CH relative to control schools for pupils whose pscore is in a 'thick' region of the pscore distribution, following Black and Smith (2004). We define as thick support pscore values in the range 0-0.2 and, additionally, a more narrow range: 0-0.1. Table 4 shows overall that our main results are robust to using only pupils in the thick support when considering the sign of the point estimates, as well as their size and significance. However, slight differences emerge across control groups. Thick support estimates when pupils in grammar schools are the controls are slightly greater than those obtained on the common support, with the greatest differences being for the probability of obtaining 1+ GCSEs at A. When looking, instead, at thick support estimates obtained with pupils in independent schools as controls, Table 4 shows that they are very similar to common support ones.[12]

ATT estimates for subsamples of pupils by gender and by SES are shown in

[12]In choosing the pscore intervals defining the thick support regions, we use as guidance the empirical distribution of pscores in Figure 4.

Table 5: Effect of attending CH for pupils' subgroups by gender and SES

	Grammar schools			Independent schools		
	ATT	Matched	All	ATT	Matched	All
			Males			
1+ GCSE with A	0.081***	0.856	0.848	0.127***	0.810	0.722
S.e.	0.028			0.034		
1+ GCSE with A*	0.177***	0.587	0.527	0.072*	0.692	0.495
S.e.	0.044			0.044		
5+ GCSE with A-A*	0.127***	0.560	0.471	0.093**	0.595	0.390
S.e.	0.046			0.047		
N		292	3,701		208	5,460
			Females			
1+ GCSE with A	-0.016***	0.944	0.890	0.151***	0.776	0.746
S.e.	0.026			0.038		
1+ GCSE with A*	0.174***	0.722	0.618	0.214***	0.682	0.518
S.e.	0.041			0.044		
5+ GCSE with A-A*	0.190***	0.675	0.558	0.271***	0.594	0.423
S.e.	0.044			0.047		
N		227	3,374		169	3,230
			IDACI below median			
1+ GCSE with A	0.002	0.930	0.881	0.117***	0.818	0.782
S.e.	0.028			0.042		
1+ GCSE with A*	0.124***	0.652	0.607	0.080	0.693	0.550
S.e.	0.049			0.056		
5+ GCSE with A-A*	0.152***	0.575	0.544	0.102*	0.620	0.451
S.e.	0.052			0.059		
N		227	3,562		126	4,393
			IDACI above median			
1+ GCSE with A	0.055**	0.877	0.854	0.089***	0.842	0.679
S.e.	0.026			0.031		
1+ GCSE with A*	0.172***	0.679	0.534	0.113***	0.733	0.456
S.e.	0.041			0.039		
5+ GCSE with A-A*	0.190***	0.601	0.481	0.147***	0.640	0.352
S.e.	0.044			0.043		
N		274	3,513		227	4,297
			FSM			
1+ GCSE with A	0.174**	0.754	0.777	0.130*	0.797	0.335
S.e.	0.080			0.073		
1+ GCSE with A*	0.275***	0.551	0.458	0.203**	0.623	0.196
S.e.	0.103			0.096		
5+ GCSE with A-A*	0.290***	0.449	0.343	0.159	0.580	0.149
S.e.	0.105			0.100		
N		44	251		43	409
			No FSM			
1+ GCSE with A	0.050**	0.884	0.871	0.089***	0.844	0.751
S.e.	0.021			0.025		
1+ GCSE with A*	0.174***	0.648	0.575	0.097***	0.725	0.519
S.e.	0.032			0.033		
5+ GCSE with A-A*	0.183***	0.589	0.519	0.167**	0.606	0.415
S.e.	0.034			0.036		
N		437	6,824		328	8,281

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table 5. Results separately by gender are in line with common support estimates except the very small and negative effect of obtaining 1+ GCSE at A for females, suggesting some heterogeneity by gender for those pupils who are not among top achievers at CH since the result for males is positive. Heterogeneity by gender also seems to be present among top achievers, i.e. pupils with 5+ GCSE at A-A*, as it is shown by greater point estimates for females.

We examined subsamples by SES in two alternative ways. First, we obtained estimates for pupils who live in an area with an IDACI value above the median of the distribution, i.e. poor areas, and contrast these with estimates for those with an IDACI value below the median, i.e. a more affluent areas. Second, we compared estimates for pupils according to whether or not they were on FSM. When we look at results separately by whether IDACI is low or high, we find that achievement gains arising from attending CH tend to be higher for pupils with a high IDACI. Results by FSM are similar although when the controls are pupils in grammars their precision is lower due to the low number of matched controls. However, we cannot fully test these differences in a selection on observables framework without additional assumptions.

To summarise, ATT point estimates of the effect of attending CH are greater when using as outcomes proxies for high achievers at GCSE, i.e. 1+ GCSEs at A* or 5+ GCSEs with A-A*, who are among the top 10-15% in the distribution of achievement at GCSE. For these same variables, estimates of the CH effect obtained using grammar school pupils as controls tend to be greater than those obtained using independent school pupils. When we consider only those pupils in the thick support region, we find that point estimates are very similar to those obtained using the common support. This similarity suggests that estimates are driven by pupils in the middle of the achievement distribution at GSCE rather than by those in the tails tail and hence that they are little confounded by self-selection in the right tail of the propensity score distribution. Finally, we find that the effect is greater and tends to be more precise for females and for children in poor households.

6 Sensitivity analysis

In this section we perform a sensitivity analysis of our main results. Firstly, we compare them with matching estimates obtained by allowing for multiple treatments, i.e. CH, grammar or independent schools, rather than a binary one, following Lechner

(2002). Table 6 shows matching estimates that were obtained by letting the different types of selective schools that we consider be multiple treatments. The sign and size of the two sets of point estimates, as well as their significance, are in line with our main results in Table 3, that are obtained by assuming, instead, a binary treatment. Overall, this suggests that relaxing the assumption of modelling choice of CH relative to a different selective secondary school as a binary treatment does not substantially alter our main results. The only difference is that the estimate of the probability of obtaining 1+ GCSE at A with pupils in grammar schools as controls is smaller and no longer significant. This may be due to the lower number of matched controls in grammar schools and may lead to a poorer match relative to our main results, particularly when looking at the probability of obtaining 1+ GCSE at A, as grammar school pupils tend to be very high ability pupils achieving top grades at Key Stage 2 and at GCSE.

Table 6: Matching estimates of CH effect using pscore from multinomial logit

	Grammar schools			Independent schools		
	ATT	Mean for controls		ATT	Mean for controls	
		Matched	All		Matched	All
1+ GCSE with A	0.023	0.907	0.868	0.096***	0.837	0.754
S.e.	0.032			0.024		
1+ GCSE with A*	0.202***	0.622	0.570	0.147***	0.676	0.527
S.e.	0.049			0.032		
5+ GCSE with A-A*	0.183***	0.574	0.513	0.184***	0.583	0.422
S.e.	0.051			0.034		
N		175	7,075		370	8,118

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Secondly, we compare our main results, obtained by using nearest neighbour matching, with results obtained using different matching methods, as one of the limitations of nearest neighbour is finding a match for all CH pupils and not controlling for the 'quality' of the matching, i.e. how similar to pupils at CH are pupils in control schools in terms of their predetermined characteristics. With kernel and radius matching, instead, a pupil at CH can be matched with more than one pupil in the control group and the estimated counterfactual outcome for that pupil at CH is a weighted average of the outcome value for matched pupils in the control group, with the weight increasing with the quality of the matching. In kernel matching a CH pupil is matched with all pupils in the control group and the weight is inversely proportional to the distance between the propensity score value for that CH pupil and for controls.

In radius matching, instead, only control group pupils whose value of the propensity score is within a fixed radius from the one of a given CH pupil are matched with her/him. The weight is equal to the inverse of the number of matched pupils, which is the same for all controls matched to the same pupil at CH. Finally, instead of relying on the propensity score as a metric to match treated and controls, we use Mahalanobis distance. In the context of matching, this is a scalar measure of the square of the distance between the vector of covariates for a pupil at CH relative to the one for a pupil in the control group, multiplied by the inverse of the covariance matrix of the difference between the vectors.

Table 7: Matching estimates of CH effect using different matching methods

	Grammar schools			Independent schools		
	ATT	Matched	All	ATT	Matched	All
			Kernel			
1+ GCSE with A	0.047***	0.885	0.868	0.129***	0.803	0.754
S.e.	0.013			0.015		
1+ GCSE with A*	0.204***	0.619	0.570	0.165***	0.658	0.527
S.e.	0.020			0.022		
5+ GCSE with A-A*	0.203***	0.564	0.513	0.202***	0.564	0.422
S.e.	0.022			0.023		
N			7,075			8,118
			Radius with size 0.1			
1+ GCSE with A	0.056***	0.876	0.868	0.146***	0.787	0.754
S.e.	0.013			0.015		
1+ GCSE with A*	0.230***	0.593	0.570	0.206***	0.617	0.527
S.e.	0.019			0.021		
5+ GCSE with A-A*	0.231***	0.536	0.513	0.244***	0.522	0.422
S.e.	0.021			0.023		
N			7,075			8,118
			Mahalanobis			
1+ GCSE with A	0.049**	0.883	0.868	0.058***	0.874	0.754
S.e.	0.020			0.022		
1+ GCSE with A*	0.210***	0.613	0.570	0.131***	0.692	0.527
S.e.	0.029			0.028		
5+ GCSE with A-A*	0.203***	0.564	0.513	0.182***	0.585	0.422
S.e.	0.028			0.030		
N			7,075			8,118

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table 7 shows ATT estimates separately by matching method across different horizontal panels. The top panel shows estimates obtained using kernel matching, estimates in the central panel were obtained using radius matching and, finally, those in the bottom panel using a Mahalonobis distance. Overall, the table shows that the sign, size and precision of point estimates is in line with our main results in Table 3. As for the size of point estimates, those of the probability of obtaining 1+ GCSE with A or with A-A* are slightly greater than our main results.[13]

Thirdly, we assess whether our main results vary if we focus only on pupils whose primary school was located at a low distance from CH or also add those pupils living further away, as evidence on England shows that distance matters in secondary school (Burgess *et al.*, 2006). This distance offers a proxy for the opportunity cost of attending CH, in terms of transportation time, its monetary cost and also the psychological cost associated with leaving parents' home. Potential differences might suggest that the decision to attend CH is correlated with distance and with unobservables such as the psychological cost of displacement. We obtained measures of linear distances in miles from pupils' primary school to CH, as well as to the closest grammar and independent schools.[14]

To simplify results description, we plot in Figure 6 how the ATT of attending CH on the probability of obtaining, for example, 1+ GCSEs at A along the vertical axis varies when we consider pupils living at a different distance between primary school and CH. The ATT is measured along the vertical axis on the left-hand side and it is shown as a thick dashed line while its 95% confidence intervals are shown as thin dashed lines. Distance along the horizontal axis is measured in miles and the ATT is shown for the following values along the distance distribution: 5th, 10th, 25th, 50th, 75th, 90th, 95th and 99th percentiles, which are marked by circles along the line reporting the ATT. In addition, the number of observations is measured along the vertical axis on the right-hand side and it is shown as a scatterplot separately for treated and for controls, that are marked as triangles and squares respectively.

[13]The number of matched controls is not shown in the Table 7 as the matching methods used to obtain the estimates shown in it are not one-to-one, i.e. each treated is not matched to a single control but rather to several ones.

[14]Distances are computed by using publicly available data on schools' postcodes and on longitude and latitude coordinates associated to postcodes, measured using the World Geodetic System 1984 (Ordnance Survey website). They are then converted into Ordnance Survey Maps northing and easting coordinates thanks to a Helmert transformation (Watson, 2006) to eventually obtain distances in miles. We use the postcode of pupils' primary schools rather than of their home as the latter information is not publicly available. We argue that our results would not change substantially after obtaining pupils' postcodes to compute a more precise measure of the distance between home and secondary school, as anecdotal evidence suggests that distance to primary school tends to be typically low and, in addition, it seems to be subject to moderate variation across pupils (Burgess *et al.*, 2015).

Figure 6: ATT estimates by value of the distance between primary school and CH

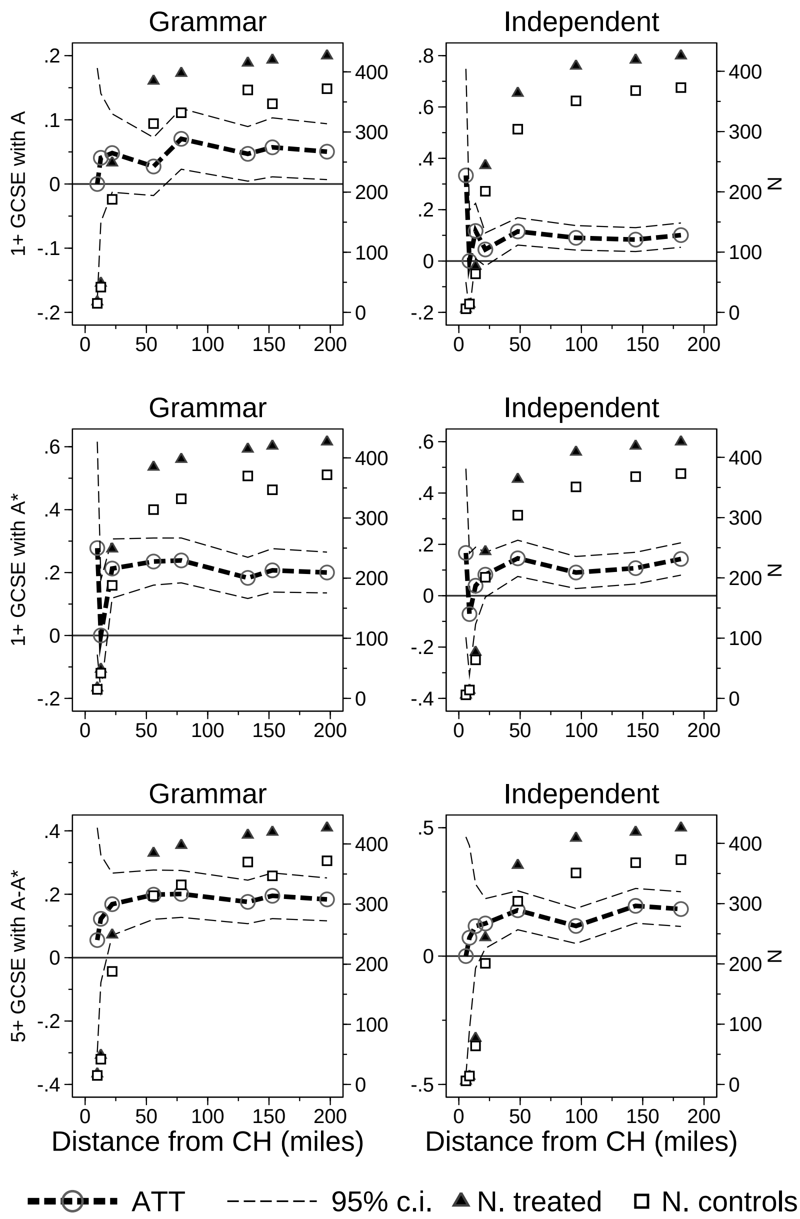


Figure 6 shows overall that the sign and size of the estimates are in line with our main results even when estimates are obtained using subsamples of pupils who live within a given distance from CH. As for estimates significance, it is smaller than or equal to 5% for distance values greater than or equal to 20-25 miles. This corresponds approximately to the 25th or 50th percentile of the distance distribution and it is slightly smaller than the mean distance of about 30 miles for treated and matched controls.

Finally, we apply the methodology proposed in Ichino *et al.* (2008) to assess the sensitivity of our main results to a failure of pscore matching identifying assumption, i.e. the CIA, due to the presence of an unobservable covariate whose distribution is similar to the empirical distribution of an observable covariate. We let U be an unobserved term, assumed binary in Ichino *et al.* (2008) for simplicity, and its distribution be fully determined by four parameters $p_{ij} = Pr(U = 1 \mid D = i, A = j, X)$ measuring the probability that the unobserved term is equal to 1 given that the treatment D, i.e. school choice in our setting, is equal to i and the outcome A, i.e. achievement, is equal to j, with $i, j = \{0, 1\}$.

$$\Gamma = \frac{\frac{Pr(A=1 \mid D=0, U=1, X)}{Pr(A=0 \mid D=0, U=1, X)}}{\frac{Pr(A=1 \mid D=0, U=0, X)}{Pr(A=0 \mid D=0, U=0, X)}} \tag{2}$$

By assuming $p_{01} > p_{00}$, i.e. that the unobserved confounder has a positive effect on the untreated outcome, and accounting for the relationship between U and X, Ichino *et al.* (2008) define the outcome effect Γ as the effect of U on the probability of a positive outcome A and compute it as the odds ratio of U after estimating the logit model of $Pr(A = 1 \mid D = 0, U, X)$, as shown in equation (2). In addition, the selection effect Δ is defined as the effect of U on the probability of treatment, i.e. $D = 1$, and it is computed as the odds ratio of U after estimating the logit model of $Pr(D = 1 \mid U, X)$, as shown in equation (3).

$$\Delta = \frac{\frac{Pr(D=1 \mid U=1, X)}{Pr(D=0 \mid U=1, X)}}{\frac{Pr(D=1 \mid U=0, X)}{Pr(D=0 \mid U=0, X)}} \tag{3}$$

Based on values of p_{ij}, with $i, j = \{0, 1\}$ obtained by using the empirical distribution of a relevant covariate, a value of U is imputed for each pupil in the dataset. The variable U is then treated as any observed covariate in X to first estimate the pscore and then the ATT using nearest neighbour matching. Varying the values of the sensitivity parameters p_{ij} and repeating the pscore and ATT estimation in a simulation with 1000 repetitions, the average of the ATT over the distribution of U is obtained.[15]

In our setting achievement in Key Stage 2 tests at age 11 and SES are observable characteristics used by CH to select its pupils while suitability for boarding is unobservable to the econometrician, due to the impossibility to match CH admission

[15] A more detailed description of the econometric details behind the sensitivity analysis is found in section 4 in Ichino *et al.* (2008).

data with NPD administrative data on all pupils. Hence, we assess the sensitivity of our main results to unobserved binary covariates whose distribution is similar to the one of observed measures of pupils' ability, as at least part of a pupil's ability is typically unobserved and may be correlated with the pupil's resilience to adapt to boarding.

As ability proxies, we use dummies equal to 1 if a pupil achieved in the Key Stage 1 Maths test a level greater than the expected one, i.e. 2, as it is typically a more precise measure of ability than using the English test, and if the level is greater than the expected one, i.e. 4, in all Key Stage 2 tests. In addition, we use the distance between primary school and CH or the closest grammar or independent secondary school as an observable measure of the opportunity cost of attending CH. This may be a relevant factor for secondary school choice as the further away a pupil lives from CH the higher the psychological effort required to adapt to boarding.

Table 8 shows in Panel A estimates of the effect of CH obtained on our three measures of achievement at GCSE by using pupils in grammar schools as controls. Estimates on each row are obtained by using a confounder U distributed according to a different covariate. Along a row, the first four columns on the left-hand side show values of the probabilities p_{ij} characterising the distribution of U by using the empirical distribution of a covariate, then the outcome and selection effect are shown and, finally, ATT estimates.

For each outcome variable, Table 8 shows firstly estimates obtained using a neutral confounder, i.e. with all p_{ij} set equal to approximately 0.5. On the two following rows the unobserved confounder is distributed similarly to observed measures of ability, proxied by dummies measuring achievement at Key Stage 1 and at Key Stage 2. In the three final rows, instead, the confounder is distributed following the empirical distribution of a dummy equal to 1 if the distance in miles between primary school and CH is greater than the median value, as well as two additional dummies equal to 1 if the distance to the closest grammar secondary or to the closest independent secondary is greater than the median.

Estimates in Table 8 show overall that both their magnitude and precision are in line with our main results. When we look instead at the outcome effect, i.e. the effect of U on the probability of higher achievement, and at the selection effect, i.e. the effect of U on the probability of attending CH, the table shows that the value of both effects is very close to one in the case of neutral confounder, which is expected as by setting all p_{ij} to 0.5 the confounder is close to i.i.d. When we look at proxies

Table 8: Sensitivity analysis of CH effect using calibrated confounders

	p_{11}	p_{10}	p_{01}	p_{00}	Outcome effect Γ	Selection effect Δ	ATT	S.e.
				Panel A: grammar schools				
				1+ GCSEs with A				
Neutral conf.	0.507	0.448	0.499	0.521	0.915	1.012	0.044**	0.020
KS1 Mat >2	0.993	0.966	0.996	0.993	1.804	0.687	0.049**	0.023
All KS2 >4	0.752	0.655	0.676	0.391	3.261	1.640	0.035	0.025
Miles pri.-CH > median	0.100	0.138	0.505	0.597	0.692	0.108	0.031	0.026
Miles pri.-gram. > median	0.530	0.517	0.496	0.498	0.999	1.145	0.048*	0.026
Miles pri.-indep. > median	0.270	0.345	0.494	0.520	0.903	0.393	0.043*	0.026
				*1+ GCSEs with A**				
Neutral conf.	0.467	0.500	0.506	0.496	1.042	0.893	0.170***	0.031
KS1 Mat >2	0.994	0.974	0.996	0.994	1.617	0.661	0.185***	0.036
All KS2 >4	0.796	0.513	0.777	0.453	4.190	1.563	0.164***	0.039
Miles pri.-CH > median	0.099	0.118	0.472	0.577	0.659	0.111	0.156***	0.041
Miles pri.-gram. > median	0.521	0.566	0.487	0.508	0.919	1.149	0.190***	0.038
Miles pri.-indep. > median	0.261	0.342	0.473	0.529	0.801	0.392	0.179***	0.040
				*5+ GCSEs with A-A**				
Neutral conf.	0.523	0.560	0.497	0.497	1.002	1.156	0.174***	0.033
KS1 Mat >2	0.994	0.980	0.997	0.993	2.432	0.637	0.185***	0.037
All KS2 >4	0.815	0.520	0.804	0.463	4.751	1.540	0.157***	0.040
Miles pri.-CH > median	0.109	0.080	0.475	0.562	0.704	0.110	0.155***	0.043
Miles pri.-gram. > median	0.517	0.570	0.483	0.510	0.901	1.157	0.189***	0.040
Miles pri.-indep. > median	0.255	0.340	0.469	0.527	0.795	0.392	0.173***	0.042
				Panel B: independent schools				
				1+ GCSEs with A				
Neutral conf.	0.490	0.586	0.505	0.510	0.983	0.966	0.103***	0.024
KS1 Mat >2	0.993	0.966	0.976	0.828	8.721	7.321	0.094***	0.030
All KS2 >4	0.752	0.655	0.511	0.169	5.023	3.633	0.059*	0.029
Miles pri.-CH > median	0.417	0.345	0.490	0.542	0.818	0.707	0.110***	0.032
Miles pri.-gram. > median	0.388	0.379	0.511	0.491	1.085	0.616	0.112***	0.031
Miles pri.-indep. > median	0.463	0.414	0.497	0.515	0.935	0.857	0.114***	0.031
				*1+ GCSEs with A**				
Neutral conf.	0.513	0.592	0.503	0.499	1.017	1.115	0.107***	0.031
KS1 Mat >2	0.994	0.974	0.987	0.885	9.194	6.704	0.122***	0.038
All KS2 >4	0.796	0.513	0.656	0.178	8.846	2.867	0.054	0.038
Miles pri.-CH > median	0.431	0.329	0.463	0.546	0.723	0.741	0.126***	0.039
Miles pri.-gram. > median	0.374	0.447	0.494	0.517	0.912	0.629	0.128***	0.039
Miles pri.-indep. > median	0.462	0.447	0.471	0.534	0.781	0.892	0.129***	0.039
				*5+ GCSEs with A-A**				
Neutral conf.	0.483	0.420	0.495	0.492	1.015	0.898	0.163***	0.034
KS1 Mat >2	0.994	0.980	0.991	0.899	12.351	6.379	0.160***	0.041
All KS2 >4	0.815	0.520	0.733	0.207	10.469	2.740	0.085**	0.042
Miles pri.-CH > median	0.438	0.330	0.480	0.521	0.849	0.715	0.160***	0.042
Miles pri.-gram. > median	0.377	0.420	0.507	0.504	1.013	0.618	0.166***	0.042
Miles pri.-indep. > median	0.459	0.460	0.472	0.522	0.820	0.887	0.168***	0.041

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

for unobserved ability, both the outcome and selection effect are greater than 1, with the outcome effect being greater. This suggests a positive selection into CH and a positive effect on achievement for CH pupils with high unobserved ability.

Finally, when we look at proxies for the opportunity cost of attending CH, Table 8 shows that both the outcome and selection effect are smaller than 1, which suggests that a high unobserved opportunity cost leads to a lower probability of high achievement and of attending CH respectively. In addition, the outcome effect is closer to one than the selection effect, suggesting that the opportunity cost affects selection more strongly. These results hold qualitatively for all the three outcomes we consider and for both our control groups, that are shown in Panel A and B respectively.[16]

7 Discussion

In this paper we tested the hypothesis that attending Christ Hospital (CH), a boarding school admitting a high share of high ability pupils with low socio-economic status (SES), improves achievement in the compulsory school final exams (GCSEs), by using administrative data on England. Our propensity score matching estimates are substantial: the probability of achieving A or A* in five or more GCSEs is 17.4 percentage points higher with respect to 59% for matched pupils in grammar schools, i.e. a 29% increase, with similar results when the control group are independent school pupils. As an additional hypothesis, we tested for heterogeneous effects and find that the CH effect is higher for low SES pupils and for girls.

Since CH differs from independent day schools in that it is boarding and tends to be more selective based on ability, when pupils in independent day schools are the control group we estimate the joint effect of boarding and of ability selection. However, since independent schools display a high variability in pupils' ability, ranging from very high for pupils admitted with a bursary to a lower level for fee-paying pupils, the quality of the matching is preserved as high ability pupils at independent schools can be repeatedly matched to similar pupils at CH. Boarding, therefore, is the most plausible mechanism underlying our estimated effect.

When, instead, pupils in grammar schools are the controls, our estimates capture the overall effect of substituting family with school inputs and of having access to better school inputs since CH is boarding and has more resources. Although

[16] The analysis of killer confounders in Ichino *et al.* (2008), which consists in increasing jointly the extent of selection and of outcome effects until a pair of values for these effects that 'kills' the main results is found, is not shown as it is little informative about the nature of the unobserved information, e.g. ability or opportunity cost, that may bias our main results. However, these estimates are available upon request.

we cannot separately quantify the boarding effect and the resources effect without additional assumptions, the fact that we obtain similar results with the independent day schools control group, where resources are much closer to those of CH, suggests that boarding is an important part of the explanation for the difference between CH's and the grammar schools' exams performances.

Our paper contributes to two related and recent studies exploiting lottery-based admission into oversubscribed boarding schools in the US (Curto and Fryer Jr, 2014) and in France (Behaghel *et al.*, 2017). Firstly, in our setting we can test hypotheses on pupils with high ability coming from low SES, thanks to an overall sample size of approximately 8,000 observations while the lottery studies cannot as they have fewer observations, approximately 400 in total. Secondly, by estimating a treatment on the treated (ATT), we offer complementary evidence to the quasi-experimental one obtained using a local average treatment effect (LATE). On the one hand, ATT has a somewhat "stronger" identification assumption based on selection on observables while, on the other, it relies on a bigger control group than quasi-experimental studies. However, what our study and the two related ones on boarding schools have in common is low external validity as they all use as treated group either a single boarding school or a small number of them, which makes them little representative of the universe of boarding schools in a country.

Our paper also contributes to empirical studies estimating an educational production function to assess the effect of those school-based policies set up to counteract the low SES negative influence on pupils' achievement (see for a survey Todd and Wolpin, 2003). We isolate the boarding, i.e. school, effect in a simple setting in which parental responses are low for all boarders while it cannot be done in the production function, where family inputs may either decrease if school and family inputs are substitutes or increase if they are complements.

Our additional results that the impact of CH is substantially higher for girls over boys are novel as they show that the documented gender achievement gap in favour of girls in England also holds for high ability children in selective schools. However, this result is based on a low number of observations for pupils in selective schools and additional work is required to fully test for gender differences based on a greater number of pupils.

Our econometric strategy based on propensity score matching relies on the unconfoundedness assumption that unobservable characteristics, such as ability or motivation, are unlikely to be different for CH pupils relative to their match in selective day schools if the set of observables used is rich enough to capture the most relevant factors driving selection into a selective secondary school. We showed in our sensitivity analysis that our results are robust to a number of assumptions on the correlation between unobservable and observable characteristics, such as ability. A complementary approach would consist in making a different set of assumptions on the role played by unobservables to quantify the value of the ratio between the extent of selection on unobservables and of selection on observables such that our results would be completely driven by selection, using the methodology in Altonji *et al.* (2005).

Our analysis paves the way for a number of extensions, on some of which we plan to work in the future. We have not yet looked at the probability of continuing with post-compulsory education, namely sixth form, achievement in A-levels, admission into prestigious universities, degree choice and achievement and, finally, labour market outcomes. In addition, we have so far focused on a single selective and boarding school while also considering state boarding schools, a number of which of which are Academies, may help us obtaining as treatment group one that is more representative of secondary school pupils than the highly selected one at CH. Finally, an extension that is particularly relevant to inform policy-decisions over the role of boarding education for high ability pupils with low SES is performing a cost-benefit analysis of subsidising these pupils.

References

Altonji, J. G., Elder, T. E. and Taber, C. R. (2005). Selection on Observed and Unobserved Variables: Assessing the Effectiveness of Catholic Schools. *Journal of Political Economy*, **113** (1), 151–184.

Becker, S. O. and Ichino, A. (2002). Estimation of Average Treatment Effects Based on Propensity Scores. *The Stata Journal*, **2** (4), 358–377.

Behaghel, L., de Chaisemartin, C. and Gurgand, M. (2017). Ready for Boarding? The Effects of a Boarding School for Disadvantaged Students. *American Economic Journal: Applied Economics*, **9** (1), 140–164.

Black, D. A. and Smith, J. A. (2004). How Robust is the Evidence on The Effects of College Quality? Evidence from Matching. *Journal of Econometrics*, **121** (1), 99–124.

Burgess, S., Briggs, A., McConnell, B. and Slater, H. (2006). *School Choice in England: Background Facts*. The Centre for Market and Public Organisation 06/159, Department of Economics, University of Bristol, UK.

—, Greaves, E., Vignoles, A. and Wilson, D. (2015). What Parents Want: School Preferences and School Choice. *The Economic Journal*, **125** (587), 1262–1289.

Chowdry, H., Crawford, C., Dearden, L., Goodman, A. and Vignoles, A. (2013). Widening Participation in Higher Education: Analysis Using Linked Administrative Data. *Journal of the Royal Statistical Society: Series A (Statistics in Society)*, **176** (2), 431–457.

Crawford, C., Macmillan, L. and Vignoles, A. (2014). Progress Made by High-attaining Children from Disadvantaged Backgrounds. *Social Mobility and Child Poverty Commission (Available from https://goo.gl/0emqZ8f)*.

Curto, V. E. and Fryer Jr, R. G. (2014). The Potential of Urban Boarding Schools for the Poor: Evidence from SEED. *Journal of Labor Economics*, **32** (1), 65–93.

Davies, P. and Davies, N. M. (2014). Paying for Quality? Associations Between Private School Income, Performance and Use of Resources. *British Educational Research Journal*, **40** (3), 421–440.

Dearden, L., Sibieta, L. and Sylva, K. (2011). The Socio-economic Gradient in Early Child Outcomes: Evidence from the Millennium Cohort Study. *Longitudinal and Life Course Studies*, **2** (1), 19–40.

Department for Education (2014). Boarding Schools. An Opportunity to Improve Outcomes for Vulnerable Children. *Department for Education (Available from https://goo.gl/T0rosq)*.

Department for Education (2016). Schools That Work for Everyone. Government Consultation. *Department for Education (Available from https://goo.gl/uYjmFc)*.

Epple, D., Romano, R. and Zimmer, R. (2016). Charter Schools: A Survey of Research on Their Characteristics and Effectiveness. In S. M. Eric A. Hanushek and L. Woessmann (eds.), *Handbook of the Economics of Education*, vol. 5, *3*, Elsevier, pp. 139 – 208.

Eyles, A., Hupkau, C. and Machin, S. (2016). School Reforms and Pupil Performance. *Labour Economics*, **41**, 9–19.

— and Machin, S. (2015). *The Introduction of Academy Schools to England's Education*. CEP Discussion Papers dp1368, Centre for Economic Performance, LSE.

Fisher, S., Frazer, N. and Murray, K. (1986). Homesickness and Health in Boarding School Children. *Journal of Environmental Psychology*, **6** (1), 35–47.

Frölich, M., Heshmati, A. and Lechner, M. (2004). A Microeconometric Evaluation of Rehabilitation of Long-term Sickness in Sweden. *Journal of Applied Econometrics*, **19** (3), 375–396.

Gerfin, M. and Lechner, M. (2002). A Microeconometric Evaluation of the Active Labour Market Policy in Switzerland. *Economic Journal*, **112** (482), 854–893.

Green, F., Machin, S., Murphy, R. and Zhu, Y. (2012). The Changing Economic Advantage from Private Schools. *Economica*, **79** (316), 658–679.

Hodges, J., Sheffield, J. and Ralph, A. (2016). Staff and Boarders Perspectives of the Boarding Environment. *Journal of Child and Family Studies*, **25** (4), 1045–1056.

Ichino, A., Mealli, F. and Nannicini, T. (2008). From Temporary Help Jobs to Permanent Employment: What Can We Learn from Matching Estimators and their Sensitivity? *Journal of Applied Econometrics*, **23** (3), 305–327.

Independent Schools Council (2014). *ISC Annual Census 2014 (Available from https://goo.gl/qktHh4)*. Tech. rep.

Jerrim, J. (2017). Comparing Socio-economic Gaps in the Performance of Highly Able UK Pupils Internationally. *Sutton Trust Report (Available from https://goo.gl/3nZ4Lw)*.

Kirabo Jackson, C. (2010). Do Students Benefit from Attending Better Schools? Evidence from Rule-based Student Assignments in Trinidad and Tobago. *The Economic Journal*, **120** (549), 1399–1429.

Lechner, M. (2002). Some Practical Issues in the Evaluation of Heterogeneous Labour Market Programmes by Matching Methods. *Journal of the Royal Statistical Society: Series A (Statistics in Society)*, **165** (1), 59–82.

Lester, L., Mander, D. J. and Cross, D. (2015). Bullying Behaviour Following Students' Transition to a Secondary Boarding School Context. *International Journal of Child and Adolescent Health*, **8** (2), 141.

Leuven, E. and Sianesi, B. (2015). PSMATCH2: Stata Module to Perform Full Mahalanobis and Propensity Score Matching, Common Support Graphing, and Covariate Imbalance Testing. *Statistical Software Components*.

Machin, S. and McNally, S. (2005). Gender and Student Achievement in English Schools. *Oxford Review of Economic Policy*, **21** (3), 357–372.

—, — and Meghir, C. (2004). Improving Pupil Performance in English Secondary Schools: Excellence in Cities. *Journal of the European Economic Association*, **2** (2-3), 396–405.

Martin, A. J., Papworth, B., Ginns, P. and Liem, G. A. D. (2014). Boarding School, Academic Motivation and Engagement, and Psychological Well-Being: A Large-Scale Investigation. *American Educational Research Journal*.

Reardon, S. F. (2011). The Widening Academic Achievement Gap Between the Rich and the Poor: New Evidence and Possible Explanations. *Whither opportunity*, pp. 91–116.

Schaverien, J. (2004). Boarding School: the Trauma of the Privileged Child. *Journal of Analytical Psychology*, **49** (5), 683–705.

— (2011). Boarding School Syndrome: Broken Attachments a Hidden Trauma. *British Journal of Psychotherapy*, **27** (2), 138–155.

Shu, B. and Tong, Y. (2015). Boarding at School and Students' Well-Being: The Case of Rural China. In *Population Association of America 2015 Annual Meeting*, vol. 30.

Sirin, S. R. (2005). Socioeconomic Status and Academic Achievement: A Meta-Analytic Review of Research. *Review of Educational Research*, **75** (3), 417–453.

Todd, P. E. and Wolpin, K. I. (2003). On the Specification and Estimation of the Production Function for Cognitive Achievement. *The Economic Journal*, **113** (485), F3–F33.

— and — (2007). The Production of Cognitive Achievement in Children: Home, School, and Racial Test Score Gaps. *Journal of Human Capital*, **1** (1), 91–136.

Wang, A., Medina, A., Luo, R., Shi, Y. and Yue, A. (2016). To Board or Not to Board: Evidence from Nutrition, Health and Education Outcomes of Students in Rural China. *China & World Economy*, **24** (3), 52–66.

Watson, G. (2006). Computing Helmert transformations. *Journal of Computational and Applied Mathematics*, **197** (2), 387 – 394.

Whitty, S. P., G. and Edwards, T. (1998). The Assisted Places Scheme: Its Impact and Its Role in Privatization and Marketization. *Journal of Education Policy*, **13** (2), 237–250.

Wires, J. W., Barocas, R. and Hollenbeck, A. R. (1994). Determinants of Adolescent Identity Development: A Cross-sequential Study of Boarding School Boys. *Adolescence*, **29** (114), 361.

Ivie

Guardia Civil, 22 - Esc. 2, 1º
46020 Valencia - Spain
Phone: +34 963 190 050
Fax: +34 963 190 055

Website: www.ivie.es
E-mail: publicaciones@ivie.es