Out-of-school-time study programmes: do they work?

Nicola Pensiero & Francis Green

Routledge
Taylor & Francis Group

OPEN ACCESS

Out-of-school-time study programmes: do they work?

Nicola Pensiero and Francis Green

UCL Institute of Education, UK

ABSTRACT

We analyse the prevalence and effectiveness of out-of-school-time (OST) study programmes among secondary aged students, focusing on their potential for reducing socio-economic gaps in educational achievement. Compared to several extant studies, including the only prior study for Britain, whose findings could be affected by heterogeneous participation in the programmes, our results derive from a rich dataset with multiple controls for social background, personal motivation, and school characteristics. We find that programme participation in England is relatively low among students from families with long-term unemployed parents and those in routine occupations. Participation is also lower outside London, and especially outside large cities. Our results show that OST programmes, as long as they are teacher-led, are moderately effective in improving the academic performance at the end of lower secondary education as measured by the GCSE total score. Teacher-led OST programmes compensate for previous social disadvantage. The policy implications include a focus on expanding programme availability and on incentives for participation, and attention to the regional disparities.

1. Introduction

In this paper we examine the effectiveness of out-of-school-time (OST) study programmes in improving academic performance at secondary level in England. It is well-established that the quality and intensity of non-school-time activities and parenting are related to children's academic performance and account, at least partly, for the social class gap in academic performance that is found in many countries (Apsler, 2009; Bonke & Esping-Andersen, 2011; Pensiero, 2011; Stafford & Yeung, 2005). Contemporary research on parenting and after-school time has identified two general trends. First, there has been an increase in both parents' child care time and expenditure on child development, especially among highly educated parents. Second, parents have become increasingly aware of the importance of structured activities for their children's development and engage their children in multiple cultivating activities, again with differences between social classes (Bianchi, Cohen, Raley,

---

CONTACT Nicola Pensiero ✉ n.pensiero@ucl.ac.uk

& Nomaguchi, 2004; Bianchi, Robinson, & Milkie, 2006; Hsin, 2008; Lareau, 2002; McLanahan, 2004).

In this light, programmes designed to support after-school learning activities assume considerable importance for education and social policy. For many years, out-of-school-time (OST) study programmes have been provided in US schools, targeted on disadvantaged children, and these have been variously evaluated (for a review see Apsler [2009] and Lauer et al. [2006]). Relatively little is known, however, about their effectiveness elsewhere, either in raising academic performance generally or in compensating for the social disadvantages that limit other forms of out-of-school learning activity. Similar programmes, available to children from all backgrounds, have also been provided in most British secondary schools. We consider not only the overall average effects, but also the extent to which they can compensate for the social class investment gap in children's education: we test whether the OST programmes are more beneficial for lower-achieving children and children with lower socio-economic origins.

The relevance and potential effectiveness of school-based OST study programmes rests on the proposition that schools, by providing greater structure and well-qualified staff, can compensate for the inadequacy of the learning environment at home and provide a channel to promote social mobility. By delivering high quality, well-resourced activities to those disadvantaged children who otherwise could have no or limited access to them, it can contribute to closing the learning opportunity gap between children of different socio-economic backgrounds. Choosing the programmes to prioritise in order to raise the achievement of disadvantaged children has become a key preoccupation of schools. Schools allocate significant resources to OST programmes: at the time of the study 55% of students attended some form of curriculum-based OST programme, costing on average £7 per session per pupil.[1] Hence knowing whether and how far OST programmes are successful in raising achievement of children in general and of disadvantaged children in particular should be a useful aid for schools' spending choices.

The implementation of some form of OST programme became widespread (though not ubiquitous) among secondary schools in the UK around the beginning of the 2000s (MORI 2004), but surprisingly there has not yet been a large scale investigation of the effectiveness of these programmes. Compared to the core hours of compulsory education which are largely teacher-directed, the OST programmes are voluntary, learner-centred, favour a greater sense of control for both teachers and students, and are characterised by a more relaxed and informal relationship between teachers and students. Only one UK study (MacBeath et al., 2001) has addressed their effectiveness, but this was focused on a small, unrepresentative sample dating back to the late 1990s when the introduction of OST programmes was in its initial experimental phase, and did not examine selection issues. Reliance on extrapolation of findings from US studies also has limitations, not least due to the differing cross-national context and scope of the programmes. A study of OST programmes in the UK is therefore overdue.

This paper uses Next Steps, formerly known as the Longitudinal Study of Young People in England, to study participation in OST programmes, and their effectiveness in improving GCSE performance. Next Steps comprises a multi-wave cohort survey of pupils in England born between 1 September 1989 and 31 August 1990 and their parents (or carers).[2] GCSE is the high-stakes national exam taken by most English school pupils around age 16, and is a gateway for differentiated transitions into further academic or vocational education or work.

We chose Next Steps owing to the richness of the available information in this survey regarding parental resources, children's behaviours, attitudes, prior achievement, school-level characteristics, and OST activities, together with its longitudinal character. Using appropriate statistical techniques—propensity score matching and school-fixed effects regression—our objective is to find credible estimates of the impact of OST programmes on subsequent GCSE scores.

The paper proceeds with Section 2 which overviews existing research, mainly from the US, on the effectiveness of OST programmes. Section 3 sets out our analytical strategy and describes the Next Steps data. All our findings are presented in Section 4, and Section 5 discusses the policy implications.

2. OST programmes and academic performance

Research on OST programmes in the US indicates that participation improves academic performance as intended, with low to moderate gains in mathematics and reading (Apsler, 2009; Durlak, Weissberg, & Pachan, 2010; Lauer et al., 2006). Given that the students who participated in OST programmes in the studies were at risk of school failure, researchers have regarded even small improvements in academic performance as a welcome indicator of positive outcomes for such students (Miller, 2003). The finding is all the more encouraging because, as supplements to regular school learning, the programmes come at a low cost relative to most intensive interventions offering a broader curriculum/scope and targeting also the parents, such as the US Abecedarian project. An effect size of 0.10 to 0.20 of a standard deviation is typical of remedial programmes (Cooper, Charlton, Valentine, & Muhlenbruck, 2000, Lipsey & Wilson 1993). However, OST programmes are diverse in terms of focus and goals, and the estimated impact does vary across studies, suggesting that the quality and content of the programme are important. Notably, OST activities that focus on regular academic programmes in schools are more robustly related to academic performance (De Kanter, 2001; Huang, Gribbons, Kim, Lee, & Baker, 2000; Miller, 2003).

Several researchers have suggested that the effectiveness of OST might depend on the grade levels of the students. OST programmes are beneficial for reading achievements for students in both elementary and secondary grades, while benefits for mathematics achievement occur primarily in the secondary grades (Lauer et al., 2006). Older children appear to be more difficult to recruit than younger children (Grossman, Walker, & Raley, 2001). The duration of programmes is also identified as a potential factor accounting for OST effectiveness (McComb & Scott-Little, 2003). The minimum duration required for an OST programme to be effective is estimated to be roughly 45 hours. However, longer OST programmes do not necessarily have more positive outcomes (Ascher, 1990; Karweit, 1985; Lauer et al., 2006). It is also important to consider how students are grouped. The largest effect sizes are found in programmes run for small numbers of students and those that provided more individualised and small-group instruction (Cooper et al., 2000; Elbaum, Vaughn, Hughes, & Moody, 2000; Fashola, 1998).

Besides an overall positive effect, it is also predicted that OST programmes should have an equalising effect. Not only were many US OST programmes targeted at lower socioeconomic groups, it is expected that the programmes should be more effective for those groups, owing to the stratified learning opportunities of children from different socioeconomic backgrounds. Thus, disadvantaged children are less likely to have significant adults

to supervise them after school. Parents at the bottom of the occupational stratification have fewer career prospects and lower job quality, which in turn increase the chances of family disruption. Finally, disadvantaged children are offered fewer opportunities from within their families and communities to be involved in cognitively and academically stimulating structured activities. These factors suggest that disadvantaged children may have more risky and less stimulating out-of-school-time hours within families, and hence that participation in OST study programmes might be particularly effective in improving the quality of this time for this group of children. This expectation is confirmed in a few studies. Thus in comparison with middle-income children, low-income children are more likely to benefit from OST programmes (Cosden, Morrison, Albanese, & Macias, 2001; Miller, 2003) and low-achieving students tend to benefit more than students who entered programmes with higher achievement (McComb & Scott-Little, 2003).

There is only one published study of the effectiveness of OST programmes in secondary schools in England and it found that such programmes are associated with an improvement in GCSE and Key Stage 3 (KS3) performance (MacBeath et al., 2001). The study was set up in the framework of 'The Study Support National Evaluation and Development Program' (SSNEDP 1997) and followed two cohorts of students tracked for three years: 6000 seniors from Year 9 to Year 11 and 2000 juniors from Year 7 to Year 9 drawn from an opportunity sample of 51 secondary schools in disadvantaged areas in those Local Education Authorities that were willing to make a three year financial commitment towards the costs of the developmental aspects of the SSNEDP. The study found that compared to students who did not participate, students who participated in the programmes improved substantially their GCSE performance, as measured by best five scores, on the number of A–C passes, and on maths and English GCSE (by half a grade). Although sport and aesthetic activities had positive effects on attainment, OST programmes related to the curriculum, drop-in sessions, and Easter revision courses had the strongest effects. The study also found that students from minority groups participated more in and benefited more from OST programmes than white students.

These findings might be questioned, however, not only because the sample was non-random, but also because the study does not present a discussion of the factors influencing the children's participation in the programme. It is not clear how participants were different from the non-participants, and hence whether selection biases were present. Moreover the study was conducted at the end of the 1990s when the introduction of these programmes was in its initial tentative phase; it is possible that subsequent learning may have improved their effectiveness.

The questions we wish to examine, therefore, are how wide is participation in OST programmes in England, how does participation vary across socio-economic groups, and how effective are they (if at all) in improving educational outcomes in more recent years, now that they have been running for a considerable time. Moreover, how does their effectiveness vary among types of programmes and types of pupils?

3. Analytic approach

3.1. Data and evaluation methods

While it would be best to address the evaluation question through the use of a randomised controlled trial (RCT), this option is not currently open. Instead we base our evaluation on observational data drawn from the high-quality longitudinal Next Steps study. In the Next

Steps survey schools were the primary sample units and the sample size per school is 30 pupils on average, with a total sample of 15,800 pupils in wave 1 when the students were aged 13/14. The survey covers participation in OST programmes, as well as very rich information on the factors which might account for participation in the programmes and on academic achievement: social origins (parental education, social class, family income, deprivation of area), and individual factors (expectations, school engagement, frequency of homework). Next Steps is also linked to National Pupil Database (NPD) records, which include the cohort members' educational outcomes, as well as prior academic achievement, and school-level characteristics.

As with any evaluation using observational data, the challenge is to estimate a comparison of participant outcomes with credible counterfactual outcomes that would have been realised had the person not participated in the OST programme (Rosenbaum & Rubin, 1983). The estimation obtained simply by comparing a group of units exposed to a programme with a comparison group that is not exposed to the programme can be biased, because selection into the programme is likely to occur on the basis of non-random processes that are linked to outcomes of interest. Biases can arise, for example, if programme participants are differently engaged with school work or have different prior abilities from non-participants, and if these differences influence academic performance over and above the participation in the programme.

Here we address the evaluation problem with alternative techniques. In addition to standard linear regression, we derive a comparison group using Propensity Score Matching (PSM), utilising a set of pre-treatment covariates, where in our case 'treatment' is defined as participation in an OST programme (Dehejia & Wahba, 2002; Gerfin & Lechner, 2000; Heckman, Ichimura, & Todd, 1997). We also estimate a regression model including school fixed effects.

PSM has two stages: first, the estimation of a propensity score capturing the propensity to participate in the programme; second, the use of that score to define a comparison group for participants, which by its characteristics is similar to the treatment group. Thus, using PSM, the study sample (both treatment and control group) is restricted to those young individuals who had a similar prior probability of participating in the programme. Once the comparison group is constituted, we estimate the average treatment effect on those treated (ATT), and the underpinning assumption is that treatment is ignorable conditional on the observed covariates (this is termed the mean conditional independence assumption [CIA]). An important condition for satisfying this assumption is that participation is independent of the potential outcomes, conditional on observed covariates (Rosenbaum & Rubin, 1983); the validity of this depends on the quality and relevance of observed covariates available. It can be expected that schools make selective use of their resources by targeting specific types of students. According to a survey conducted by MORI (2004) commissioned by the Department for Education and Skills, more able students, students with learning difficulties, students at the threshold of the next level of achievement, and disaffected students are the types most frequently targeted by schools in designing OST programmes. These findings imply that, for the CIA to be tenable, the model of the propensity to participate in OST programmes should incorporate prior achievement and motivation.

There are two main differences between PSM and linear regression. First, PSM does not assume a linear (or any other) functional form in estimating the impact of OST programmes. Second, PSM only compares the outcomes of treated individuals with non-treated individuals in a 'support group' with similar probabilities of participation, excluding those who are

substantially different. If the support group cannot be found, the evaluation is said to fail and is not carried out.

Nevertheless, just as OLS regression controls for only the observed covariates, PSM can only match on observed characteristics. Despite the relevance and richness of observed data, there might be unobserved individual or school-level variables that have an impact on children's performance independently of the programmes. Potentially relevant school characteristics include teachers' quality and students' average ability. For example, it could be conjectured that the schools which are most effective in recruiting students to OST programmes might be the ones with the best teachers. We can include some observed school-level characteristics, drawn from the school census, in both the linear regression and the PSM analysis. As a complementary approach, aiming to remove bias arising from unobserved school-level characteristics, we also run school-fixed effects regression models, which can net out the effect on outcomes of factors operating at the school level.

Survey weights are used in both the descriptive statistics, OLS, and the estimation of the propensity score to take into consideration that more deprived schools and pupils from minority ethnic groups have been oversampled by survey design. Standard errors are adjusted to take into consideration the clustering of individuals in schools.

3.2. Indicators

The Next Steps survey contains specific information regarding the different kinds of curricular activities in which the child is involved outside lessons but within the school setting. We focus on programmes that are linked to the academic curriculum because previous research in the US has suggested these are the most beneficial for academic achievement and in particular on:

(1) Teacher-led study groups.[3] These are activities in which students work together to prepare for examinations, i.e. GCSE coursework, and to do/review homework. The role of the teacher in these activities involves a combination of supervision and instruction.

(2) Self-directed study clubs (also called drop-in sessions) in which the students do homework and GSCE coursework together, without the teacher.[4]

Pupils were asked at age 13/14 (wave 1, in 2004), 14/15 (wave 2), and 15/16 (wave 3) whether their school had such programmes and, if yes, how frequently they attended each programme (from never [0] to five times a week or more [5]). The first two waves are the most relevant as they precede GCSE examinations.

The achievement outcome measure is the total capped General Certificate of Secondary Education (GCSE) score. During Year 11 at age 15/16, students sit Key Stage 4 examinations to obtain their lower secondary certificate of education (GCSE). As pupils take different numbers of courses, performance is compared using capped point scores. This measure caps the total number of included courses at eight best GCSEs or equivalent, because the vast majority of pupils take at least this number of courses. It is created by taking Grade G, the lowest grade achieved, to be 16 points; each grade improvement thereafter, for example from G to F, C to B, or A to A*, is equivalent to an additional six points. The uncapped measure of GCSE score is used as a robustness check.

Prior academic achievement, an important control variable, is well captured through the administrative records of individual scores at education's Key Stage 3.[5] A further big advantage of Next Steps is its rich set of available control variables covering social background. Parental class is coded using the National Statistics Socio-Economic Classification (NS-SEC). In cases where both parents are employed we select the highest of the parents' class category in line with the dominance approach (Erikson, 1984). Parental education is defined as the highest academic qualification of either parent (dominance approach). It has been shown that the commonly used qualifications variable, which treats vocational and academic qualifications (NVQ) as equivalents, has less predictive power of children's educational outcomes than a variable giving prominence to academic qualifications (Sullivan et al., 2013). Accordingly we used an academic-qualification-based definition of qualifications: less than GCSE graded D–G, GCSE graded D–G (level 1), O/GCSE (level 2), A-levels (level 3), degree level qualification (level 4 and higher). We also use four other indicators of socio-economic circumstances: a dummy variable indicating whether the child lives in a two parent or single parent family; self-reported ethnic origins; eligibility for free school meals (a standard measure of social deprivation, using an administrative source); and the local area multiple deprivation index. In another specification not shown, we included family income as an additional measure of socio-economic background, but it does not have a significant independent impact on children's academic performance and the estimates of the effectiveness of the OST programmes were not different from the ones obtained excluding family income. Therefore, we decided to exclude family income from the presented analysis.

We include a measure of school engagement, which is a recognised indicator of student's motivation (Fredricks, Blumenfeld, & Paris, 2004). It is computed as the first principal component (eigenvalue 4.2) of the scaled responses to the following questions regarding the perception of schooling: 'I am happy when I am at school', 'school is a waste of time for me', 'school work is worth doing', 'most of the time I do not want to go to school', 'on the whole, I like being at school', 'I work as hard as I can in school', 'in a lesson, I often count the minutes till it ends', 'I am bored in lessons', 'the work I do in lessons is a waste of time', 'The work I do in lessons is interesting to me'. The items' scale ranges from 1 (strongly agree) to 4 (strongly disagree). The child's educational plan is included as a dummy indicator of the intention to continue education after age 16 (the end of compulsory school [Strand, 2011]), along with a variable indicating eligibility for a special education programme (using an administrative source), and the frequency of homework ranging from none (0) to five times per week (5).

School level factors are census averages drawn from the NPD data and linked to Next Steps members.[6] Those factors comprise the percentage of students eligible for free school meals (an indicator of socio-economic disadvantage), the percentage of students whose first language is not English, the percentage of students attaining level 2 qualifications (from the NPD). The indicator of programme availability is derived from Next Steps. While programme availability in the school could in principle be a dummy variable, assuming that the programme is open to all students, we suspect that it could be measured incorrectly with a downward bias in some cases, in that some students might not be aware of programme availability if they do not themselves take part. As a proxy for whether each student would have a possible programme to join, therefore, we include as our control the proportion of students affirming that a given OST programme is available in the school.

Finally, we introduced two indicators of the broader social context: the local area's multiple deprivation index (which encompasses income, employment, education, health, crime, and

Table 1. Participation in OST study programmes at age 14/15 (Year 10) by children's social and personal characteristics.

	Teacher-led study group	Self-directed study club
Social class (NS-SEC)		
VIII: Long term unemployed	0.42	0.28
VII: Routine	0.40	0.28
VI: Semi-routine	0.40	0.28
V: Lower supervisory	0.45	0.29
IV: Small employers	0.45	0.33
III: Intermediate	0.44	0.32
II: Lower managerial	0.46	0.34
I: Upper managerial	0.46	0.35
Free school meal		
No	0.44	0.31
Yes	0.43	0.27
Marital status of parents		
Single	0.41	0.29
Married	0.45	0.32
Ethnic origins		
White	0.43	0.31
Mixed	0.46	0.36
Indian	0.55	0.43
Pakistani	0.48	0.36
Bangladeshi	0.56	0.38
Black Caribbean	0.55	0.41
Black African	0.62	0.45
Gender		
Male	0.42	0.33
Female	0.46	0.30
Prior achievement (KS3)		
I quintile	0.33	0.27
V quintile	0.47	0.34
Urban/Rural		
< than 10K population	0.40	0.31
≥10K population (No London)	0.43	0.30
London	0.52	0.39
Total	0.44	0.32
N	(5098)	(3596)

housing), and an indicator of the size of the town the child lives in, which distinguishes between small towns (<10K), large towns (≥10K, with the exclusion of London), and London.

4. Results

4.1. OST programme participation: a description

Overall, 44% of our sample participated in teacher-led OST groups at age 14/15, while 32% took part in self-directed study groups. Table 1 provides a description of how participation varies across socio-economic groups. Social class of origin is associated to some extent with participation in both types of programmes: there is a (statistically significant) participation gap between the routine, semi-routine, and unemployed classes together and all the other classes. Participation in self-directed study clubs is lower by four percentage points for those with free school meal status, but participation in teacher-led study groups is virtually the same. Non-white students participate more than white students, with the highest rate (62%) shown for black African students. Students from two-parent families and with a higher prior academic achievement in KS3 show higher participation rates in both types of programmes.

Girls participate more often than boys in teacher-led study groups, but less often in self-directed study clubs. Finally, students living in London have much higher participation rates in both types of programmes than students elsewhere, including those in other large cities. Students in smaller cities (or rural environments) do not show significantly different participation rates from students in larger cities with the exception of London. Though not shown in the table, the lead in participation rates for London is especially noticeable for those in Social Classes VII and VIII, where 55% join teacher-led study groups, more than the other classes.

4.2. Propensity score matching

For the first stage of the analysis, the estimation of the determinants of participation in OST programmes, we used a probit regression model (Rosenbaum & Rubin, 1983). Drawing on previous research, the covariates include predicted antecedents of both participation in OST programmes and GCSE performance (Austin, Grootendorst, & Anderson, 2007). The detailed findings are presented in the Appendix 1.

The propensity to participate is higher among highly engaged students and students who do more homework. Having more ambitious educational plans is also related to higher participation, though only in the case of teacher-led study groups. While behavioural attributes are important determinants of participation, prior achievement has no effect on participation in teacher-led study groups. Given that these factors have been controlled for, other socio-economic and ethnic background characteristics are less important in accounting for participation. Moving to school characteristics, students attending schools with a higher proportion of economically disadvantaged and with a higher proportion of students attaining level 2 qualifications tend to participate more in self-directed study clubs. Our proxy for programme availability also has the expected strong impact on participation. Finally, as for regional differences in participation, the pre-eminence of London noted above in the description of participation is reproduced here, but in addition, it emerges that there is also a smaller advantage of other large towns over small towns or rural areas with populations of less than 10,000: after controlling for many other potential determinants of programme participation, students living in large towns, even more so in London, tend to participate significantly more in teacher-led study groups.

For the second stage, we specify matching individuals for participants. There is no single matching estimator that is valid in all situations (Dehejia & Wahba, 2002). Here, where we have a sufficiently large sample, we use 'radius matching', in which each individual is compared with all those within a caliper of a given radius. Thus, participants were matched on the logit of the propensity score using calipers of width equal to 0.2 of the standard deviation of the logit of the estimated propensity score (Rosenbaum & Rubin, 1985).[7] To check that our results were robust, the analysis was then repeated using different estimators, namely kernel matching and stratification on the propensity score, obtaining very similar results.

4.3. The effectiveness of OST programmes

4.3.1. Overall effect

In Table 2 we present our first key findings about whether OST programmes work. The mean effects of attending teacher-led and self-directed study clubs, as estimated by OLS linear

Table 2. The effect of OST programmes at age 14/15 (Year 10) on academic performance (GCSE scores).

	Teacher-led study group	N	Self-directed study club	N
OLS	3.6** (1.4)	6019	−0.03 (1.3)	6019
PSM	3.2* (2.1)	6004	−1.64 (2.4)	6004
SFE	2.7** (1.4)	6019	−0.19 (1.4)	6019

Notes: ***$p<0.01$; **$p<0.05$; *$p<0.1$, SE in parentheses; OLS: Ordinary least squares (controls: KS3, Sex, Educational plans, School engagement, Homework, SEN status, Two-parent family, Social class, Parental education, FSM, Urban/London/rural, Ethnic origins, multiple deprivation index, % FSM [school], % first language not English [school], % students saying programme is available [school], % Lev 2 [school]).
PSM uses radius matching, caliper: 0.2 of SD of PS, Bootstrap; for PS model see Appendix.
SFE: school fixed effects.

regression with all above-listed covariates as controls, are given in the first row. As can be seen, teacher-led programmes raise GCSE achievement score by an average of 3.6, while the effect of self-directed study groups is small, negative, and statistically insignificant.

The second row shows the result of the PSM estimation, where the parameter of interest is the mean effect of attending OST programmes, or the average treatment effect on the treated (ATT), estimated in the region of common support. Again, the improvement in GCSE achievement associated with teacher-led OST programme attendance is positive (at 3.2 points), while it is non-significant in the case of self-directed study groups.

As a robustness check, the analysis of the programme average effectiveness was repeated using the uncapped measure of GCSE achievement and the same matching method. The results, which are not reported, should be interpreted with caution as the uncapped measure conflates the number of GCSEs taken and the grades attained.[8] Results show even larger estimates of the effectiveness of teacher-led OST programmes across the different techniques (12 points), while estimates regarding the self-directed study group are non-significant.

A three points average gain is modest, being equivalent to half a grade in one subject out of the eight best GCSE results used in the computation of the total GCSE score. This magnitude is smaller than the one found by the previous UK study carried out in the 1990s (MacBeath et al., 2001), which amounted to three and a half grades on the Best 5 score, but more in line with the moderate effectiveness found in US studies by looking only at disadvantaged children. The other indication emerging from the results is that OST programmes are effective when they are teacher-led, but that one cannot reject the presumption that self-directed study groups have no effects. In contrast with the study of MacBeath et al. (2001), which found that OST programmes were more effective as long as they focused on the curriculum, our results suggest that the academic focus of the programme is not a sufficient condition for the programme to be effective and that the leadership of the programme is also important.

While the PSM method has the advantage of being non-parametric, it is notable that the estimated effect of teacher-led study programmes is quite close to the OLS estimate. Yet it remains possible that unobserved characteristics of the school environment play a role. The third row shows the estimated effects obtained from a school-fixed effects regression model. The estimated effect size remains significant. Though it is smaller than the estimates obtained using PSM and OLS, the fact that this estimate is not radically different from those of the PSM and OLS models suggests that the effects of school characteristics may already have been reasonably well captured by the included indicators.

4.3.2. Effects of OST programmes for different socio-economic and achievement groups

The second research question is whether OST programmes are effective at reducing the achievement gap between children from different socio-economic backgrounds or with different prior academic achievements. For the purposes of this analysis, social class is grouped in four categories, combining class categories in pairs, to achieve a sufficiently large sample size. It was hypothesised that those from more disadvantaged backgrounds would have more to gain. The argument was that a lack of economic and cultural resources implied that their effective use of non-school time was much below that of advantaged socio-economic groups, and that teacher-led OST programmes could be a more effective substitute for them than they would be for socially-advantaged children. Thus we expect that material deprivation and disadvantage associated with lower social origins negatively influences the quality and quantity of the effective learning time spent at home once children finish school, especially compared to upper middle class children. Child poverty is known to have severe detrimental consequences for a child's cognitive achievement, and the UK child poverty rate—14% in the mid-2000s—is high compared to Nordic and continental European countries. The connection of poverty with class in the Next Steps sample is strong. Those from routine occupations and long-term unemployed households have strikingly higher chances of being materially deprived than the remaining classes: more than half (56%) receive free school meals compared with 16% for semi-routine and lower supervisory households, 8% from lower supervisory and small employers households, and 2% from upper and lower managerial households. Similarly, children in the lowest socio-economic group are much more likely to be living in single-parent households, with consequent scarce time resources for aiding child learning. Taken together, there is ample reason to expect that, just as earlier studies have shown for the United States, and despite the different context, OST programmes in Britain should also have an equalising effect.

Table 3 presents propensity score matching estimates that address this question, showing the effectiveness of the OST programme separately for each group of social class. The effectiveness of teacher-led study groups is, as predicted, related to the class of origin of the student. Children from long-term unemployed parents or routine occupation households benefit more than all other groups from attending teacher-led OST programmes. For this group the improvement of 11 in GCSE scores amounts to two grades higher for one GCSE (e.g. going from a D to a B) or two grade improvements for two GCSEs, e.g. two As instead of two Bs (11 points)—substantially greater than the average effect reported above. By contrast, there is no statistically significant effect for the more advantaged social classes.

We used the uncapped measure as a robustness check. The unemployed and routine classes are still the greatest beneficiaries of those programmes (13 points improvement).

Table 3. Propensity score matching effect of participation in teacher-led study group at age 14/15 (Year 10) by social class: a model for each group.

Social class		N
VIII & VII	10.7* (6.9)	759
VI & V	−2.8 (5.0)	1311
IV & III	−0.1 (5.5)	1004
II & I	2.3 (2.6)	2937

Notes: ***$p<0.01$; **$p<0.05$; *$p<0.1$; for further notes, see Table 2.

Table 4. Propensity score matching effect of participation in teacher-led study group at age 14/15 (Year 10) by quintiles of KS3 total score: a model for each group.

Prior academic performance (KS3)		N
KS3 I & II Q	5.1 (5.4)	1449
KS3 III Q	1.3 (2.5)	1255
KS3 IV Q	5.5** (2.6)	1547
KS3 V Q	0.5 (1.8)	1726

Notes: ***$p<0.01$; **$p<0.05$; *$p<0.1$; for futher notes, see Table 2.

The managerial classes show similar improvements, while the routine and lower supervisory classes benefit less (seven points). Small employers and intermediate occupations do not show a significant improvement with the uncapped measure.

The analysis of OST programme by sub-group of prior academic performance is conducted using quintiles of KS3 scores—see Table 4. The first two quintiles are merged in one group because the sample size in the region of common support for the first quintile is too small (<600).[9] The students with a low prior performance (quintiles I and II) show a five-point gain in GCSE performance associated with OST programme attendance, although the effect is not statistically significant. Students scoring in the III and V quintiles of prior academic performance do not benefit from OST programmes, while students in the IV quintile show a significant five-point improvement when they attend OST programmes. Using the uncapped measure all groups show equally a significant improvement, which is also higher than the one obtained using the capped measure (around 10 points).

The analysis of the effectiveness of OST programme by sub-groups of prior academic performance and social class has been replicated using the OLS and school fixed effects models. In both models an interaction term between programme participation and social class (or prior academic performance) was introduced in order to inspect the extent to which the programme is more beneficial to lower class (least achieving) children compared to higher class children. The results from both models, which are not shown here, confirm overall the findings from the sub-group analysis discussed above. This indicates a high level of robustness of the sub-group results in respect to the different modelling strategies used.

We have presented the results so far only on teacher-led study groups. Even though we found no overall effect of self-directed study programmes, it is of course possible that they are effective for some socio-economic classes or prior achievement groups. However, in results not shown above we found no sub-group differences in the effectiveness of the self-directed study groups: they had no significant effects for any group.

As a further extension, we have also analysed the effect of the intensity of exposure of OST programmes. Focusing on teacher-led study groups, among the students who participate in the programme, 49% participate less than once a week, 41% once or twice a week, while 9% participate three times or more. In line with previous research in the US, in results (not shown, available on request) we found that attending the programme more than occasionally does not seem to improve GCSE total score. There are also no incremental effects: participation in the previous year (Year 9) does not affect achievement later on. What matters is the most recent attendance in the OST programme. There are also no cumulative effects in respect of other types of OST programmes and activities such as homework. That is, programmes' effectiveness does not vary in relation to other type of activities the child is engaged in after school time. Students who more frequently do their homework and

participate in other OST activities are not advantaged when attending teacher-led study groups. Our interpretation of these findings is that the goal and effect of the teacher-led programme, whether overall and for disadvantaged groups, is to complement and consolidate the learning processes of topics that had been already presented during normal school hours, in which case their marginal efficacy drops relatively rapidly.

5. Conclusions

This study has analysed the effectiveness of OST study programmes among secondary aged students. Compared to several extant studies, including the only prior study for Britain, whose findings could be affected by heterogeneous participation in the programmes, the present results derive from a rich dataset with multiple controls for social background, personal motivation, and school characteristics. We have used appropriate methods to construct credible counterfactual estimates of programme participation and to compare the programme effectiveness across students with differing socio-economic class and prior academic achievements. The data are from English secondary schools, where a large-scale investigation of OST programme is lacking and where these programmes are widely available. Unlike studies in the US, where most of the programmes target students at risk of academic failure, the results can be generalised to the overall population. Moreover, they enable us to look at the compensatory effects of OST programme with respect to social and academic disadvantage through sub-group analysis.

The results show that OST programmes, when they are teacher-led, are moderately effective in improving the academic performance at the end of lower secondary education as measured by the GCSE total capped score (age 16). By contrast, when children attend self-directed OST programmes, there is no statistically significant effect; arguably, the reason for this might be that children, when left unsupervised, tend to work less effectively than when they are guided by trained teachers. This conclusion is consistent with US research showing that the best programmes are those which provide greater structure, a stronger link to the curriculum, well qualified staff, and small group settings (Apsler, 2009; Durlak et al., 2010; Lauer et al., 2006; Miller, 2003).

The analysis has also explored whether the OST programme can compensate for previous disadvantage and reduce the achievement gap between low and high achieving students and between children from differing socio-economic groups. The sub-group analysis showed that children from parents who are unemployed or in a routine occupation benefit the most. Indeed, we found no evidence that children from higher classes benefit significantly from participating in OST programmes. Attending the programme does not seem to reduce overall achievement gaps according to previous academic performance. Using the uncapped measure of GCSE achievement, the analysis showed that children from an unemployment or routine background and children from a managerial background equally benefit from the programme, while other classes benefit significantly less.

The main caveat to these findings is that, although Next Steps contains rich data on the factors that are expected to account for attendance in and returns to OST programmes, we cannot exclude that our estimates of the effectiveness of OST programmes remain affected by some degree of bias. The main biases might arise as a consequence of the omission of relevant personal factors or of the failure to measure properly the relevant factors. While we are confident that, more than with earlier studies, we have extensively covered known

relevant personal, social origins, and school-level factors, it remains possible that the used indicators capture those factors with some error, and that relevant unknown factors are excluded.

Notwithstanding these caveats, the findings are relatively optimistic, compared with the view that compensatory educational interventions in late adolescence are highly costly or ineffective, owing to a presumed low level of skill malleability among adolescents (Cunha & Heckman, 2007; Heckman, 2006). Our results suggest that, even among 16 year old students, it is possible, by investing moderate resources, to compensate partially for a disadvantageous home learning environment and previous low performance. While OST programmes do not remove disadvantage, they do appear to make a difference, and the estimated effect of the teacher-led OST programmes for the lowest socio-economic group is large enough to warrant policy-makers' and schools' attention. The programmes require qualified and well-trained staff, but they are only moderately expensive. The Education Endowment Foundation estimates that OST programmes cost, on average, £7 per session per pupil.[10] Given that most participants attend either 'occasionally' or 'once or twice a week', one might envisage attendance for 25 weeks in the course of a 39-week school year, giving a roughly-estimated cost of £175 per pupil year. The return for the unemployed and routine occupation group, as estimated in this study using PSM, is an improvement of 11 points on their GCSE score, equivalent to about two grade increases out of eight best GCSEs—assuredly worth having.

There are thus three indicative policy implications. First, with only 42% of children from a social background of unemployed parents and 40% of those with a routine socio-economic background participating in OST programmes, there would seem to be some scope for improvements. It should be possible to ensure that OST programmes are available in all schools that take in children from families with long-term unemployed and routine class parents. With further encouragement and incentives for participation among these groups, a notable contribution could be made to reducing the socio-economic gap in achievement.

Second, the regional differences we have found also suggest that there is a potential for improving participation in small cities and rural areas. Students living in large towns, and particularly in London, tend to participate significantly more in teacher-led study groups. Not only is the participation in London higher, but also the participation gap between social classes, which in general favours the advantaged classes, is overturned within London. There, children from routine and unemployed classes participate significantly more than all other classes. If more than half of children in London in general and 55% of children from unemployed and routine social classes can be incentivised to participate in teacher-led OST programmes, it should be possible to achieve similar levels elsewhere.

Third, our findings imply that, where OST programmes are introduced, the impact is negligible unless they are teacher-led. Unless it can be argued that other benefits accrue, there seems to be little or no value in promoting self-study OST programmes, even if their provision cost is low.

Notes

1. https://educationendowmentfoundation.org.uk/toolkit/toolkit-a-z/extended-school-time/
2. http://www.cls.ioe.ac.uk/page.aspx?&sitesectionid=1246&sitesectiontitle=Welcome+to +the+Longitudinal+Study+of+Young+People+in+England+

3. Relevant survey questions are: ssexamYP 'Whether have teacher-led study groups outside lessons', ssexamfYP 'How many times a week YP works with teacher to prepare for exams outside lessons'.

4. Relevant survey questions are ssdropYP: 'Whether school has times outside lessons when can study (with other students) without teachers', ssdropfYP: 'How many times a week YP goes to study club'. We have no information as to whether 'teacher-led' involves a qualified teacher in all groups.

5. Key Stage 3 are national examinations administered in all state schools at age 14 in the three core subjects of English, Mathematics, and Science.

6. NPD data for Next Steps members are only available from 2006, the year of GCSE examinations (11). Therefore, we use data from Year 11 and excluded from the analysis the students which have changed school between Year 10 and 11—roughly 200 cohort members.

7. While Rosenbaum and Rubin (1985) used a caliper of 0.25 standard deviations, Austin (2011) and Wang et al. (2013) recommend a more conservative caliper of 0.2 standard deviations of the PS.

8. For example, attaining 10 Grade D GCSEs would be equivalent (in terms of total points scored) to receiving eight Grade C GCSEs (Dearden, Vignoles, Crawford, Goodman, & Chowdry, 2013).

9. The sample size for each quintile varies depending not only on the size of the quintile, but also on the propensity to participate in the programme. As the parameter of interest is the average treatment effect on the treated (ATT) and not an overall effect, where there are no participants a counterfactual cannot be estimated. Given that the first quintile has a lower propensity to participate (Table 1), the sample size for this group is consequently smaller.

10. https://educationendowmentfoundation.org.uk/toolkit/toolkit-a-z/extended-school-time/. These estimates cover school budget expenditures, but not the opportunity costs of qualified teachers' time when this is not counted as an added expense.

11. Using the prevailing answer rather than the individual one reduces the measurement error due to self-reporting: those not participating might erroneously think that the programme is not provided even if it is. By contrast, if the majority of students consistently report that the programme is (not) provided in their school, this gives more confidence in establishing that the school does (not) provides the programme.

Disclosure statement

No potential conflict of interest was reported by the authors.

Funding

Notes on contributors

Nicola Pensiero is a quantitative sociologist at the Centre for Learning and Life Chances in Knowledge Economies and Societies (LLAKES), UCL Institute of Education. His research interests lie in the sociology of stratification and inequality, sociology of education, and income inequality.

Francis Green is Professor of Work and Education Economics, LLAKES Centre, UCL Institute of Education. His general interests lie in labour economics, education, and political economy.

References

Apsler, R. (2009). After-school programmes for adolescents: A review of evaluation research. *Adolescence, 44*, 1–19.

Ascher, C. (1990). *Summer school, extended school year, and year-round schooling for disadvantaged students* (ERIC/CUE Digest Number 42). New York, NY: ERIC Clearinghouse on Urban Education (ERIC Document Reproduction Service No. ED 298 213).

Austin, P. C. (2011). Optimal caliper widths for propensity-score matching when estimating differences in means and differences in proportions in observational studies. *Pharmaceutical statistics, 10*, 150–161.

Austin, P. C., Grootendorst, P., & Anderson, G. M. (2007). A comparison of the ability of different propensity score models to balance measured variables between treated and untreated subjects: A Monte Carlo study. *Statistics in Medicine, 26*, 734–753.

Bianchi, S., Cohen, P., Raley, S., & Nomaguchi, K. (2004). Inequality in parental investment in childrearing. In K. Neckerman (Ed.), *Social inequality* (pp. 189–219). New York, NY: Russell Sage.

Bianchi, S. M., Robinson, J. P., & Milkie, M. A. (2006). *Changing rhythms of American family life*. New York, NY: Russell Sage.

Bonke, J., & Esping-Andersen, G. (2011). Family investments in children—productivities, preferences, and parental child care. *European Sociological Review, 27*, 43–55.

Cooper, H., Charlton, K., Valentine, J. C., & Muhlenbruck, L. (2000). Making the most of summer school: A meta-analytic and narrative review. *Monographs of the Society for Research in Child Development* (Serial No. 260), *65*, 1–118.

Cosden, M., Morrison, G., Albanese, A. L., & Macias, S. (2001). When homework is not home work: After-school programmes for homework assistance. *Educational Psychologist, 36*, 211–221.

Cunha, F., & Heckman, J. (2007). *The technology of skill formation* (No. w12840). National Bureau of Economic Research.

De Kanter, A. (2001). After-school programmes for adolescents. *NASSP Bulletin, 85*, 12–21.

Dearden, L., Vignoles, A., Crawford, C., Goodman, A., & Chowdry, H. (2013). Widening participation in higher education: Analysis using linked administrative data. *Journal of the Royal Statistical Society: Series A (Statistics in Society), 176*, 1–27.

Dehejia, R. H., & Wahba, S. (2002). Propensity score-matching methods for nonexperimental causal studies. *Review of Economics and Statistics, 84*, 151–161.

Durlak, J. A., Weissberg, R. P., & Pachan, M. (2010). A meta-analysis of after-school programmes that seek to promote personal and social skills in children and adolescents. *American Journal of Community Psychology, 45*, 294–309.

Elbaum, B., Vaughn, S., Hughes, M. T., & Moody, S. W. (2000). How effective are one-on-one tutoring programmes in reading for elementary students at risk for reading failure? A meta-analysis of the intervention research. *Journal of Educational Psychology, 92*, 605–619.

Erikson, R. (1984). Social class of men, women and families. *Sociology, 18*, 500–514.

Fashola, O. S. (1998). *Review of extended-day and after-school programmes and their effectiveness.* (Report No. 24). Baltimore, MD: Center for Research on the Education of Students Placed at Risk. Retrieved from http://www.csos.jhu.edu/crespar/techReports/report24.pdf

Fredricks, J. A., Blumenfeld, P. C., & Paris, A. H. (2004). School engagement: Potential of the concept, state of the evidence. *Review of educational research, 74*, 59–109.

Gerfin, M., & Lechner, M. (2002). A microeconometric evaluation of the active labour market policy in Switzerland. *Economic Journal, 112*, 854–893.

Grossman, J. B., Walker, K., & Raley, R. (2001, April). *Challenges and opportunities in after-school programs: Lessons for policymakers and funders*. Philadelphia, PA: Public/Private Ventures.

Heckman, J. J. (2006). Skill formation and the economics of investing in disadvantaged children. *Science, 312*, 1900–1902.

Heckman, J. J., Ichimura, H., & Todd, P. (1997). Matching as an econometric evaluation estimator: Evidence from a job training program. *Review of Economic Studies, 64*, 605–654.

Hsin, A. (2008). *Parenting, investments in children and the social reproduction of skills and status* (PhD dissertation). Los Angeles, CA: Department of Sociology, University of California.

Huang, D., Gribbons, B., Kim, K. S., Lee, C., & Baker, E. L. (2000). *A decade of results: The impact of the LA's best after school enrichment initiative on subsequent student achievement and performance*. Los Angeles, CA: UCLA Center for the Study of Evaluation, Graduate School of Education and Information Studies.

Karweit, N. (1985). Should we lengthen the school term? *Educational Researcher, 14*, 9–15.

Lareau, A. (2002). Invisible inequality: Social class and childrearing in black families and white families. *American Sociological Review,* 747–776.

Lauer, P. A., Akiba, M., Wilkerson, S. B., Apthorp, H. S., Snow, D., & Martin-Glenn, M. L. (2006). Out-of-school-time programs: A meta-analysis of effects for at-risk students. *Review of Educational Research, 76*, 275–313.

Lipsey, M. W., & Wilson, D. B. (1993). The efficacy of psychological, educational, and behavioral treatment: Confirmation from meta-analysis. *American Psychologist, 48*, 1181–1209.

MacBeath, J. E., Kirwan, T., Myers, B., McCall, J., Smith, I., & McKay, E. (2001). *The impact of study support: A report of a longitudinal study into the impact of participation in out-of-school-hours learning on the academic attainment, attitudes and school attendance of secondary school students* (DfES Research Report RR273).

McComb, E. M., & Scott-Little, C. (2003). *After-school programs: Evaluations and outcomes*. Greensboro, NC: SERVE.

McLanahan, S. (2004). Diverging destinies: How children are faring under the second demographic transition. *Demography, 41*, 607–627.

Miller, B. M. (2003). *Critical hours: Afterschool programmes and educational success*. Quincy, MA: Nellie Mae Education Foundation. Retrieved from http://www.nmefoundation.org/getmedia/08b6e87b-69ff-4865-b44e-ad42f2596381/Critical-Hours

MORI. (2004). *Study support survey 2004* (DfES Research Report RR591).

Pensiero, N. (2011). Parent-child cultivation and children's cognitive and attitudinal outcomes from a longitudinal perspective. *Child Indicators Research, 4*, 413–437.

Rosenbaum, P. R., & Rubin, D. B. (1983). The central role of the propensity score in observational studies for causal effects. *Biometrika, 70*, 41–55.

Rosenbaum, P. R., & Rubin, D. B. (1985). Constructing a control group using multivariate matched sampling methods that incorporate the propensity score. *The American Statistician, 39*, 33–38.

Stafford, F., & Yeung, W. J. (2005). The distribution of children's developmental resources. In D. S. Hamermesh & G. A. Pfann (Eds.), *The economics of time use* (pp. 289–313). Amsterdam: Elsevier.

Strand, S. (2011). The limits of social class in explaining ethnic gaps in educational attainment. *British Educational Research Journal, 37*, 197–229.

Sullivan, A., Ketende, S., & Joshi, H. (2013). Social class and inequalities in early cognitive scores. *Sociology, 47*, 1187–1206.

Wang, Y., Cai, H., Li, C., Jiang, Z., Wang, L., Song, J., & Xia, J. (2013). *Optimal caliper width for propensity score matching of three treatment groups: a Monte Carlo study*.

Appendix 1

Availability of OST programmes

The indicator of programme availability is the prevailing perception among students within each school regarding the existence of such a programme in the school. According to this indicator teacher-led study groups are offered by 95% of schools at age 14 and by 98% at age 15. The provision of self-directed study clubs is less widespread, with 68% of schools offering the programme at age 14 and 81% at age 15.[11]

Determinants of participation in OST programmes

There are two models in Table A1 describing the propensity of OST programme participation, one for each type of programme. There are four types of covariates: the ones regarding socio-economic background, the ones regarding the child's characteristics, the ones regarding the school, and one regional indicator. It seems the propensity of participation does not vary by socio-economic status,

as indicated by social class, marital status, and parental education. Students who receive free school meals at school tend to participate less in self-directed study clubs. A different conceptualisation of socio-economic status, including an additional indicator of family income, was also tried and this did not exhibit any additional significant relationship. Students from an ethnic minority background do not show a different propensity to participate in OST programmes than white students, with the exception of black African students who are more likely than the white ones to attend teacher-led study groups. The participation in self-directed study clubs does not vary according to ethnic origins. Children from a two-parent household do not show a different participation than students from a single-parent household. The main factors determining differences in the propensity of participation are the behavioural/attitudinal ones and the school-related ones.

Table A1. Effect of social origins, child's and school characteristics on the propensity to participate in after-school programmes at age 14/15 (Year 10)—probit model.

	Teacher-led study group	Self-directed study club
Female	0.173**	−0.302***
	(0.0610)	(0.0653)
KS3	−0.00415	−0.0135
	(0.00667)	(0.00732)
Special education need	−0.154	0.152
	(0.132)	(0.133)
Study club	0.889***	
	(0.0690)	
Study group		0.917***
		(0.0681)
School engagement	0.219***	0.216***
	(0.0357)	(0.0387)
Homework	0.110***	0.0786**
	(0.0226)	(0.0246)
Educational plans	0.266*	0.0872
	(0.111)	(0.122)
Social class NS-SEC (Ref.: VIII)		
VII	0.0847	0.0747
	(0.228)	(0.219)
VI	0.0509	−0.0998
	(0.214)	(0.210)
V	0.306	−0.0278
	(0.219)	(0.217)
IV	0.109	−0.101
	(0.230)	(0.230)
III	0.171	−0.0754
	(0.219)	(0.220)
II	0.200	−0.0112
	(0.213)	(0.211)
I	0.137	−0.157
	(0.226)	(0.222)
Parental academic qualifications (Ref.: less than Level 1)		
Level 1	−0.205	0.169
	(0.133)	(0.147)
Level 2	−0.134	0.152
	(0.126)	(0.137)
Level 3	−0.185	0.191
	(0.138)	(0.143)
Level 4	−0.207	0.0796
	(0.144)	(0.153)
Ethnic origin (Ref.: White)		
Mixed	−0.0488	0.141
	(0.147)	(0.171)
Indian	0.0698	0.188
	(0.131)	(0.151)
Pakistani	−0.219	−0.0336
	(0.161)	(0.183)

(*Continued*)

Table A1. (*Continued*)

	Teacher-led study group	Self-directed study club
Bangladeshi	−0.113	0.0831
	(0.232)	(0.218)
Black Caribbean	0.122	0.0914
	(0.188)	(0.213)
Black African	0.383*	0.193
	(0.182)	(0.193)
Free school meal	−0.0633	−0.377**
	(0.126)	(0.139)
Two-parent family	0.0772	−0.0151
	(0.0751)	(0.0764)
School characteristics		
% FSM	−0.06333	0.0120*
	(0.00528)	(0.00478)
% English not first language	−0.00407	−0.00152
	(0.00289)	(0.00277)
% Lev 2	7.50e-05	0.00712**
	(0.00264)	(0.00250)
% saying that programme is available	0.03120***	0.02695***
	(0.333)	(0.220)
Multiple deprivation index	−0.00194	−0.000386
	(0.00267)	(0.00270)
Urban/rural (Ref.: < than 10K population)		
≥10K population (No London)	0.167*	−0.0403
	(0.0835)	(0.0838)
London	0.405**	0.0807
	(0.140)	(0.145)
Constant	−4.217***	−2.915***
	(0.448)	(0.374)
Pseudo *R*-squared	0.0817	0.0932
Observations	6019	6019

Notes: Robust standard errors in parentheses. *** $p<0.001$, ** $p<0.01$, * $p<0.05$.

School engagement, frequency of doing homework, and educational plans (in respect to teacher-led study group) are significantly and strongly associated with a higher propensity of participation. Academic performance at KS3 seems to be negatively related to the propensity of participation, although the coefficient does not reach the statistical significance bar. Being a girl is negatively related to the propensity of participation in self-directed study cubs, yet it is positively related to the propensity to participate in teacher-led study groups. Those who participate in one type of programme tend also to participate in the other type of programme as shown by a positive and significant coefficient related to participating in the other type of programme.

Schools seem to target self-directed study clubs to economically disadvantaged children as shown by the higher propensity of participation in schools where there is a larger share of students receiving free school meals. The school average achievement defined as the percentage of students attaining level 2 qualifications is related to a higher participation in self-directed study clubs. School-related factors do not matter for the participation in teacher-led study groups. The availability of the programme in school, whose indicator is the proportion of students affirming that the programme is available, is significantly related to a higher propensity to participate in both types of programmes. Moving to the broader social context, the multiple deprivation index does not show any effect on participation in either programme, while living in a large town is important for increasing the participation in teacher-led study groups but not in self-directed study clubs.

Matching procedure: the case of teacher-led study group

Table A2 presents the distributions of resulting propensity scores grouped by study-group attendance. There are differences in propensity scores between attendees and non-attendees, but there is

overall considerable overlap between the two groups, which is a confirmation of the success of the matching procedure.

Table A2. Cumulative distribution of estimated propensity scores—teacher-led study group participation at age 14/15 (Year 10).

	Participation in teacher-led study group		
Propensity score	Yes	No	Total
0	133	22	155
0.2	365	133	498
0.3	728	377	1105
0.4	400	291	691
0.45	328	315	643
0.5	446	611	1057
0.6	321	610	931
0.7	172	447	619
0.8	44	250	294
Total	2937	3056	5993

To assess the quality of the match, we analyse in Table A3 whether the propensity score adequately balances characteristics between the treatment and comparison group units. The objective of these tests is to inspect whether treatment is independent of unit characteristics after conditioning on observed characteristics as estimated in the propensity score model. We compare the before-matching and after-matching imbalance, the objective being to examine whether any differences between group means in the matched sample have been eliminated. After matching, the highest bias is about 13%, yet most of the covariates are related to a bias lower than 8%. The highest post-matching bias is found in the school-level covariates percentage receiving free school meals (12.7%) and percentage of students whose first language is not English (11.8%) and in the covariate identifying black African (10%). After matching on the estimated propensity score, observed systematic differences between treated and untreated participants appear to have been greatly reduced. Overall, the analysis indicates that the propensity score model is adequately specified.

Table A3. Covariate imbalance among teacher-led study group attendees and non-attendees at age 14/15 (Year 10).

		Treated	Control	% bias	% reduction of bias
Female	Unmatched	1.5295	1.4717	11.6	
	Matched	1.5347	1.5109	4.8	58.8
KS3	U	35.375	34.443	14.2	
	M	36.142	36.33	−2.9	79.8
Special education need	U	0.05714	0.10495	−17.6	
	M	0.04679	0.05166	−1.8	89.8
School engagement	U	0.28288	−0.0696	37.3	
	M	0.29709	0.22053	8.1	78.3
Homework	U	3	2.5929	27	
	M	3.0141	2.9036	7.3	72.9
Educational plans	U	0.92696	0.85657	22.8	
	M	0.93586	0.93077	1.7	92.8
Social class NS_SEC (Ref.: VIII)					
VII	U	0.07937	0.08552	−2.2	
	M	0.07363	0.07211	0.6	75.3
VI	U	0.11856	0.13012	−3.5	
	M	0.11158	0.11334	−0.5	84.8
V	U	0.10364	0.09488	2.9	
	M	0.10635	0.10506	0.4	85.3
IV	U	0.07281	0.0657	2.8	
	M	0.06479	0.05934	2.1	23.2
III	U	0.09847	0.10075	−0.8	

(*Continued*)

Table A3. (*Continued*)

		Treated	Control	% bias	% reduction of bias
	M	0.10177	0.09678	1.7	−118.2
II	U	0.3149	0.31052	0.9	
	M	0.33671	0.34457	−1.7	−79.2
I	U	0.13746	0.14627	−2.5	
	M	0.14431	0.15596	−3.3	−32.3
Parental qualification (Ref.: Less then level 1)					
Level 1	U	0.13589	0.14069	−1.4	
	M	0.13384	0.13266	0.3	75.5
Level 2	U	0.26588	0.26781	−0.4	
	M	0.27356	0.27312	0.1	76.9
Level 3	U	0.19688	0.19639	0.1	
	M	0.1983	0.20485	−1.6	−1247.6
Level 4	U	0.22579	0.2288	−0.7	
	M	0.23037	0.23475	−1	−45.4
Ethnic origin					
Mixed	U	0.05132	0.05332	−0.9	
	M	0.05007	0.04897	0.5	45.4
Indian	U	0.08224	0.05388	11.3	
	M	0.08181	0.07189	3.9	65
Pakistani	U	0.06203	0.05203	4.3	
	M	0.05825	0.05168	2.8	34.3
Bangladeshi	U	0.0586	0.03592	10.7	
	M	0.0553	0.04697	3.9	63.3
Black Caribbean	U	0.04061	0.02611	8.1	
	M	0.03698	0.02936	4.2	47.5
Black African	U	0.04789	0.0224	13.9	
	M	0.04418	0.0258	10	27.9
Free school meal	U	0.16231	0.14661	4.3	
	M	0.13973	0.12524	4	7.7
Two-parent family	U	0.76225	0.73838	5.5	
	M	0.77651	0.78208	−1.3	76.7
% FSM	U	18.46	15.349	20	
	M	18.047	16.07	12.7	36.4
% English not first language	U	18.245	13.109	21.4	
	M	17.477	14.655	11.8	45.1
% Lev 2	U	60.621	60.229	2	
	M	60.114	60.991	−4.5	−123.9
% saying that teacher-led study group is available	U	0.88126	0.83892	33.3	
	M	0.88527	0.87762	6	81.9
% saying that self-directed study club is available	U	0.64899	0.61889	17.9	
	M	0.64827	0.65171	−2	88.6
Multiple deprivation index	U	25.589	23.593	11.3	
	M	24.439	23.136	7.4	34.7
Urban/rural (Ref.: <10K population)					
≥10K, no London	U	0.63515	0.66468	−6.2	
	M	0.64823	0.65549	−1.5	75.4
London	U	0.22107	0.14348	20.2	
	M	0.19764	0.1654	8.4	58.4