# The direct and indirect effects of education policy on school and post school outcomes*

Steve Bradley and Giuseppe Migali

June 24, 2012

**Abstract**

Successive British governments since the early 1980s have introduced a host of educational policy reforms in an attempt to raise pupil performance at school. One of the most important educational policies in the secondary education sector was the specialist schools policy, which was introduced in 1994. Using data from the YCS for pupils who left school in either 2002 or 2004, a period of rapid expansion of the specialist schools programme, we seek to evaluate the effects of the policy. Unlike most previous work in this area we investigate the effects of the policy on test scores and truancy for pupils at school, but also assess whether the policy had direct and/or indirect effects on post-school outcomes, such as labour market status, wages and A-Level scores. We show that specialist schools did raise test scores during compusory schooling, and that the policy had a positive and statistically significant effect in raising the probability of employment. The evidence on A-level scores suggests a negative effect and, due to data limitations, no effect on wages is apparent. Although we stop short of claiming that our findings are causal, they do imply that policy makers need to take a more comprehensive view of the effects of education policies when trying to address whether they deliver value for money to the taxpayer.

*The authors would like to thank Nuffield Foundation for the financial support to this work.

# 1 Introduction

Successive British governments since the early 1980s have introduced a host of educational policy reforms, such as the 1988 Education Reform Act, which led to the creation of a quasi-market in education. Funding for secondary schools has also been increased substantially since 1997, rising from £9.9bn to £15.8bn in 2006/7, equivalent to an increase in real expenditure per pupil increased by over 50% (from £3206 to £4836 in 2005/6 prices). One of the most important educational policies in secondary education in this time period was the specialist schools policy, which was introduced in 1994. To obtain specialist status, state maintained schools were required to raise unconditional sponsorship from the private sector of £50,000 and to have a development plan. Selected schools then received a capital grant of £100,000 from central government, and around £130 per pupil over a four year period, equivalent to a 5% increase in per pupil funding.[1] In addition to increasing the funding of schools, the specialist schools policy also simultaneously enhanced parental choice of the school their child would attend, and competition between schools for pupils, because schools were encouraged to specialise in particular subjects.[2] The earliest specialist schools were Technology schools, starting in 1994, constituting approximately 20% of all schools by 2007, with significant proportions of schools focusing on Arts, Sport and Science. Other specialisms, such as Business and Maths were introduced more recently in 2002.[3] Specialist schools are encouraged to spread good practice to non-specialist schools in the same educational district with respect, for instance, to teaching methods. Over 80% of secondary

[1] The capital grant was reduced to £25,000 in recent years but was higher for the time period covered by this study.

[2] It is worth noting that although specialist schools are encouraged to focus on particular subjects, all schools are also required to deliver a national curriculum. Thus, most pupils will typically study around 10 subjects in their final two years of compulsory schooling between the ages of 14 and 16. They then sit nationally recognised tests, the General Certificate of Secondary Education (GCSE), in each subject.

The GCSE is a norm-based examination taken by almost all pupils, and the grades range from A* to G. Grades A* to C are considered acceptable for entry to university, together with the acquisition of advanced qualifications obtained two years later. Pupils of lower ability may also take General National Vocational Qualifications instead of GCSEs.

[3] See the Department for Children, Schools and Families website for more details: www.standards.dcsf.gov.uk/specialistschools.

schools are now specialist and the policy has since been abandoned as priorities have shifted in favour of the academy schools initiative.

The key objective of the specialist schools policy was to improve the test score performance of secondary school pupils. There are two different mechanisms by which the policy could affect test scores - *funding* and *specialisation* effects. The increase in resources to specialist schools creates a funding effect whereby increased spending on books and equipment, for instance, improves the quality of the educational experience throughout the school and hence may improve test scores in all subjects. However, by allowing greater subject specialisation, parents can select those schools that 'match' the aptitudes and skills of their children, thereby increasing allocative efficiency. 'Better' subject specialist teachers may also move to schools that specialise in their subject area. Hence test scores in particular subjects may increase - a specialisation effect. Clearly, if specialist schools are shown to increase the test scores of pupils then, indirectly, the policy may also affect post-school outcomes since test scores have been shown to be highly correlated with staying on rates and wages, for instance. We test for such indirect effects. However, it is also important to know whether the specialist schools policy had a direct effect on post-school outcomes. This may arise via the specialisation effect insofar as pupils acquire skills and knowledge in those subjects in which the school specialises which are also highly valued in the labour market. Obtaining estimates of the direct and indirect effects of the specialist schools policy is helpful to policy-makers interested in determining whether educational initiatives are 'value for money'. As far as we are aware this is the first paper to investigate the direct and indirect effects of the policy on post-school outcomes.

The main questions we address in this paper are as follows: First, we begin by asking whether the specialist schools policy has had an effect on test scores and truancy behaviour of pupils. Second, conditional on test score performance, are pupils who have attended specialist schools more likely to continue their education beyond the compulsory school leaving age? Thirdly, for those young people who enter the labour market is there a wage premia associated with attendance at a specialist school? Fourthly, has the specialist school policy led to improvements in educational attainment at the post-compulsory level, such as A-level qualifications. To answer these questions we use the Youth Cohort Surveys, cohorts 11-12, which refer to young people who finish compulsory schooling between 2002 and 2004, a period of time over which the number of specialist schools expanded rapidly.

It is not straightforward to evaluate the effect of the specialist schools policy on educational or labour market outcomes because these schools were not randomly selected to participate in the policy. Indeed, there are likely to be several sources of bias that must be mitigated in an evaluation of the specialist schools policy. One source of bias arises directly from the non-random selection of schools into the policy (hereafter *school selection bias*). Figure 1 shows that specialist schools have tended to outperform non-specialist schools throughout the history of the policy, and Bradley and Migali (2011) present evidence to suggest that specialist schools are increasingly likely to have test scores in the highest quintile of test scores, which is strongly suggestive of non-random assignment of certain types of school into the specialist schools initiative. A closely related source of bias is the non-random selection of pupils into specialist schools (*pupil selection bias*), insofar as unobservably more able pupils are 'cream-skimmed' by 'good' (specialist) schools. Specialist schools are, for instance, more likely to have a higher percentage of pupils from richer social backgrounds and are more likely to have a lower percentage of ethnic minority pupils (Bradley and Migali, 2011). It is difficult to disentangle these two sources of bias with available data, nevertheless in trying to measure the impact of the specialist schools policy on educational and labour market outcomes it is important to try and do so. We do so by adopting matching methods for the analysis of test scores at the pupil and school level, and a control function approach for post-school outcomes.

The remainder of this paper is structured as follows. In section 2 we discuss the small literature that has attempted to measure the impact of various educational policies. This is followed by a discussion of our data and statistical methods. Section 4 presents our results which is followed by our concluding comments.

# 2 Literature

## 2.1 The effect of the specialist schools policy on test scores

There is a growing literature on the evaluation of the specialist schools policy on test score outcomes at school. For example, Gorard (2002), Jessson (2002, 2004), Jesson and Crossley (2004) and OFSTED (2005) find a positive effect on test scores. Most of this early work use school level or pupil level data to estimate cross-sectional OLS models. For instance, Jesson finds that specialist schools increased the percentage of pupils obtaining 5 or more GCSEs grades A*-C by between 4.5 percentage points and 5 percentage points, and total points score is increased by 4.2%. These are large effects but no allowance is made for the different types of selection bias. Schagen, Davies, Rudd and Schagen (2002) do raise some issues regarding the methodological approach of this early work, arguing that pupil level data and multi-level modelling techniques should be used. Estimated effects of the policy arising from this study are considerably smaller at between 0.02-0.11 of a GCSE point with some variation by subject - maths (0.04-0.06), English (0.03-0.09) and science (-0.03-0.07). Benton *et al* (2003), again using multi-level modelling techniques with pupil level data, find that the specialist schools policy raised GCSE grades, or points, by 1.1. Similarly, Levacic and Jenkins (2006) using data for 2001 find a very similar effect at 1.4 GCSE points. These authors also found some variation by subject of specialisation and the duration of the policy.

Taylor (2007) finds that the specialist schools policy has had very little impact on average test scores, though there is evidence of more substantial impacts for specific areas of specialisation, for example, business and technology. Bradley and Taylor (2010) estimate the impact of the specialist schools policy, as well as other educational policies, using school-level panel data, and find a small positive effect of specialist schools on test scores. However, many of these papers fail to allow for the bias that often arises in programme evaluation settings, which calls into question whether they have been able to identify a causal effect of the specialist schools policy. Furthermore, many of these studies do not explicitly consider the mechanisms by which the specialist schools policy could affect the test score outcomes of pupils. Bradley, Migali and Taylor (2012) use matching methods combined with a difference-in-differences approach on pupil data drawn from several versions of the NPD (2002-2005) and find a modest effect of the policy. In an extension of this work Bradley and Migali (2012) investigate the effects of both the specialist schools and EiC policies and find that, whilst the latter has a greater effect on pupil test scores, there is also a complementary effect of the two policies.

There is a broader literature (see Hanushek (1998) for a survey) which focuses on the effectiveness of school resources, which is also relevant to this paper. For example, using US data on expenditure and National Assessment of Education Progress test scores, Krueger (1998, 2000) finds modest gains in test scores due to increase in expenditures and Hanushek (1998) does not find any strong relationship between resources and student performance. Machin and McNally (2011) also provide a summary of the effects of a variety of British educational policies.

There is very little prior research on the labour market effects of educational policies and none on the effect of the specialist school initiative. There is a very large literature on the returns to education but this tends to focus on the wage returns to an extra 'unit' of education, rather than on the effect of a particular education policy. In fact, studies of this kind are hard to find. Nevertheless, education economists have in recent years investigated the effects of different aspects of education policy on labour market outcomes, such as the effect of a central exam system on wages (Backes-Gellner and Veen, 2006) and the effect of school type (general versus technical, public versus private) on employment probabilities, wages and the transition to HE (Cappellari, 2004; Margolis and Simonnet, 2002).

Other literature..

# 3 Data and descriptives

The data used in this analysis are the Youth Cohort Surveys (YCS) of England and Wales, versions 11 and 12, which refer to the 2002-2004 period. The YCS is a major programme of longitudinal research designed to monitor the behaviour and decisions of representative samples of young people (around 14,000 per survey) aged 16 and upwards. The survey records educational outcomes and provides a wealth of information on the socio-demographics of the pupil's family. By combining school level data with the pupil level data in the YCS we are able to observe precisely when each school switches from non-specialist ('policy-off') to specialist ('policy-on') status to investigate how the test scores of different cohorts of pupils change. Specifically, we link schools in YCS11 with the same schools in YCS12 and restrict attention to those pupils in a non-specialist school in 2001/02 ('policy-off') and compare them with pupils in the same school which acquired specialist status during 2002/04 ('policy-on'). This reduces the original sample to 5,244. This approach allows us to go some way to controlling for school selection bias since we essentially difference out unobserved school fixed effects. Pupil selection bias should also not be a problem since all of the pupils in the analysis had chosen a non-specialist school, which then becomes specialist during their period of secondary schooling (policy on). One potential drawback of this type of analysis is the fact that we cannot distinguish year effects and cohort effects, because the inclusion of these dummies would be perfectly collinear with the treatment variable. Consequently, we assume that temporal exogenous shocks affect all schools in the same way, hence shifting the distribution of pupil attainment.

Each YCS also comprises three sweeps conducted at the ages 16, 17 and 18. For each sweep the young person is asked about their educational and labour market status. In the first sweep the information includes their experiences and achievements at school, and their personal and family characteristics. For young people who proceed to post-compulsory education, the Survey also collects information on the type of course taken, and the grades achieved. We link to the YCS data information on the school, including information on whether a school is specialist or not, which is obtained from the annual School Performance Tables and the School Census. Using these data we are able to investigate 'early career' outcomes.

Tables 1 shows three measures of test score outcome for the combined sample YCS11-12. The first is the total GCSE score (*GCSEscore*), taken at age 16, that is, the number of points achieved in all GCSE subjects, where grades are ranked from A*=8 points to fail=0. The second is a binary variable indicating whether a pupil obtained 5 or more GCSE grades A*-C (*GCSEbin*). The third measure is also a dummy variable indicating whether a pupil obtained 10 or more GCSE grades A*-C (*GCSEbin10*), which refers to the upper end of the ability distribution. In each case we distinguish between pupils who were in non-specialist schools in YCS11 and compare their test score performance against pupils in specialist schools in YCS12, that is, schools that had acquired specialist status between YCS11 and YCS12. Table 1 shows that pupils who attended specialist schools have a higher GCSE points score (*GCSEscore*) although a lower proportion achieve 5 or more GCSEs A*-C. The latter finding is probably explained by the fact that specialist school pupil performance is massively greater than their non-specialist school pupils at the upper end of the test score distribution (GCSE-10+A*-C). Pupils in schools that have specialist status have a much lower propensity to truant by about 14 percentage points.

The bottom of Table 1 and Table 2 provide descriptive statistics on post-school test scores (A-Level score). Due to data limitations we focus on YCS12 and we append, to the original sample of 10270 pupils, school level data. Once again this enables us to to distinguish between schools that became specialist between 1994 and 1999 (the start of the specialist school programme and the year before pupils enrolled in secondary school) and non-specialist schools. The latter schools actually acquired specialist status between 2004 and 2006 on the grounds that these schools are likely to be most similar to those schools that were specialist at the time of YCS12. This restriction makes our comparison groups more homogenous but we lose around 50% of the initial sample. Panel B of Table 1 shows A-level test score and mean number of A levels studied.[4] We observe that pupils that attended specialist schools on average outperform those from non-specialist schools in terms of A-level qualifications.

Looking at Table 2, the majority of pupils continue their full time studies in wave 1, although those coming from a specialist school are 3 percentage points more likely to do so. Furthermore, the raw data also shows that pupils from specialist schools have a slightly greater probability of securing employment after leaving and a consequently lower risk of entering the NEET categority (Not in eduction, employment or training). The percentage of pupils continuing their education at wave 2 is much lower, although pupils from specialist schools are more likely to study for longer (52% of pupils stay-on from non-specialist schools versus 55% from specialist schools). Interestingly, the percentage of pupils from non-specialist schools in employment by wave 2 is higher when compared to their counterparts from specialist schools, perhaps reflecting the greater success in the labour market of the group of non-specialist school pupils who leave college after 1 year. However, pupils that had been educated in specialist schools were still less likely to enter the NEET categort. By wave 3 the outcomes of the two groups of pupils have more or less equalised, which suggests that any advantage that a specialist school education provides may be short lived.

Finally, for those pupils in employment, log real wages are very similar for pupils from specialist and non-specialist schools. Wages rise from wave 1 to wave 2 and then drop in wave 3.

# 4 Statistical methodology

We follow two econometric methodologies, the first is based on cross-sectional matching and estimates the effect of specialist schools on education test scores. The second is a control function approach and estimates the effect of specialist schools on further education and labor market outcomes.

[4] A Level qualifications are taken during post-compulsory schooling, typically by pupils aged 16-19. They are pre-requisites for entry to HE in the UK.

## 4.1 Matching methods

Our approach is based on the concept of the education production function wherein educational outcomes, such as test scores are a function of personal, family and school inputs, as well as specialist school status. However, to estimate the effect of the specialist schools policy on pupils' test scores requires a solution to the counterfactual question of how pupils would have performed had they not attended a specialist school. We adopt the non-parametric matching method which does not require an exclusion restriction, or a particular specification of the model for attendance at a specialist school. Thus, the main purpose of matching is to find a group of non-treated pupils who are similar to the treated in all relevant pre-treatment characteristics, $\mathbf{x}$, the only remaining difference being that one group attended a specialist school and another group did not. In the first stage we therefore estimate the propensity score (PS) using a discrete response model of attendance at a specialist school.

One assumption of the matching method is the *common support* or overlap condition which ensures that pupils with the same $\mathbf{x}$ values have a positive probability of attending a specialist school. A second, and key assumption is the *conditional independence assumption (CIA)*, which implies that selection into treatment is solely based on observable characteristics.[5] There may, however, be a problem of hidden bias due to unobserved effects, and any positive association between a pupil's treatment status and test score outcomes may not therefore represent a causal effect. If the assumption of ignorability (i.e. no hidden bias) fails, the treatment is endogenous and the matching estimates will be biased (Heckman et al. 1998). Several tests have been developed to assess whether hidden bias is a problem in cross-sectional models and we adopt the method proposed by Ichino et al. (2008) and Rosenbaum (1987). The details of these tests are reported in Appendix A.

Given these two assumptions, the matching method allows us to estimate the average treatment effect on the treated (ATT). The ATT estimator is the mean difference in outcomes

[5] Conditional on a set of pre-treatment observable variables $\mathbf{x}$, potential outcomes are independent of assignment to treatment.

over the common support, weighted by the propensity score distribution of participants.

All matching estimators are weighted estimators, derived from the following general formula:

$$\tau_{ATT} = \sum_{i \in T} ( Y_{1i} - \sum_{j \in C} W_{ij} Y_{0j} ) \, w_i \tag{1}$$

where $T$ and $C$ represent treatment and control groups, respectively. $W_{ij}$ is the weight placed on the $j$th observation in constructing the counterfactual for the $i$th treated observation. $Y_1$ is the outcome of participants and $Y_0$ of non-participants; $w_i$ is the re-weighting that reconstructs the outcome distribution for the treated sample. A number of well-known matching estimators exist which differ in the way they construct the weights, $W_{ij}$. We use two matching algorithms, nearest neighbor (NN) and kernel with bandwidth 0.1. Analytical standard errors are provided for the first estimator (Abadie and Imbens, 2008) and bootstrapped standard errors for the second (Heckman et al. 1998).

A further issue arises insofar as expenditure per pupil has risen for reasons other than the specialist schools policy and we must control for this. Therefore, we estimate a post-matching regression including real expenditure per pupil, measured in 2003 prices, of the form:[6]

$$Y_{is} = \alpha + \beta Spec_{is} + \gamma \widehat{p}(x_i) + \rho Spec_{is}(\widehat{p}(x_i) - \mu_p) + \theta Exp_{is} + \varepsilon_{is} \tag{2}$$

where $\widehat{p}(x_i) = P(Spec_{is} = 1 | x_i)$ is the estimated propensity score, $\mu_p$ is the sample average of $\widehat{p}(x_i)$ and $Exp_{is}$ are the expenditure per pupil in school $s$. Since we only have data on expenditure from 1999, we restrict our sample to schools that become specialist from 1999 to 2002, and schools that are non-specialist in the same period but which become specialist between 2003 and 2005.

[6] In the post-matching analysis we adopt a control function approach using the propensity score (Wooldridge, 2005 and Rosembaum and Rubin, 1983). We regress our dependent variable (test scores) on the treatment dummy variable (*Spec*), the estimated propensity score and its deviation from the mean interacted with *Spec*. The coefficient of the *Spec* dummy consistently estimates ATE. In our specific case, we add in the post-matching regression the expenditure per pupil variable and its deviation from the mean interacted with *Spec*. In this way we want to get the treatment effect corrected for the fact that the control group changes over time.

## 4.2 Control Function method

We turn now to the statistical methods adopted to assess the effect of the specialist schools policy on post-school outcomes. Our baseline model involves the estimation of the following equation

$$Y_i = \alpha + \gamma Spec_i + \beta X_i + \epsilon_i \tag{3}$$

where $Y$ variable represents either further education outcomes or labor market outcomes, *Spec* indicates whether a pupil has attended a specialist or non-specialist school. The vector, $\mathbf{x}$, includes a series of controls. The main problem for the estimation of equation 3 is the presence of ability bias which implies that the effect of specialist school on outcome reflects both unobservable differences across pupils as well as the effect of the policy *per se*, so that OLS are biased upwards. A common solution to this empirical problem is to adopt instrumental variable (IV) methods, however it is not easy to find variables both uncorrelated with the unobservables determinants of outcomes and correlated with the policy variable. Therefore, we adopt a two-step selection model or control function estimator, which is more restrictive than IV. We estimate as first step a propensity score model of school participation in the specialist school programme. We match schools similar in all pre-treatment characteristics except being specialist or non-specialist and we estimate a probit

$$S_s^* = \delta + \rho \mathbf{Z}_s + \upsilon_s \tag{4}$$

where $S_s = 1$ if schools $s$ acquired specialist status between 1994 and 1999 and $S_s = 0$ if school $s$ is non-specialist but will acquire specialist status after 2004.

We use the estimated probit coefficients from equation 4, to compute the predicted probability of a school acquiring specialist status, denoted *specscore*. We then merge this variable, at school level, with the pupil data. All pupils in the same school are associated with the same value of *specscore*. As a second step we estimate Equation 5:

$$Y_i = \alpha + \lambda_1\, specscore_i + \lambda_2\, Gcsescore_i + \lambda_3\, (Gcsescore_i \times specscore_i) + \beta X_i + u_i \tag{5}$$

By including *specscore* in equation 5 we hope to correct for the fact that pupils in specialist schools are unobservably different to those pupile in non-specialist schools for the reasons cited in the Introduction. We also include a test score variable, *Gcsescore*, to control for prior attainment, and we interact this variable with *specscore* to try to capture the indirect effects. The vector X groups individual characteristics (such as age and gender), family background (whether the parents are in a profession/managerial or less qualified job, whether the pupil has only one parent) and school characteristics (whether the pupils has attended a comprehensive or modern school, and whether the schools was a only boys or a only girls school).

This approach is more restrictive than IVs because it relies on the assumption of mean independence of $u$ from $Spec$ and $Z$ conditional on $v$:

$$E[u|v, Spec, Z] = E[u|v] = f(v) \tag{6}$$

where $f$ is the control function.

However, the advantage of this method is that we estimate the effect of specialist schools on outcomes across the whole distribution of unobservables, in particular at the mean. Thus, the control function method, unlike IV, yields estimates that are comparable to the much simpler least squares regression method. Furthermore, to better control for school self-selection in the specialist school programme, we repeat our approach by considering in the first step only school in the first 65% percentiles of the *specscore* distribution. We define *spec*65 this new variable which replaces *specscore* in equation 5.

In general, we estimate four different models according to the definition of the dependent variable. The first model considers post-school (labour market) outcomes, we define a categorical variable

$$\begin{aligned} Y_i = & 1 \quad \text{if individual i is in full-time education} \\ = & 2 \quad \text{if individual i is in full-time job} \\ = & 3 \quad \text{if individual i is NEET} \\ = & 0 \quad \text{otherwise.} \end{aligned}$$

In this case we estimate equation 5 as a multinomial logit and we compute the marginal effects. In the second model we use the natural logarithm of real hourly wages as a continuous dependent variable and we estimate equation 5 using OLS regressions. The third model uses as dependent variable the average between the number of A2 points and AS points cumulative to wave 3. We estimate equation 5 as Poisson regressions and we compute the marginal effects. Finally, in the fourth model the dependent variable is the total number of A2 and AS achieved cumulative to wave 3. Equation 5 is estimated by OLS regressions.

# 5 Results

## 5.1 A 'policy-on' versus 'policy-off' analysis

Table 3 shows that, prior to matching, pupils in a given school during a 'policy-on' period obtain around 2.8 GCSE points more than their counterparts in the *same* school in the 'policy-off' period (*Gcsescore*).[7] After matching we observe a reduction in the effect on GCSE points score by between 15-37%, with the estimated impact falling to between 1.7-2.0 points, depending on which estimator is used. Interestingly, there is no statistically significant difference in the proportion of pupils obtaining 5 or more GCSEs graded A*-C (*Gcsebin*). However, at the very top of the attainment distribution (*Gcsebin10*) a positive

[7] The probit model for selection (choice) into (of) specialist schools contain personal (gender, ethnicity), parental (e.g. occupational status, parental qualifications) and school characteristics that could affect the choice or selection decisions of school administrators and pupils. In terms of school variables, we use the lagged GCSE performance of the school. Most variables are highly statistically significant.

and statistically significant effect is observed. In fact, the pre-match estimate of 0.09 falls to between 0.07-0.08 implying that the specialist schools policy increased the probability of obtaining 10+ GCSE grades A*-C by between 7-8 percentage points.

The inclusion of a confounder variable does not dramatically change our results. This is also confirmed by the Rosembaum test (see Table 3, lower panel) in a situation of up to 50% hidden bias. As an additional robustness check we estimate equation 2 which corresponds to a post-matching regression that includes expenditure per pupils at school level. We notice that the effect of the policy remains substantially unchanged compared to the matching estimates. In particular, for *Gcsescore* the effect is 1.86, a values that lies between the nearest neighbour and kernel estimates. The same happens for *Gcsebin10*, while the effect for *Gcsebin* is still insignificant. Note that in this analysis we mitigate the bias arising from school selection bias, this is because we remove unobserved school fixed effects and due to the short time framework of the analysis it is also unlikely to be affected by pupil selection bias.

With regard to truancy, the evidence suggests that prior to using matching methods, specialist school pupils have 4.7 percentage point lower probability of truanting, which is very close to the mean difference reported above. Post matching our estimates suggest that this is about 4 percentage points depending on whether the nearest neighbour of kernel matching method is used. Inclusion of a confounder variable, designed to pick up the effects of unobserved differences between specialist and non-specialist pupils, shows that the probability of a pupil from a specialist school is 5.4 percentage points lower that it is for their counterpart from a non-specialist school. This is about 1.5 percentage points higher than the estimates from the models without the confounder variable and implies that pupils from non-specialist schools are, not only observably different to their counterparts from specialist schools with respect to truancy behaviour, but they are also unobservably different.

In sum, this section of the paper has sought to demonstrate that specialist school pupils have a higher test scores and lower probabilities of truanting after allowance is made for observable (and unobservable) differences between pupils. Although our models do seek to control for the different types of selection bias discussed in the Introduction, we stop short of claiming that the relationship between specialist school attendance has a causal effect on these school outcomes. Nevertheless, they are indicative of an effect, which in turn implies that specialist schools may have an indirect effect on post school outcomes (e.g. via test scores) as well as a possible direct effect.

### 5.2 The effect of the specialist school policy on labour market outcomes

We report all the results of estimating of the control function models. As explained in Section 4.2, the first step in each model is always the same and allows use to obtain the two variables *specscore* and *specs*65. The school matching passes the balancing test. The probit model of attendance at a specialist school contains pre-treatment school characteristics, such as the (lagged) GCSE performance of the school, the type of school (e.g. modern, comprhensive), and whether the school is a girls-only or a boys-only school. All variables are highly statistically significant.

Looking Table 4 we show the results for the labor market outcomes. We distinguish on the one hand between a direct effect of the policy and an indirect effect arising via test scores, and between two measures of the policy, that is, the continuous variable *specscore* (columns 2-4) and the categorical variable *specs*65 (columns 5-7). In terms of the former, we find that pupils who attend schools with a higher propensity of becoming a specialist school are less likely to stay on, conditional on a variety of covariates and pupil test score, This effect declines over time. However, there is a direct effect of specscore on the probability of pupils finding a job after leaving school, raising this probability by 1.8 percentage points in wave 1. Moreover, the effect increases over time and remains statistically significant - by wave 2 the effect is 2 percentage points and 3.6pp by wave 3. These effects are strong especially when compared against the effect of pupil test scores (*gcsescore*). All of these results are largely confirmed when we look at *specs*65, the categorical measure of the policy effect, although the magnitude of the effects reduces somewhat.

In the matching analysis we showed that attendance at a specialist school is at least correlated with higher test scores, which raises the possibility of an indirect effect on post-school and labour market outcomes. To capture this indirect effect we interact gcsescore and specscore, and our estimates suggest that the combined effect of a school with a higher underlying propensity of becoming a specialist school and pupils with higher test scores lead to increase in the probability that pupils will stay on beyond the age of 16. This effect persists for waves 1 and 2 and in magnitude implies a modest 1 percentage point increase in that probability. The effect becomes insignificant by wave 3 which is unsurprising given the descriptive evidence presented above. The indirect effect of specialist schools is to reduce by 0.5-1 percentage point the probability of being in employment in waves 1-3, which counters the direct effect discussed above. However, since the indirect effect of specialist schools on the probability of a pupil being employed is much lower than the direct effect, overall we can still claim that attendance at a specialist school does seem to improve pupil prospects in the labour market. Interestingly, there is also a very small but statistically significant indirect effect of specialist schools on the probability of being NEET.

Table 5 reports the wage premia associated with prior attendance at a specialist school.[8] There is some evidence that graduates from such schools earn more than their counterparts in non-specialist schools, however, these effects are statistically insignificant. The indirect effects are also negative, which does seem perverse, and statistically insignificant.

## 5.3 Post-school educational attainment

We turn now to the effects of the specialist school policy on post-secondary school educational attainment, which refer to qualifications obtained either through further education at colleges

[8] Note that the sample in each YCS survey ranges between 1000 and 3800 wage observations. In YCS9-10 and YCS11 the proportion of individuals in specialist schools is small and in each wave we observe between 200 and 700 wages. In YCS12 the sample is quite balanced and more than half of the observed wages are for individuals in specialist schools. Sample size is likely to effect the precision of our estimates.

or schools. We distinguish between performance in academic subjects (A level exams), which are the main stepping stone into higher education, and the number of A level subjects studied. The results of our analysis are reported in Table 6 and are limited to wave 3 of YCS12. In terms of performance in A level exams we compute the total points score and it is clear that the direct effect of attendance at a specialist school is to reduce this by about 18 points. There is a modest positive counteracting indirect effect. These findings appear odd at first sight, however, it may be that because pupils from specialist schools take more A level subjects (see the largest positive direct effect) there efforts are spread and hence test performance falls.

# 6 Conclusion

This paper is the first paper to attempt to measure the effect of a UK government education policy, the specialist schools initiative, on post-school outcomes. Specifically, we seek to evaluate whether the specialist schools policy had direct and/or indirect effects on the post-compulsory school outcomes of school leavers and the early labour market outcomes of young adults. To do this we use data for several cohorts of the YCS, which is combined with school level data, and then employ matching methods and a two stage control function approach to attempt to control for various forms of bias that are likely to be present in cross-sectional analyses.

Initially we demonstrate that specialist school pupils have higher test scores and lower probabilities of truanting after allowance is made for observable (and unobservable) differences between pupils. Our effects are somewhat larger than those found in studies that employ difference-in-differences methods. Consequently, although our models do seek to control for the different types of of school and pupil selection bias, we stop short of claiming that specialist school attendance has a causal effect on these school outcomes. Nevertheless, they are indicative of an effect, which in turn implies that specialist schools may have an indirect effect on post school outcomes (e.g. via test scores) as well as a possible direct effect. Our analysis of the direct and indirect effects of the specialist schools policy on post-school outcomes, such as labour market status, A-Level qualifications and wages, is again suggestive of statistically significant effects. There are several interesting findings. First, we find a statistically significant direct effect of the policy on the probability of finding a job after leaving school and this effect increases and persists up to the age of 19 (wave 3). It should be noted, however, that the indirect effect of the policy on the employment probability is smaller and negative, which implies that the net effect is still positive. Second, we find no effect of the policy on wages although we have very few observations for this part of the analysis. Third, the direct effect of the policy on A-Level test scores is negative, statistically significant, and large but there is a small countervailing indirect effect. This is counter-intuitive, however, we also find that pupils who attended specialist schools study more subjects and hence their effort is spread more thinly which may explain the fact that their average A-Level scores are lower than their counterparts from non-specialist schools.

What this analysis suggests is that when evaluating the costs and benefits of educational policies it is important to look beyond the effects on school outcomes, which is itself a contentious issue because of the problems of identifying causal effects, to also consider their effects (direct and indirect) on post-school outcomes. By doing so policy makers will be able to get closer to evaluating whether a particular policy initiative delivers value for money. Our own work is preliminary. Ideally, we would like to have more cohorts of data to analyse and we would like to observe individuals in later stages of life so that we can obtain more robust, and reliable, estimates of the policy effects but also so that we can address the question of whether educational policies have long term labour market effects.

# A Appendix: Testing the validity of the CIA assumption

The CIA is not directly testable, because the data are uninformative about the distribution of $Y_i(0)$ for the treated and of $Y_i(1)$ for the control group. We therefore use two indirect tests from the literature. One test, developed by Imbens (2004), proposes an indirect way of assessing the CIA, based on the estimation of a 'pseudo' confounding factor that should, if the CIA holds, have zero effect. We adopt the method proposed by Ichino et al. (2008). This is based on the prediction of a confounding factor, $A$, by simulating its distribution for each treated and control unit. Then, estimates of the average treatment effect of the treated (ATT) are derived by including the confounding factor in the set of matching variables. Different assumptions on the distribution of $A$ imply different possible scenarios of deviation from the CIA.

For simplicity, let $A$ be a binary variable, its distribution is given by fixing the following parameters

$$P(A = 1|T = i, Y = j) = p_{ij} \quad i, j = 0, 1$$

Where $Y$ is a binary test score outcome (e.g. 'high ability'$= 1$ and 'low ability'$= 0$) and $T$ is the pupil's treatment status. In this way, we can define the probability of $A = 1$ in each of the four groups identified by the treatment and the outcome.[9] In our analysis, we assume that the confounding variable follows the same distribution as that of the pupil test scores prior to entry to secondary school (i.e. Key Stage 2 scores). Therefore $A$ can be thought as a measure of the ability of the pupil that secondary schools 'observe' in making selection decisions. The effect of bias on the estimation of the policy varies depending on the dataset used.[10] The variable $A$ is included in the set of variables used to estimate the

[9]The simplifying assumption that the simulation of $A$ does not depend on $X$ does not change the interpretation of the test. For a complete explanation of the test see Ichino et al. (2008).

[10]For example, for the NPD we obtain the following parameters: $p_{11} = 0.76,\ p_{10} = 0.30,\ p_{01} = 0.73,\ p_{00} = 0.29$; $p_{11}$ can be interpreted as the proportion of 'high ability' pupils in specialist schools who get high test scores. In contrast in the LSYPE we get $p_{11} = 0.72,\ p_{10} = 0.28,\ p_{01} = 0.75,\ p_{00} = 0.33$, where there is a

propensity score and the ATT is estimated using the nearest neighbour algorithm.[11] The ATT is re-estimated 500 times, and the values presented in our Tables are an average over the distribution of $A$(See *NN with confounder*).

Given this set up, if the confounded estimates are still significant, but with the same sign and (similar) in magnitude when compared to the 'true' estimates, we can be fairly confident of the robustness of our results.

The second method we adopt has been proposed by Rosenbaum (1987) and involves only one parameter, representing the association of $T$ and $A$, and derives bounds for significance levels and confidence intervals. Specifically, it computes the upper and lower bounds on the Mantel and Haenszel (MH, 1959) test-statistic used to test the null hypothesis of no treatment effect. In particular, $e^{\gamma}$ measures the degree of departure from a situation that is free of hidden bias ($e^{\gamma} = 1$) and we use $e^{\gamma}$ in the range [1,2]; $\gamma$ represents the effect of an unobserved variable on the probability of attendance at a specialist school.[12] The test can be interpreted as the difference in the relative odds of attending a specialist school for two pupils that appear similar in terms of observable covariates, $\mathbf{x}$. If those most likely to go to specialist schools are more able, then there is positive unobserved selection and the estimated treatment effects overestimate the true treatment effect. In general, DiPrete and Gangl (2004) stress that the results of this test are worst-case scenarios, insofar as they only reveal how the hidden bias might alter inference.

---

slightly higher probability of 'high ability' pupils obtaining high test scores attending non-specialist schools.

[11] We omit the results for different matching methods because they are very similar.

[12] Thus $Pr(D_i = 1|x_i, u_i) = F(\beta x_i + \gamma u_i)$ is the probability of attending a specialist school and $F$ is the logistic distribution. The odds that pupil $i$ attends a specialist school is given by $\frac{P_i}{(1-P_i)} = \exp(\beta x_i + \gamma u_i)$, and the odds ratio of receiving this treatment is $\frac{\frac{P_i}{(1-P_i)}}{\frac{P_j}{(1-P_j)}} = \exp(\gamma(u_i - u_j))$. For simplicity, $u$ is assumed to be a dummy variable and the previous equation may be rewritten as $\frac{1}{e^{\gamma}} \leq \frac{\frac{P_i}{(1-P_i)}}{\frac{P_j}{(1-P_j)}} \leq e^{\gamma}$. In our work, we apply the routines *mbound* and *rbounds* available in Stata. A detailed explanation of the method can be found in Rosenbaum (1995), Aakvik (2001), DiPrete and Gangl (2004).

# B References


Backes-Gellner, U. and Veen, S. (2006) 'Incentives for schools, educational signals and labour market outcomes', Working Paper No. 9, Institute for Corporate Strategy and Business Economics, University of Zurich.

Bradley, S. Migali, G. and Taylor, J. (2009) 'Evaluation of the impact of funding and subject specialisation on school quality using matching models', Working Paper, Department of Economics, Lancaster University Management School.

Bryson, A. Dorsett, R. and Purdon, S. (2002) 'The use of propensity score matching in the evaluation of labour market policies', Working Paper No.4, Department of Work and Pensions.

Cappellari, L. (2004) 'The effects of high school choices on academic performance and early labour market outcomes', mimeo, Università del Piemonte Orientale, Italy.

DiPrete, T. and Gangl M. (2004) 'Assessing bias in the estimation of causal effects: Rosenbaum bounds on matching estimators and instrumental variables estimation with imperfect instruments', *Sociological Methodology* 34, 271-310.

Heckman, J. Hichimura, H. and Todd, P. (1997) 'Matching as an econometric evaluation estimator: Evidence from evaluating a job training programme', *Review of Economic Studies* 64, 1017-1098.

Levacic, R. and Jenkins, A. (2004) 'Evaluating the effectiveness of specialist schools', Centre for the Economics of Education, London School of Economics, London.

Margolis, D.N. and Simonnet, V. (2003) 'Educational track, networks and labor market outcomes, Institute for the Study of Labor, Bonn, Discussion Paper 699.

Smith, H. (1997) 'Matching with multiple controls to estimate treatment effects in observational studies', *Sociological Methodology* 27, 325-353.

Table 1: Dependent variables

| *Panel A: test scores* | YCS11-12 | | | | | | |
|---|---|---|---|---|---|---|---|
| | *Gcsescore* | *Gcsebin* | | *Gcsebin10* | | *Truancy proportions* | |
| | mean | −5 | +5 | −10 | +10 | Non-Truant | Truant |
| Non-specialist (YCS11) | 43.124 | 54.19 | 52.90 | 55.73 | 37.85 | 51.39 | 57.04 |
| n | 2,796 | | | | | | |
| Specialist (YCS12) | 45.904 | 45.81 | 47.10 | 44.27 | 62.15 | 48.61 | 42.96 |
| n | 2,448 | | | | | | |
| Total | 5,244 | 1,705 | 3,539 | 4,536 | 708 | 3,573 | 1,648 |

| *Panel B: post-school test scores* | YCS12 | | | | | |
|---|---|---|---|---|---|---|
| | *A level score* | | | *A level number* | | |
| | *C* | *T* | *n* | *C* | *T* | *n* |
| mean | 5.01 | 5.35 | 5.12 | 2.24 | 2.40 | 2.30 |
| s.d. | 6.06 | 6.19 | 6.10 | 2.12 | 2.08 | 2.11 |
| n | 1565 | 802 | 2367 | 1565 | 802 | 2367 |

T = Specialist

C = Non-Specialist

Figure 1: The test score performance of specialist and non-specialist schools over time

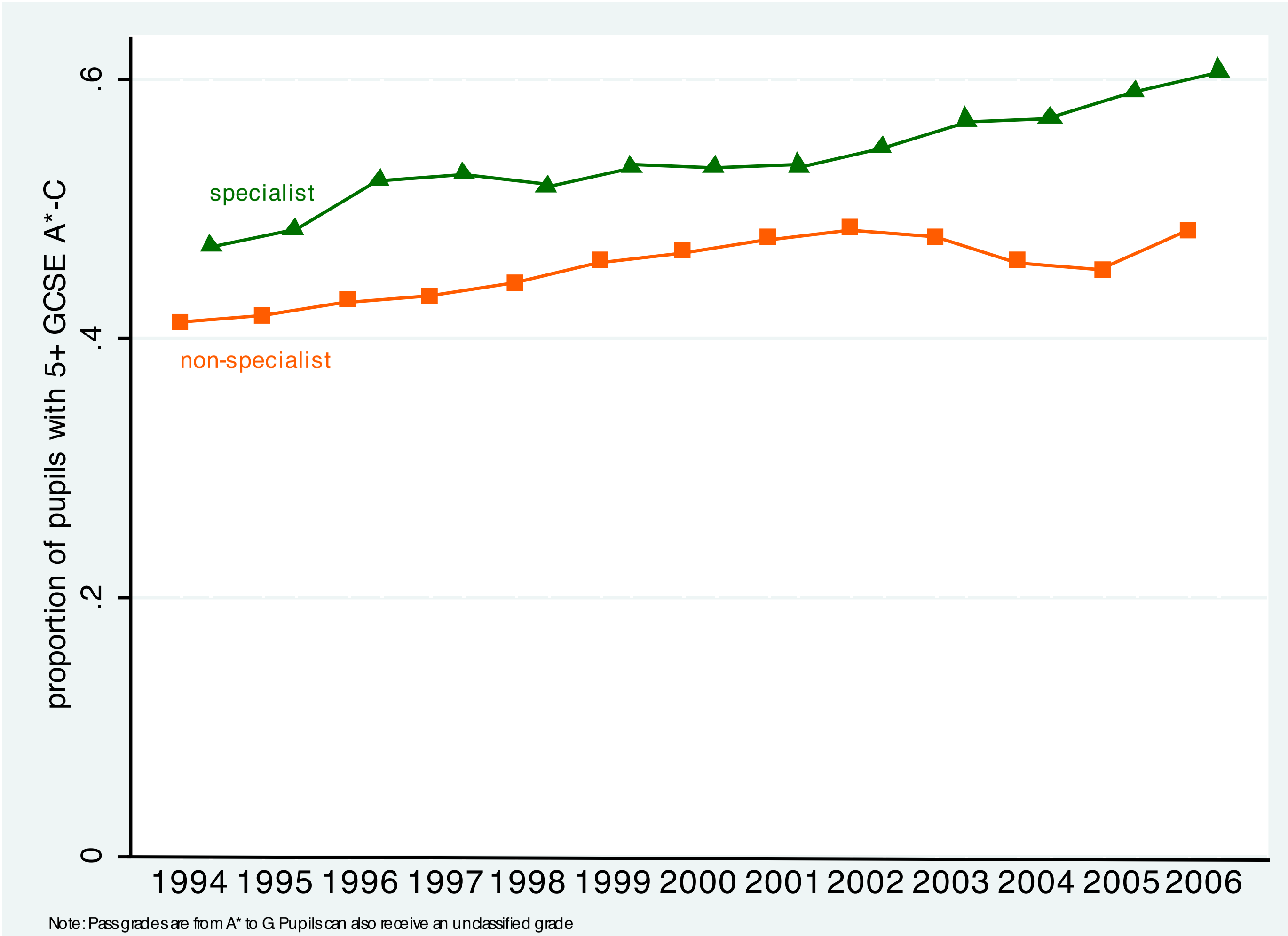

Table 2: Dependent variables Labor market outcomes

| | *Wave 1* | | *Wave 2* | | *Wave 3* | |
|---|---|---|---|---|---|---|
| | *C* | *T* | *C* | *T* | *C* | *T* |
| | | | *Percentages* | | | |
| Stay on =1 | 76.67 | 79.46 | 52.03 | 55.79 | 28.39 | 29.24 |
| Stay on =0 | 23.33 | 20.54 | 47.97 | 44.21 | 71.61 | 70.76 |
| Employed =1 | 8.46 | 7.51 | 7.85 | 5.82 | 12.62 | 12.15 |
| Employed =0 | 91.54 | 92.49 | 92.15 | 94.18 | 87.38 | 87.85 |
| Neet =1 | 6.89 | 5.45 | 4.65 | 3.57 | 5.90 | 5.57 |
| Neet =0 | 93.11 | 94.55 | 95.35 | 96.43 | 94.10 | 94.43 |
| n | 3,121 | 1,597 | 3,121 | 1,597 | 3,121 | 1,597 |
| | | | *mean* | | | |
| Log real wage | 1.22 | 1.20 | 2.54 | 2.63 | 1.68 | 1.66 |
| n | 398 | 188 | 314 | 136 | 434 | 219 |

T = Specialist

C = Non-Specialist

Table 3: Policy-off policy-on analysis using the YCS11-12

| | *Gcsescore* | *Gcsebin* | *Gcsebin10* | *Truancy* | St.Bias |
|---|---|---|---|---|---|
| unmatched | 2.761 | 0.012 | 0.085 | -0.047 | 5.201 |
| | (0.443) | (0.013) | (0.009) | (0.013) | (4.202) |
| NN1 | 1.732 | -0.013 | 0.072 | -0.035 | 1.601 |
| | (0.593) | (0.018) | (0.013) | (0.018) | (1.438) |
| NN1 $_{\text{with counfounder}}$ | 1.385 | -0.030 | 0.063 | -0.054 | |
| | (0.691) | (0.021) | (0.015) | (0.023) | |
| $\text{Kernel}_{0.1}$ | 1.988 | -0.012 | 0.078 | -0.043 | 0.937 |
| | (0.461) | (0.012) | (0.001) | (0.013) | (0.571) |
| *post-matching regression* | | | | | |
| | 1.866 | -0.006 | 0.080 | | |
| | (0.453) | (0.013) | (0.009) | | |

| | *Bounds M-H statistics* | | | | |
|---|---|---|---|---|---|
| | $e^{\gamma} = 1$ | $e^{\gamma} = 1.25$ | $e^{\gamma} = 1.50$ | $e^{\gamma} = 1.75$ | $e^{\gamma} = 2$ |
| Gcsescore | 2.65*** | 1.0-4.3*** | -0.40-5.6 | -1.55-6.7 | -2.55-7.65 |
| Gcsebin10 | 4.0*** | 2.60-5.45*** | 1.47-6.66* | 0.53-7.72 | 0.11-8.66 |

Std. err. in parenthesis.

Balancing Property and Common Support satisfied.Analytical s.e. for NN, Bootstrap 500 repetitions for Kernel

Confounder follows same KS2 test score distribution in the NPD sample.

Standardized Bias= = $\frac{100\overline{x}_{non-sp} - \overline{x}_{spec}}{\sqrt{s^2_{non-sp} + s^2_{spec}/2}}$ where: $\overline{x}_{non-sp}$ = mean of the non-specialist schools group

$\overline{x}_{spec}$ = mean of the specialist school group, $s^2_{non-sp}$ = variance of the non-specialist schools group

$s^2_{spec}$ = variance of the specialist school group.

Significance of MH statistic bound indicates treatment effect is not sensitive to selection bias.

Note: bounds computed with the kernel method .

Table 4: The effect of the specialist schools policy on labour market outcomes

| | *wave1* | *wave2* | *wave3* | | *wave1* | *wave2* | *wave3* |
|---|---|---|---|---|---|---|---|
| *full time education* | | | | | | | |
| | | | | *restriction 65% pct* | | | |
| specscore | -0.436 | -0.444 | 0.161 | specs65 | -0.058 | -0.057 | 0.026 |
| | (0.141) | (0.283) | (0.244) | | (0.021) | (0.040) | (0.034) |
| Gcsescore | 0.005 | 0.010 | 0.009 | Gcsescore | 0.008 | 0.013 | 0.009 |
| | (0.001) | (0.002) | (0.001) | | (0.000) | (0.001) | (0.000) |
| Gcsescore × specscore | 0.013 | 0.011 | -0.003 | Gcsescore × spec65 | 0.001 | 0.001 | -0.001 |
| | (0.003) | (0.006) | (0.005) | | (0.000) | (0.001) | (0.001) |
| *full time job* | | | | | | | |
| specscore | 0.177 | 0.198 | 0.357 | specs65 | 0.024 | 0.031 | 0.039 |
| | (0.077) | (0.073) | (0.118) | | (0.011) | (0.010) | (0.017) |
| Gcsescore | -0.002 | -0.001 | 0.001 | Gcsescore | -0.003 | -0.002 | -0.001 |
| | (0.001) | (0.000) | (0.001) | | (0.000) | (0.000) | (0.000) |
| Gcsescore × specscore | -0.005 | -0.006 | -0.009 | Gcsescore × spec65 | -0.001 | -0.001 | -0.001 |
| | (0.002) | (0.002) | (0.002) | | (0.000) | (0.000) | (0.000) |
| *NEET* | | | | | | | |
| specscore | 0.058 | 0.127 | -0.115 | specs65 | 0.009 | 0.010 | -0.015 |
| | (0.045) | (0.045) | (0.101) | | (0.007) | (0.006) | (0.014) |
| Gcsescore | -0.002 | -0.001 | -0.001 | Gcsescore | -0.002 | -0.002 | -0.001 |
| | (0.000) | (0.000) | (0.001) | | (0.000) | (0.000) | (0.000) |
| Gcsescore × specscore | -0.002 | -0.003 | 0.003 | Gcsescore × specs65 | -0.000 | -0.000 | 0.000 |
| | (0.001) | (0.001) | (0.002) | | (0.000) | (0.000) | (0.000) |

Std. err. in parenthesis.

Table 5: The effect of the specialist schools policy on wages

| | *wave1* | *wave2* | *wave3* |
|---|---|---|---|
| specscore | -0.112 | 0.900 | 0.057 |
| | (0.533) | (1.713) | (0.468) |
| Gcsescore | 0.000 | 0.016 | 0.003 |
| | (0.004) | (0.015) | (0.003) |
| Gcsescore × specscore | 0.006 | -0.027 | -0.003 |
| | (0.014) | (0.054) | (0.010) |
| | | *restriction 65% pct* | |
| specs65 | 0.025 | 0.066 | 0.001 |
| | (0.045) | (0.186) | (0.026) |
| Gcsescore | 0.002 | 0.013 | 0.003 |
| | (0.003) | (0.013) | (0.001) |
| Gcsescore × specs65 | -0.000 | -0.015 | -0.002 |
| | (0.009) | (0.045) | (0.005) |

Std. err. in parenthesis.

Table 6: The effect of the specialist schools policy on post school qualifications

| | *A level score* | *A level Number* |
|---|---|---|
| specscore | -17.767 | 5.150 |
| | (2.519) | (1.483) |
| Gcsescore | 0.129 | 0.096 |
| | (0.016 | (0.007) |
| Gcsescore × specscore | 0.371 | -0.074 |
| | (0.055) | (0.025) |
| | *restriction 65% pct* | |
| specs65 | -0.804 | 0.043 |
| | (0.194) | (0.068) |
| Gcsescore | 0.194 | 0.075 |
| | (0.010) | (0.003) |
| Gcsescore × specs65 | 0.134 | 0.003 |
| | (0.031) | (0.008) |

Std. err. in parenthesis.