Very preliminary
Comments welcome

Ability Tracking and Student Performance in Secondary Schools in England in Wales

Alan Manning
Jörn-Steffen Pischke
LSE

March 2004

Abstract:

British secondary schools moved from a system of extensive selection and tracking to one with comprehensive schools during the 1960s and 70s. Before the reform, students would take an exam at age eleven, which determined whether they would attend an academically oriented grammar school or a lower level secondary school. We make use of differences in the timing of the reform at the local level to study the impact of the system a student attended on their performance in school, educational attainment, and labor market success. This paper uses data from the National Child Development Study, a cohort panel following individuals who entered secondary school in 1969, in midst of the transition to comprehensive education. We show that areas which switched to a comprehensive system earlier tended to have poorer households and more poorly performing students, so that the raw difference of early comprehensive areas and those remaining selective is not informative. We present regression adjusted results using detailed controls as well as instrumental variables estimates, using the political composition of a county as an instrument. The results indicate that selective schools tend to perform at least as well or better overall, but there may be an advantage for certain children from attending comprehensive schools, particularly those with high ability but poor family background.

Introduction

Schools and school systems differ in the amount of ability tracking of students they provide in secondary school. Some systems (e.g. the US) are based on comprehensive schools, where students of all abilities attend the same school, although there is typically some tracking within schools. Other systems (e.g. Germany) channel students at an early age into different types of schools based on academic ability. British schools moved from a system of extensive tracking to one with comprehensive schools in the 1960s and 70s. We make use of this reform to study the impact of the system a student attended on their performance in school, educational attainment, and labor market success. The British experience is interesting, because it involved a major and well defined change in terms of the ability grouping of secondary school students. This contrasts markedly with some of the US literature on tracking, where the treatment of interest is often difficult to define (see, for example, the debate between Rees, Brewer, and Argys, 2000, and Betts and Shkolnik, 2000a, b, over the classification for US schools).

Our question differs somewhat from the one that has often been asked by education economists so far. Many previous studies have analyzed the effect of attending a more selective school or a higher ability track on the performance of the concerned individual. These studies often find that students attending higher ability tracks or selective schools fare better than those who do not. But these studies cannot address the question how many students should attend a selective school: if everybody attended a selective school, it would be a comprehensive school. Instead, we ask: what is the effect of the availability of selective schools in an area on the performance of students, irrespective of whether they actually attend a selective school.

The British school reform in the 1960s and 70s provides a useful experiment to address this issue. In the traditional system, students were tracked into either an academically selective grammar school at age 11, or they would attend a secondary modern school, which was academically less demanding. Starting in the 1950s, there was dissatisfaction with selection at the local level, and some local authorities began to experiment with comprehensive schools. In 1965, the central government asked the Local Education Authorities (LEAs) to draw up plans to switch to a comprehensive system. The implementation proceeded slowly, with faster growing, more Labour leaning LEAs moving to comprehensive schools more quickly, while those without

expanding numbers of students, and more Conservative leaning Authorities implemented the change more slowly. In fact, there are still a number of LEAs to date, which provide grammar schools as an option.

The LEA level implementation provides ample regional variation in the type of school system students were exposed to in the 1970s. We use the National Child Development Study (NCDS), a panel which tracks members of the 1958 birth cohort. For LEAs in England and Wales, we can match the NCDS with the school composition of each student's LEA around the time when they were age 11. While there were many LEAs, which were in the middle of the reorganization process, we can identify LEAs which were more clearly comprehensive or selective. We argue that it is important to focus on those LEAs only. The characteristics of selective and comprehensive LEAs are not the same, however. We use two strategies to address the selection problem. The NCDS provides a wide array of family background and student ability measures which allows us to adjust for area characteristics when comparing student outcomes. We also use an instrumental variables strategy, using political composition of the LEA conditional on characteristics as an instrument.

The NCDS also provides a wide variety of outcome measures: test scores at age 16, the type of secondary school credentials (O-levels and A-levels) obtained, the highest education level completed, as well as labor market outcomes at age 33. A useful aspect of the data set is that we can analyze results by family background and ability of a student, in order to address the distributional impacts of tracking. While our results are not particularly precise, they indicate that the selective system seems to outperform the comprehensive system. The benefits of the selective system are concentrated on high ability students but low ability students also benefit. Medium ability students are about as well off in either system. However, among high ability students, the benefits of the selective system are concentrated on those students from high SES backgrounds.

At a theoretical level, there are good arguments for selection as well as for comprehensive education. The main argument for selection or tracking is presumably that it is much easier to teach lower variance classes. Since teachers can focus on the ability level of particular groups of students, students of all ability levels might benefit from selection. One argument against

selection is that there might be positive peer effects from the most able students. By tracking these students into separate classrooms, the most able students may benefit from being with each other. However, the lower ability ranges loose from not having this peer group around. We know very little about the different impact of peer group effects on different types of students empirically, so it is difficult to judge a priori whether this indeed leads to lower average performance in a selective system. Another argument, particularly against early selection as in post-war Britain, is that eventual ability levels are difficult to predict at an age as early as 10 or 11. Moreover, secondary selection is based on a single exam, clearly a noisy mechanism. This may result in some kids ending up in the wrong track. This suggests that particularly middle ability kids may loose in the British style selective system. Since there are arguments going either way, the issue eventually is an empirical one.

The findings in the empirical literature about tracking and selection differ widely. Many researchers, looking at tracking in the US and at the British secondary reorganization conclude that the evidence does not support claims to the superiority of either system (see, for example, Figlio and Page, 2002, for the US, and Crook, Power, and Whitty, 1999, for the British case). On the other hand, there are also studies which find more pronounced effects going one way or the other. Jesson (2000) argues that comprehensive LEAs systematically outperform selective LEAs in Britain. Kim, Lee, and Lee (2003) find the exact opposite comparing cities in Korea with selective versus mixed secondary school systems. Since none of the previous studies is without problems, we feel that it is worthwhile to reanalyze the experience in England and Wales.

The remainder of this paper is organized as follows. The next section describes the institutional background of British secondary education, and the history of comprehensive reorganization. Section 3 is devoted to laying out the empirical framework, and compares our approach to some of the existing literature. The following section describes the data and key variables. Results are presented in section 5, and the final section concludes.

Secondary Education in Britain[1]

The classic state supported secondary school in 19th century Britain was the grammar school, an academically oriented and class based school for 11 to 18 year olds, administered independently of the local municipality. Only in the 20th century became municipal grammar schools available to the middle classes but they remained academically selective. When the school leaving age was being raised to 14 in the interwar years, the question arose how to provide this secondary schooling for all. The dominant view both among Labour and Conservative education politicians was that the role of grammar schools should not be touched, and separate secondary schools should be created for the children who would not attend grammars. It was a popular view at the time that psychometric testing could accurately determine the abilities of students at an age as early as 10 or 11. Tracking students into different types of schools was therefore seen as a natural way to proceed, and was typical of most European countries at the time.

During the second world war, the idea emerged that the secondary curriculum should be further differentiated, and in addition to the grammar schools, both secondary modern and secondary technical schools should be created. However, education policy making in Britain was always rather decentralized, with the local educational authorities (LEAs) being the main administrative units, which retained a lot of decision making power about the exact make-up of the local school system. The technical schools never became popular with the LEAs and few were ever created. The 1944 Education Act did not require LEAs to create all three types of schools but rather required them to submit detailed plans for the scrutiny of the Education Ministry. But the Act prescribed separate secondary schools and a transfer at age 11. A typically LEA would therefore offer the selective grammar schools alongside lower level secondary modern schools during the post-war period.

Nevertheless, the 1944 Act left open the possibility for LEAs to experiment with other schemes, including comprehensive schools, even in the immediate post-war era. However, the post-war Labour government was not particularly supportive of comprehensive education plans, particularly because of fears that these schools would have to be too large. On the other hand,

[1] This section draws heavily on the descriptions in Kerkhoff et al. (1996) and Griffith (1971).

early drives for the establishment of comprehensives at the local level were typically, although not exclusively, in Labour dominated urban areas, like London, Bristol, and Coventry.

Some of the Conservative Education Ministers in the 1950s and 60s, like David Eccles and Edward Boyle, were actually more progressive in allowing experimentation with comprehensive schools than their Labour predecessors. The existing consensus about secondary schooling at the national level also began to crack in the 1950s. The differentiation into three types of schools had failed. There was growing unease about the selection process using the 11+ examination, which determined the admission to grammar schools. Moreover, it became obvious that there were other alternatives to a selective system than large comprehensive schools serving all children age 11 to 18 at once. The Leicestershire LEA, for example, began to experiment with abandoning the 11+ and creating a comprehensive school up to age 14.

By the early 1960s, many, if not most, LEAs were working on reorganization plans, which were trying to end the traditional selective system, or were challenging it in one way or another. While Labour led LEAs played a leading role in this development, the trend cut across party lines, with some Conservative authorities being among the most fervent advocates to ending selection. As a response, the 1964 Education Act passed by the Conservative government abandoned the principle of school transfer at age 11.

Only in 1963 had the Labour Party fully embraced the comprehensive principle and called for the end of selection at their party conference. When Labour came to power in 1964, its goal was to accelerate the existing trend for comprehensive reorganization, which was well underway at the local level. However, there was a great diversity of views on how this reorganization should be achieved. Rather than compelling LEAs on a particular system of comprehensive schooling, the government issued Circular 10/65 in July 1965, requesting local authorities to submit detailed plans on how to establish comprehensives. The circular itself suggested no fewer than six different models of comprehensive reorganization.

The model of a three tiered school system with a comprehensive middle school received a boost from the Plowden Report in 1967, a government commissioned study on the best age of school transfer. While the members of the Plowden Committee favored an 8 – 12 year middle school, many LEAs opted for a 9 – 13 model. The middle school model often turned out to be a popular

model with Conservatives, because it allowed to retain the grammar schools as the upper schools (an example is the Leeds LEA). Nevertheless, in practice, most comprehensive schools eventually became 11 – 16 or 11 – 18 schools.

By the mid-1960s, secondary education in England and Wales was rather diverse. In some LEAs, comprehensive reorganization had been well under way for years. The majority were drawing up plans to respond to the central government's request in some way or another. A small number of LEAs resisted the call by the central government completely. This diversity was not limited to the LEAs but often extended to divisions and districts within the LEAs. Overall, the period from 1965 to 1974 was one where new comprehensive schools were established in most parts of the country, and the fraction of students served by these schools increased dramatically. In 1965, there were still only 262 comprehensive schools in England and Wales. By 1974 this number had grown to 2,677, and these were attended by 62 percent of secondary students in LEA maintained schools (see Figure 1).

One of the key issues of contention during this period was less the establishment of comprehensive schools, as the closing of existing grammar schools, some of which had a long tradition. Parents, teachers, and school administrators often resisted the closing of grammar schools, and together with the support of local or national policy makers, this resistance was often successful.[2] Some grammar schools avoided closure by leaving the state sector and becoming independent (i.e. private) schools. As a result, new comprehensive schools often coexisted with the traditional grammar schools in many places, hence perpetuating selection despite reorganization.

An even thornier issue than the grammar schools maintained by the LEAs were posed by the existence of direct grant grammar schools and independent schools. The direct grant schools are privately run, often denominational schools, financed by endowments and fees, who receive part of their funding as a grant from the central government. In return for this funding, they were required to offer 25 percent of their places to local authorities. In practice, the LEAs paid fees

[2] But old age was no insurance for a grammar school's continued existence. For example, Derby School, founded about 1160, was abolished in 1972 in comprehensive reorganization. Katharine Lady Berkeley's School in Wotton Under-Edge, Gloucestershire, founded in 1384, was changed into a comprehensive school in 1973. In fact, we found no correlation between the number of grammar schools in existence in the 17th century and comprehensive reorganization.

for about 60 percent of the students at the 178 direct grant schools in the late 1960s. Since these schools were outside the local control, there was little LEAs could do to integrate them into their comprehensive schemes, other than withdrawing the local take up of places.

There was also great variety in the form the new comprehensive schools took. Some were purpose built, others were amalgamated out of existing schools, often on different physical sites. Some served pupils until they reached the university entrance exams at age 18 (A-levels), while others ended at 16 and students who wanted to stay in school longer were served by separate so called sixth form colleges.

The pace of reorganization was very different in different places as well. Because of the reluctance of many LEAs to close grammar schools, there was a growing view within the Labour Party during the late 1960s that the government should pass legislation to force laggard LEAs to adopt the comprehensive principle. But the Labour government lost the general election in 1970, before a bill could be passed. The new Conservative Education Secretary, Margaret Thatcher, immediately issued a new circular, 10/70, effectively giving LEAs more freedom in determining the pattern of secondary schooling they wanted to pursue. This change in national policy was unable to stem the existing tide, and, in fact, more comprehensives were opened under Mrs. Thatcher's helm at the Education Department than under any of her predecessors or successors. Nevertheless, while the change in the government did not reverse the longer-term trend, the comprehensive movement never regained the same thrust it had during the 1960s. Comprehensive reorganization, while concentrated in the 1960s and 70s, does continue into the 1990s (see Figure 1). A few LEAs still maintain grammar schools, and the political discussion about the merits of selection continue in Britain until this day.

Our research design relies on the difference in the speed of reorganization at the local level. What aspects mattered for whether an LEA was sufficiently advanced to offer an essentially comprehensive education for the cohort turning 11 in 1969? The political makeup of the LEA often mattered in the speed of reorganization. Urban, Labour dominated LEAs were often the vanguards of the comprehensive idea in the 1950s, and hence probably responsible for the national trend. On the other hand, they were not necessarily the LEAs to implement

comprehensive reorganization most completely. An example is London, which had opened 59 comprehensives as early as 1961, but was slower to push reorganization further in the 1960s.

Conservatives or particular individuals in the Conservative party were often at the forefront of ending selection in certain LEAs (examples are Leicestershire and Leeds), typically using a middle school model. The Conservative landslide in the 1967 municipal elections did not really turn back the momentum towards reorganization but it often fell to Conservative councils to implement the reorganization plans drawn up earlier, sometimes leading to small changes or slowing down reorganization. Determined education officials at the local level clearly mattered. Reorganization often happened quickly in LEAs, where a Chief Education Officer (CEO) with interests in the comprehensive idea worked hand in hand with an interested Labour Council (examples are Stoke-on-Trent and Bristol). Where the ideas of politicians and administrators clashed, the process would be slower (as in Manchester). Where the local politicians were uninterested in education matters, the CEO could often emerge as a powerful figure, and enable wide ranging transformations, sometimes with national implications (as in Leicestershire, which pioneered the middle school model). However, these powers were often checked by forces in the local districts within the LEA, which could thwart the pace of change (as in the West Riding of Yorkshire or West Sussex).

Kerkhoff et al. (1996) provide a statistical analysis of the determinants of reorganization using the NCDS data. They find that local political control as the only significant predictor. However, their results also imply that political control predicted strongly what type of comprehensive scheme was be implemented. Conservative LEAs were more likely to prefer middle school models, and they were more likely to have separate sixth form colleges. These features are often associated with outcomes most resembling the selective system: the students entering the schools leading to A-levels would be more likely to come from more advantaged family backgrounds. This is a potential problem for our strategy of using political control as an instrument.

Our investigation focuses on England and Wales, which went through the protracted transformation process just described. Scotland transformed to a comprehensive system more quickly, and without much local discretion. Nevertheless, Scotland is not particularly useful as a comparison group because it has a very different educational system from England and Wales

with its own school leaving exams, university system (undergraduate degrees taking 4 years compared to 3 in England). Northern Ireland kept the selective system during this period but it is not included in the NCDS, and also differs sufficiently from the rest of the UK to make a comparison difficult.

Research Design and Comparison with the Previous Literature

We would like to investigate how a student or set of students who were educated in a comprehensive school would have fared, had they been part of the selective system instead. Beyond the impact of the secondary system on the average outcomes on students, we are interested in the distributional aspects of the policy change: for example, do high ability or low ability students benefit more from one system or the other. The main challenge in the evaluation of these questions is the definition of an adequate control group for the comprehensive treatment. Ideally, the control group should be as similar as possible to the treatment group, except for the secondary school selection or tracking that we are interested in.

It is our view it is necessary to compare students in comprehensive *LEAs* to students in selective *LEAs* rather than comparing students in comprehensive *schools* to students in selective *schools*. The US equivalent to this would be to compare tracked schools to un-tracked schools rather than comparing students in different tracks within a school to students who are not tracked. However, the latter has been the predominant approach in the literature. In the US literature on tracking, it has been customary to run a regression of the form

$$y_i = \alpha + \beta_1 H_i + \beta_2 L_i + X_i \gamma + \varepsilon_i$$

where y_i is a student outcome, like a test score, H_i is an indicator for a student attending a high ability track, and L_i is an indicator for a student attending a low ability track, and untracked students are the base category. The control variables X_i might contain characteristics like previous test scores. Argys, Rees, and Brewer (1996) is a widely cited paper proceeding along these lines. In the British context, the analogy would be to run a similar regression on students in all LEAs, and let H_i denote student attending grammar school, L_i students attending secondary

modern schools with comprehensive students being the base category. Galinda-Rueda and Vignoles (2004) present some regressions of that type.

Not surprisingly, studies proceeding along these lines find that students in high ability tracks or grammar schools do better while students in low ability tracks or secondary modern schools do worse than untracked or comprehensive school students. The authors have often interpreted these results as evidence that tracking tends to exacerbate educational inequality, and hurt low ability students. We find this approach highly problematic. Test scores and other controls may be imperfect, so assignment to a high or low track may easily pick up unmeasured student characteristics. The results may simply reflect omitted variables bias.

Figlio and Page (2002) criticize this approach for the same reasons. In order to assess the distributional aspects of tracking, they comparing high ability students in tracked schools to high ability students in untracked schools, and similarly for low ability students. Using this approach, they find no difference between tracked and untracked students in three different ability ranges. Figlio and Page (2002) also discuss the potential problem that tracked and untracked schools may not be completely comparable. To address this issue, they present IV estimates, using interactions between state graduation requirements, the number of schools in the county, and voting patterns as instruments. However, the motivation for these instruments, as the authors admit, is "not immediately transparent."

In the British context, it is therefore more natural to compare comprehensive schools to selective schools. This has been the approach in Kerkhoff et al. (1996) and Galinda-Rueda and Vignoles (2004), both studies utilizing the NCDS data we also use below. However, as we argued above, comprehensive and selective schools coexisted in the majority of LEAs at the time the NCDS cohort entered secondary schools. The bulk of the identification will therefore come from these mixed LEAs. But if grammar schools coexist alongside comprehensive schools in the LEA it is unclear whether the comprehensive schools will really have an intake which is representative of all students or whether the grammar schools will keep cream skimming the best students. The comprehensive schools would effectively just be relabeled secondary modern schools in this case.

Kerkhoff et al. (1996) acknowledge this problem, and also compare only LEAs which are either purely selective or purely comprehensive. However, they downplay this approach because it results in relatively small sample sizes. However, this is the approach we prefer, because it is the approach most likely to compare students facing a truly comprehensive system with those in a selective system. This approach is not without problems, because LEAs which switched to the comprehensive system early enough for the 1969 entering cohort may be different from those which switched later. Although Kerkhoff et al. (1996) claim to find few differences other than political affiliation of the area, we document below that comprehensive LEAs are systematically poorer and have lower ability students.

We try to deal with this problem in two ways. First, we rely on the wealth of information collected at age 7 and 11 in the NCDS and use these pre-treatment characteristics as conditioning variables. Our comparisons will identify the treatment effect of being in a comprehensive LEA, as long as these conditioning variables include all relevant factors, which determine student success apart from the comprehensive treatment. Obviously, these same covariates are also available when comparing selective and comprehensive schools within an LEA, as in the previous literature. Nevertheless, we believe that this strategy is more believable when using only the between LEA variation. Whether a student attended, say, a comprehensive or a grammar school within an LEA is a question of individual school choice. The covariates therefore need to reflect all the factors relevant to that individual choice. This is a tall order, since many of the covariates are presumably measured with error. For example, a student with an intermediate test score at age 11, who actually attends a grammar school, may in fact be more able than reflected in the test score. Across LEAs we only need to control for the fact that the average characteristics of students in comprehensive and selective LEAs differ. The same type of measurement error is therefore likely to be of much less consequence.

Alternatively, we also present IV estimates using political control of the LEA, conditional on the same covariates, as an instrument. This instrument will be valid if, conditional on the covariates, which include socio-economic variables, Labour and Conservative controlled areas are similar apart from comprehensives being more likely in Labour areas. Galinda-Rueda and Vignoles (2004) also present IV results using election outcomes. Since comprehensive reorganization was decided at the LEA and possibly the district level, it fits more naturally in our setting of

comparing LEAs rather than schools. We use results from the local elections in 1961 and 1967, while Galinda-Rueda and Vignoles (2004) rely on the national election results in 1974. Local control in the 1960s was clearly what was relevant for local education decisions at the time. In fact, one might be worried that election outcomes after comprehensives were established reflect the experience with these schools, and hence are an outcome of the process.

We also tried a number of other instruments, including the prevalence of direct grant grammar schools in the area, measures of the fraction of grammar schools with a long heritage (both of these may make it more difficult to establish a comprehensive system), and population growth in the 1950s and 60s (which should have led to more funding for new school construction, which could have been used to build comprehensives). None of these variables resulted in a viable first stage.

Kim, Lee, and Lee (2003) study a similar reorganization from a selective to a comprehensive secondary school system in Korea. Access to Korean high schools used to be based on entrance exams, which made the system highly selective. Since the 1970s, the Korean government has pursued an equalization policy, abolishing entry exams and forcing high schools to admit mixed ability cohorts. This was introduced at different times in different areas. The approach of Kim, Lee, and Lee (2003) is very similar to ours, as they compare mixed and selective areas. They find students at all ability levels to perform better in the selective system. Their study is not without problems either. Areas differ in terms of student test scores at high school entry. Since they use data from a national testing system, they only have a limited set of covariates, and they rely heavily on pre-treatment test scores for a different cohort of students. Given the evidence they present, it is unclear whether test score gains are expected to be similar in selective and mixed areas.

Data and Definition of Treatment

Our main data source is the National Child Development Study (NCDS). The survey includes everyone born between March 3 and 9, 1958. After a parental survey at birth, there were follow-ups at ages 7, 11, 16, 23, 33, and 41. The original survey included 17,414 births. However, the

study has had significant attrition, and by the time of the fifth wage at age 33, information is only available for 11,407 sample members. Extensive information on the life course of each individual is available, including interviews with the sample member, parents, teachers, health professionals, and results on aptitude and medical tests, as well as on official school leaving exams (O-levels and A-levels). Our sample includes all individuals in England and Wales, on whom there is information from the schools in the age 7, 11, and 16 surveys.

In terms of outcomes of interest we analyze test scores at age 16, O-level and A-level exams passed, various measures of the highest educational attainment, as well as two labor market outcomes: full time employment and wages at age 33. A variety of tests were given to the NCDS sample, and each test is scored on its own scale. In order to make results easier to read, we have converted all test scores to a 0 – 100 scale. The same reading test was given to the sample members at ages 11 and 16, and the test was not really appropriate at age 16. As a consequence, we rely mostly on the math test results at 16.

NCDS sample members will be age 11 in 1969, the typical age of secondary school transfer. They typically continue in school until age 16 or 18. Hence, we would like to compare individuals in areas in which there were no options to attend grammar schools by 1969 anymore to those, where grammar schools still served a large number of students in 1974. Many LEAs, of course, will not satisfy either of these criteria, because they will be in the middle of the reorganization.

In order to classify LEAs as either comprehensive or selective, we started with data on school attendance by LEA from Education Statistics for the years 1967, 1971, and 1974. We created two different definitions of comprehensive and selective areas: a narrow one with more stringent criteria, and a wide one with more lenient criteria.[3] We will focus on showing results for the wide definition of a comprehensive area, which requires that at least 75 percent of LEA students attend comprehensives in 1971 and 1974, more than 10 percent attend grammar schools in 1967, but less than 5 percent attend grammar schools in 1974.[4] We require a selective area to have less

[3] We also experimented with broader definitions of the treatment and control groups. The groups so defined tended to look more alike, but these definitions might also result in some degree of misclassification and hence attenuation of the estimates. The precision of the estimates was not substantially improved.

[4] We require at least 10 percent in grammar schools in 1967 because similar aggregate data are not available before 1967. Hence we cannot effectively distinguish areas which have few students in grammar schools because they

than 20 percent of students in comprehensive schools in 1971, and less than 40 percent in 1974. In addition, in 1974 at least 10 percent of students have to be in grammar schools and we also require the 1974 fraction in grammar schools to be at least 80 percent of the 1967 fraction. This is supposed to rule out that significant reorganization is already under way while our sample is in school. The narrow definition tightens the criteria for a comprehensive area, requiring that at least 75 percent of LEA students attend comprehensives already in 1969 instead of only in 1971. Some of the criteria are somewhat arbitrary, but we show below that they result in a sensible delineation of the treatment and control groups.

After this first step, we manually went through the descriptions of secondary reorganization plans in Comprehensive Schools Committee (1967) and Benn (1971) and used these to adjust the original classification. We change both definitions of comprehensive LEAs is to include 8 LEAs which had clearly reorganized before 1967. On the other hand, we excluded 4 LEAs from the selective group, because they already had a comprehensive system in place in some areas in 1971. Details are again documented in the appendix. The NCDS distinguishes 164 LEAs in 1969.[5] According to our wide classification there are 25 comprehensive and 29 selective LEAs with observations in the sample, while 110 LEAs are not used. The narrow classification excludes 8 further comprehensive LEAs.

We supplemented the NCDS data with some area characteristics from the 1971 Census, which we merged with the 1969 residential LEA of the sample member. In addition, we collected data for the political composition of the county borough, which we use as instrumental variables. Local political control in Britain is organized in a number of counties. Some cities are administered independently of the counties (these are called county boroughs). The LEAs outside Greater London in 1969 coincided either with a county or county borough. We collected information on the outcomes of local elections in 1961 and 1967 from *The Times*. Where this information was not available in *The Times*, we supplemented it with information from the Municipal Yearbook (1962, 1968). The reports on election outcomes do not match completely between these two sources. Since *The Times* appeared to present a more complete coverage, we used it as the primary source. *The Times* typically also reported which party held control of the

always offered few grammar school places (but tend to remain selective) and those which have reorganized along comprehensive lines.

[5] The NCDS combines some of the LEAs in Greater London, so that none of these LEAs can be used.

county council, which is the variable we eventually use. Where this information is missing, and one party held the majority of the seats, we assigned control to that party. There are a few counties and county boroughs, which were established between 1961 and 1969. In these cases, we assigned political control in the main predecessor area for the years when the new LEA was not in existence.

Empirical Results

Table 1 shows the percentage of students attending different types of schools during the period we analyze. In the narrow definition, the vast majority of students attended comprehensive schools by 1969, i.e. at the time when the NCDS cohort enters secondary school. Only 3 percent of students still attend grammar school in 1969, and this fraction drops to zero by 1974. In the wide treatment group definition, about 10 percent of students are still in grammar schools in 1969, although almost all are in comprehensives by 1971.

The respective control groups of selective LEAs are the same in both definitions. These areas also start opening comprehensive schools during this period, although these tend to be very few in number in 1969, and even by 1974 do not exceed 10 percent. Most importantly, the growth in these comprehensives primarily took place at the expense of secondary modern schools, and not grammar schools. The number of students in grammar school stays constant from 1967 to 1974. Hence, the control groups remained effectively selective during this period, despite opening a few comprehensives.

The NCDS allows us to compare the treatment and control LEAs along a large number of pre-treatment characteristics. This is done in Tables 2 (for age 7 variables) and 3 (for age 11 variables) for the wide treatment group definition. The tables show means in the selective LEAs and the difference in comprehensive LEAs together with the associated t-statistics. It turns out that comprehensive LEAs differ from selective LEAs in a systematic fashion: They tend to be poorer as evidenced by lower SES of the parents, a worse housing stock, and more public (council) housing. Student performance tends to be worse in comprehensive LEAs, both according to teacher assessments and according to formal tests. The school environment of the

students does not differ much between the treatment and control areas (in terms of class size, school size, age at entry, and in terms of curriculum). Interestingly, the incidence of ability tracking at age 7 or 11 also does not differ much between these areas. The table also shows that all the Welsh LEAs using these treatment group definition are comprehensive. We will use a control for Wales, which means that the effective comparison is only between English selective and comprehensive LEAs.

Table 4 demonstrates how the age 16 experience of the students differs in comprehensive and selective LEAs. About 12 percent of students in selective LEAs attend a comprehensive school, but over 90 percent in comprehensive LEAs. The opposite pattern occurs for selective schools (grammars and secondary modern schools). Moreover, students in comprehensive schools in comprehensive areas spent on average almost four years of their secondary schooling in comprehensive schools, while those in comprehensive schools in selective areas attended these schools for less than two years.

Students in comprehensive LEAs are less likely to be in independent (private) schools, both at ages 16 and 11. This presumably reflects the fact that these are poorer areas. Most important for our purpose, there is no indication that parents sent more children to private school in response to an area changing their system to comprehensive: the incidence remains around 2 percent from age 11 to 16.

The following rows in table 4 demonstrate that the schooling experience in comprehensive and selective areas differed in a variety of ways. Comprehensive schools tended to be much bigger. We defined measures of whether a student is tracked in his or her school at ages 12 and 16. We automatically classified anybody attending a selective school as being tracked. However, given the size of comprehensives, they also frequently used ability tracking within the school. Obviously, most students in selective areas will be tracked by construction but the incidence of tracking is almost as high in comprehensive areas at age 16 (although it is substantially lower at age 12). This means that even comprehensive schools are not fully comprehensive: ability tracking, particularly in math, was a widespread phenomenon. It is important to keep this in mind when interpreting the results. Although the comprehensive environment is going to be much different from the selective system, it still involves ability grouping.

Comprehensive schools also involved a number of other changes. Many selective schools in Britain tend to be single sex: about 40 percent of students in our selective sample attended single sex schools compared to only about 12 percent in comprehensive areas. Although pupil teacher ratios and class sizes are roughly comparable in table 3, the class size differences rise above one once we control for age 7 and 11 covariates. The fractions of students in the schools who attempt O-levels or A-levels or go on to university are markedly higher in comprehensive areas. It is important to keep in mind that in the selective areas, the schools were a significant fraction of students take these exams are the grammar schools, while they would be much lower in secondary modern schools. This implies that the exposure and access to exams for lower achieving students has become much greater in comprehensive schools.

Hours in core subjects like math and English are slightly lower in comprehensive areas. But while these differences are almost significant, they are numerically relatively small. Comprehensive areas are also more likely to use more severe disciplinary measures. It is unclear whether this is the result of more discipline problems in these areas because of the characteristics of their student intake, because of the fact that these schools are bigger, and hence more likely to produce discipline problems, or whether the approach to discipline simply tends to be different in comprehensive schools. Interestingly, the comprehensive areas are less likely to actually expel students for discipline problems, and this pattern remains controlling for age 7 and 11 covariates, pointing to a different approach to discipline.

Parents in comprehensive areas are basically as satisfied with the schools as those in selective areas (and the small difference is not significant). Students in comprehensive areas, on the other hand, have a much more negative attitude towards their schooling: they are more likely to dislike school and homework. These differences again remain when we control for age 7 and 11 covariates, indicating that the comprehensive school experience may indeed be worse for students. This could be due to the nature of the comprehensive schools, or to the fact that these schools were very recently established, and therefore functioning less smoothly than the schools in selective areas. Hence, it is important to keep in mind that the comprehensive treatment involved a variety of differences compared to the selective system over and apart from ability tracking and secondary school selection.

Table 5 shows our baseline results for the wide treatment group definition. Columns (1) and (3) show the raw differences. These look similar to the age 11 differences in test scores: students in the comprehensive areas performed worse than those in the selective areas. The difference for the math test is 5.6 points. The between LEA standard deviation using all LEAs is about 6.6, so that comprehensive LEAs are worse by almost an entire standard deviation.[6] However, some of this difference will be likely due to the fact that comprehensive LEAs tend to have lower ability students to begin with, and not the result of comprehensive schooling.

Columns (2) and (4) attempt to control for these differences using a large set of age 7 and 11 control variables (all the starred variables in Tables 2 and 3). The differences between selective and comprehensive areas shrink dramatically. Nevertheless, students in comprehensive areas still score about 2 points lower in math, and this difference is significant at the 7 percent level. This difference is still economically sizeable, about 30 percent of the between LEA standard deviation in test scores. The difference for reading is zero once controls are introduced. This is not surprising, since the reading test is not really appropriate at age 16.

In order to probe these results, we also provide instrumental variables estimates using the political control in the LEA as instruments. We use the political composition both in 1961 and in 1967. Areas that were more or less fully comprehensive by 1969 must have started planning for comprehensive reorganization early on, most likely before Circular 10/65 was published. Hence, political control in 1961 should be a good indicator for whether an area might have started to move towards comprehensives before 1965. Nevertheless, most plans were only approved after 1965. Since political control shifted to the Conservatives in many areas in the 1967 local elections, this has likely influenced the implementation of reorganization plans. Hence, political control in 1967 might matter independently of 1961 control.

Table 6 shows the first stage regression results. Consistent with Kerkhoff et al. (1996), we find in columns (1) and (2) that both Conservative control and Labour control in 1961 are significant predictors of treatment status with the expected signs. The two measures are not the same, many areas were controlled by independents or no single party had political control. Conservative control works slightly better as a predictor of comprehensive status. Using Conservative and

[6] The student level standard deviation of math scores is 22.6.

Labour control in both 1961 and 1967 together also yields expected results, although Conservative control in 1961 remains the strongest single predictor, and the remaining instruments are not individually or jointly significant. These covariates become more important in column (5), which adds an interaction for Labour control in both 1961 and 1967. Labour control throughout was actually detrimental to comprehensive reorganization, while Labour control in only one of these years was associated with a higher likelihood of becoming comprehensive. Areas that were controlled by Labour throughout this period were probably predominantly working class areas, with a much higher Labour vote than the areas that change control. More middle and working class children were likely to attend grammar school in these steady Labour areas. Hence, there might have been much less demand for comprehensives among Labour voters and politicians in these areas. However, it may also be that these areas differ from other areas in aspects that we cannot control for. We only use Conservative control as a single instrument in column (5) for the narrow treatment group definition. It works similarly, although the coefficient is slightly less than in column (1).

Table 7 shows the instrumental variables results. The difference in math test scores between comprehensive and selective areas was consistently larger according to the IV results than the OLS results, which are repeated in column (1). Moreover, the IV results differ markedly depending on which instrument set is being used. Using Labour control rather than Conservative control yields a more negative estimate, and a coefficient of almost twice the size (in absolute value). Large negative estimates are also obtained when the instrument set includes a variable for Labour control in both 1961 and 1967. It is difficult to believe that these large negative coefficients, larger than the raw difference in test scores, reflect causal effects of comprehensive schooling. If the controls in the OLS regressions are insufficient, it would be curious for the IV estimates to be more negative rather than more positive. Moreover, given that comprehensive LEAs have a worse student population to begin with, it seems natural that the causal estimates should be smaller than the raw differences. This is the case for the OLS estimate, and maybe for the IV estimates in columns (2) and (4). Finally, policy effects in the order of 6 points seem basically too large to be believable. This casts some doubt on the validity of the instruments. Either the instruments are correlated with unobserved area conditions which influence student outcomes, or comprehensive schools set up in Labour controlled LEAs were systematically

performing worse than comprehensive schools in other areas. However, none of the estimates points towards a superior performance of comprehensive areas in terms of test score outcomes.

Table 8 compares the math test score results using the wide treatment group definition with the narrow treatment group definition. Students in comprehensive LEAs score slightly lower according to the narrow definition, but the OLS results using controls are very similar. However, the IV results using the narrow definition are again almost twice as large in absolute value and very similar to the raw difference. The IV regressions only use Conservative control in 1961 as a single instrument. This further highlights that the IV results are rather sensitive to small redefinitions of the sample, possibly casting some doubt on the validity of the instruments.

Table 9 compares our results to the estimation approach used in much of the previous literature, like Kerkhoff et al. (1996) and Galindo-Rueda and Vignoles (2004). Instead of comparing students in selective and comprehensive LEAs, this table compares students in selective and comprehensive schools, using all LEAs. The raw difference in test scores for students in comprehensive schools in column (1) is bigger than in comprehensive LEAs: about 7 points compared to 5.5. After adjusting for covariates in column (2) the difference is smaller, on the other hand, although still significant in our sample. This differs slightly from the results in Kerkhoff et al. (1996), who report no significant differences. The IV estimate in column (3) is much more negative in this case. However, the use of this instrument is more curious in this context because the instrument only varies at the LEA level, while comprehensive versus selective attendance varies within LEAs now. It is therefore not surprising that the IV point estimate is closer to the estimate comparing LEAs while it is also less precise, because the first stage is weaker using all LEAs.

Columns (4) and (5) repeat the OLS estimates using the school attended, grammar versus secondary modern, as the treatment variables, comparable to some of the specifications in previous research. The raw differences in test scores are large. But this is unsurprising since grammar school students are positively selected. Their score advantage drops markedly once controls are introduced in column (5) but is still quite significant at 6 points. This might be unsurprising because it compares grammar school students to all comprehensive school students. In addition, as we argued above, some of the difference may simply reflect the fact that the

conditioning variables are too noisy to effectively control for all differences of grammar school and comprehensive school students.

Table 10 shows results for school leaving exams (O-levels and A-levels) passed and for post 16 educational attainment. Comprehensive school students pass significantly fewer O-level and A-level exams but these differences become small or vanish once controls are introduced and comprehensive status is instrumented. About 15 percent of sample members hold a college degree. The incidence is lower in comprehensive LEAs but the difference is not significant, and again vanishes completely once controls are introduced. Results are slightly different for the number of years of schooling, which include vocational qualifications. On average, students in the sample went to school for about 1.3 years past the compulsory schooling age of 16. In comprehensive LEAs, students obtained about 0.4 years less of schooling. This difference shrinks to slightly less than half after controlling for covariates. The IV estimate, however, is as large as the raw difference. Since we do not find a similar pattern for other education outcomes, we do not want to stress these results too much.

Table 11 shows labor market outcomes. There is little evidence for any employment effects. Among those working, wages are higher for workers who went to school in comprehensive LEAs. This is true both for the raw effect (although the difference is not significant), and after adjusting for covariates and in the IV regressions. In fact, adjusted estimates are bigger than the raw effects, and the IV estimate is bigger still. This is curious, because children growing up in comprehensive LEAs come from worse backgrounds, and we found their comprehensive experience to depress their test scores at age 16. Hence, we would have expected them to do worse in the labor market. Both the OLS and the IV estimates of 5 and 10 percent are also too large to be plausible. However, standard errors are now so large for the results also to be consistent with sizeable negative effects.

The overall picture which emerges from these results is some negative effects on learning in comprehensive schools, as reflected in test scores. On the other hand, we do not find any consistent evidence of effects on school leaving exams and further educational attainment (although the estimates for years of schooling were negative). We also find no evidence on

negative labor market effects, if anything the opposite seems to be true. Hence, it seems that any detrimental impact of comprehensive schooling might have been short lived.

Previous studies have often concluded that there is little overall impact of comprehensive versus selective schooling or tracking, but different types of children do better under one system or the other. Hence, we also investigate these distributional aspects in tables 12 to 14. Table 12 starts by presenting results for the math test scores at age 16. Column (1) shows a specification also run by Kerkhoff et al. (1996), interacting the indicator for a comprehensive LEA with a single indicator for ability at age 11. We have used the average of the four test scores for reading, math, and the verbal and non-verbal scores on the general ability test. Like in Kerkhoff et al. (1996), we find a negative ability interaction, although the result is noisy. The point estimate implies that in a comprehensive area an individual with a test score of 100 would score 4.7 points lower in math at age 16 than an individual with a score of 0. However, few individuals have a score of 0, and nobody has a score of 100. The difference between the 5th and 95th percentile of the score distribution is about 61. This difference would result in a 2.9 lower score at 16. This suggests that the highest ability individuals loose the most from the comprehensive experience. These are, of course, the individuals who are most likely to attend a grammar school in the selective system.

Column (2) suggests that things are not quite that simple. Instead of a linear ability indicator we split ability into three ranges with roughly equal numbers of students (high ability are scores above 60, while medium ability is in the range from 40 to 60). The results suggest that it is high ability and low ability individuals who loose from comprehensive schools the most, although the effect for high ability individuals is slightly bigger. Again, these estimates are not terribly precise.

Not only ability but also family background might impact the schooling experience, particularly in the selective system.[7] We therefore constructed an indicator for low SES, which we set to one for children with fathers in the two lower social classes (social class 4 and 5) as well as those with no father present.[8] Columns (3) and (4) show results introducing this indicator, and

[7] There is actually no clear evidence in our selective sample that low SES children are less likely to attend grammar schools, conditional on their ability, although this is the case using other sample definitions.
[8] Children with no father present look much like the children in the two lower social classes.

interacting it with high ability. The results in column (4) indicate that the main losers in comprehensive schools are high ability-high SES children and low ability-low SES children. Both these groups have math scores of about 3.7 points less than in the selective system. Medium ability-high SES and high ability-low SES children do almost as well in comprehensive schools as in the selective system. However, even with these interactions, no group clearly gains from the comprehensive experience.

Table 13 repeats these results but instruments comprehensive status with Conservative control in 1961. The instruments for the interactions are corresponding interactions of conservative control with the ability indicators. The results tend to be somewhat more pronounced than for the OLS estimates, but the general pattern of results is very much the same.

Table 14 shows results for the ability interactions using the other outcome variables. With a few qualifications, the patterns of results are pretty much the same as that found in table 12 for the math test scores. High ability-high SES individuals tend to fare worse under the comprehensive system, but not high ability-low SES individuals. An important difference emerges for the O-levels and A-level exam results. Only high ability individuals have a significant number of O-level and A-level passes to begin with, and no interesting patterns emerge for the lower ability groups. High ability-high SES individuals pass about 0.7 fewer O-levels in comprehensive areas, and about 0.3 fewer A-levels. High ability-low SES individuals, on the other hand, have 1.4 more O-level passes and 0.3 more A-level passes in comprehensive areas. All these calculations are based on the OLS results. Given that the average number of O-levels and A-levels passed by high ability individuals in our sample are about 5 and 1, respectively, these are quite sizeable effects. This is interesting, because we did not find any superior performance of high ability-low SES individuals in terms of test scores. Hence, the results indicate that these individuals may have had better access to these school leaving exams in comprehensive areas. A similar patter emerges for post-secondary degrees and years of schooling, but these results are rather imprecise.

Conclusion

We have argued in this paper that the experience in England and Wales provides a useful experiment to study the impact of comprehensive versus selective schooling. Compared to previous studies, we pay particular attention to the definition of the treatment and control groups, and we only compare children educated in comprehensive and selective LEAs, and not in different schools within LEAs. Because the NCDS cohort enters secondary school right in the middle of the comprehensive reorganization period most children actually attend schools in LEAs that are neither clearly comprehensive nor clearly selective. Hence, our strategy comes at the cost of relatively small sample sizes, and many of our results are not particularly precise.

Nevertheless, a few results emerge from the analysis. While many studies have concluded that there are no differences between children in comprehensive and selective schools, we actually find children in the selective system to do at least as well or better in terms of test score outcomes. One explanation for this might be that the nature and timing of the comprehensive reorganization in England and Wales meant that children in comprehensive schools faced more turmoil in their school system, since the comprehensive schools they attended were only very recently established. In addition, we documented that the comprehensive schools looked somewhat different in terms of a variety of other characteristics. For example, comprehensive schools were bigger, more often mixed gender, taught slightly fewer hours of math and English in somewhat bigger classes, and used different disciplinary measures than selective schools. Moreover, students liked their schooling experience in comprehensive schools less. The lower performance in comprehensive schools may also be a result of these other characteristics, which tended to be associated with comprehensive status but which might affect education outcomes directly.

Another finding that emerges from our analysis is that certain students tend to benefit more in the selective system, so it is important to study the impacts on different subgroups. Both high and low ability children might do better in the selective system. Hence, it is important to analyze effects which are non-linear in ability. Among high ability children, only high SES children benefit in the selective system, while low SES children might actually be better off in the comprehensive system. This seems to be true particularly with respect to O-level and A-level

results, the British school leaving exams at age 16 and 18. These patterns seem to make a lot of sense if the cause for these results is mixing versus selection. Bright low SES children might be less likely to benefit from grammar schools, and hence be better off in the comprehensive system. It is less clear how these results should be explained by other school characteristics. For example, if class sizes were indeed larger in comprehensive schools, it is difficult to see how this effect would hurt high SES children but not low SES children. In fact, the analysis in Krueger (1999) suggests the exact opposite. Hence, the most obvious explanation for our results is that they are indeed related to the nature of selection.

References

Argys, L.M., D.I. Rees, and D.J. Brewer (1996) "Detracking America's schools: Equity at zero cost?" *Journal of Policy Analysis and Management* 15, 623-645.

Benn, Caroline (1971) *1971 Survey of Comprehensive Reorganization Plans and Lists of Comprehensive Schools in England, Wales, and Scotland*. Comprehensive Schools Committee.

Betts, Julian R. and Jamie L. Shkolnik (2000a) "The effects of ability grouping on student achievement and resource allocation in secondary schools," *Economics of Education Review* 19, 1-15.

Betts, Julian R. and Jamie L. Shkolnik (2000b) "Key difficulties in identifying the effect of ability grouping on student achievement," *Economics of Education Review* 19, 21-26.

Comprehensive Schools Committee (1967) *Comprehensive Education. Secondary Reorganization in England and Wales*. Survey No. 1 1966/7.

Crook, D.R., S. Power, and G. Whitty (1999) *The Grammar School Question*, London: Institute for Education.

Figlio, David N. and Marianne E. Page (2002) "School choice and the distributional effects of ability tracking: Does separation increase inequality?" *Journal of Urban Economics* 51, 497-514.

Galindo-Rueda, Fernando and Anna Vignoles (2004) "The heterogeneous effect of selection in secondary schools: Understanding the changing role of ability," manuscript, Centre for Economic Performance, LSE.

Griffith, A. (1971) *Secondary School Reorganization in England and Wales*, London: Routledge & Kegan Paul.

Jesson, David (2000) "The Comparative Evaluation of GCSE Value-Added Performance by Type of School and LEA," University of York Discussion Papers in Economics No. 2000/52.

Kerckhoff, Alan C., Ken Fogelman, David Crook, and David Reeder (1996) *Going Comprehensive in England and Wales. A Study of Uneven Change,* London: The Woburn Press.

Kim, Taejong, Ju-Ho Lee, and Young Lee (2003) "Mixing versus sorting in schooling: Evidence from the Equalization Policy in South Korea," KDI School Working Paper 03-07.

Krueger, Alan B. (1999) "Experimental Estimates of Education Production Functions," *Quarterly Journal of Economics* 114, 497-532.

Municipal Yearbook (1962, 1968) London: Municipal Journal Ltd.

Rees, D.I., D.J. Brewer, and L.M. Argys (2000) "How should we measure the effect of ability grouping on student performance?" *Economics of Education Review* 19, 17-20.

Figure 1
Shares of Publically Supported Pupils by School Type

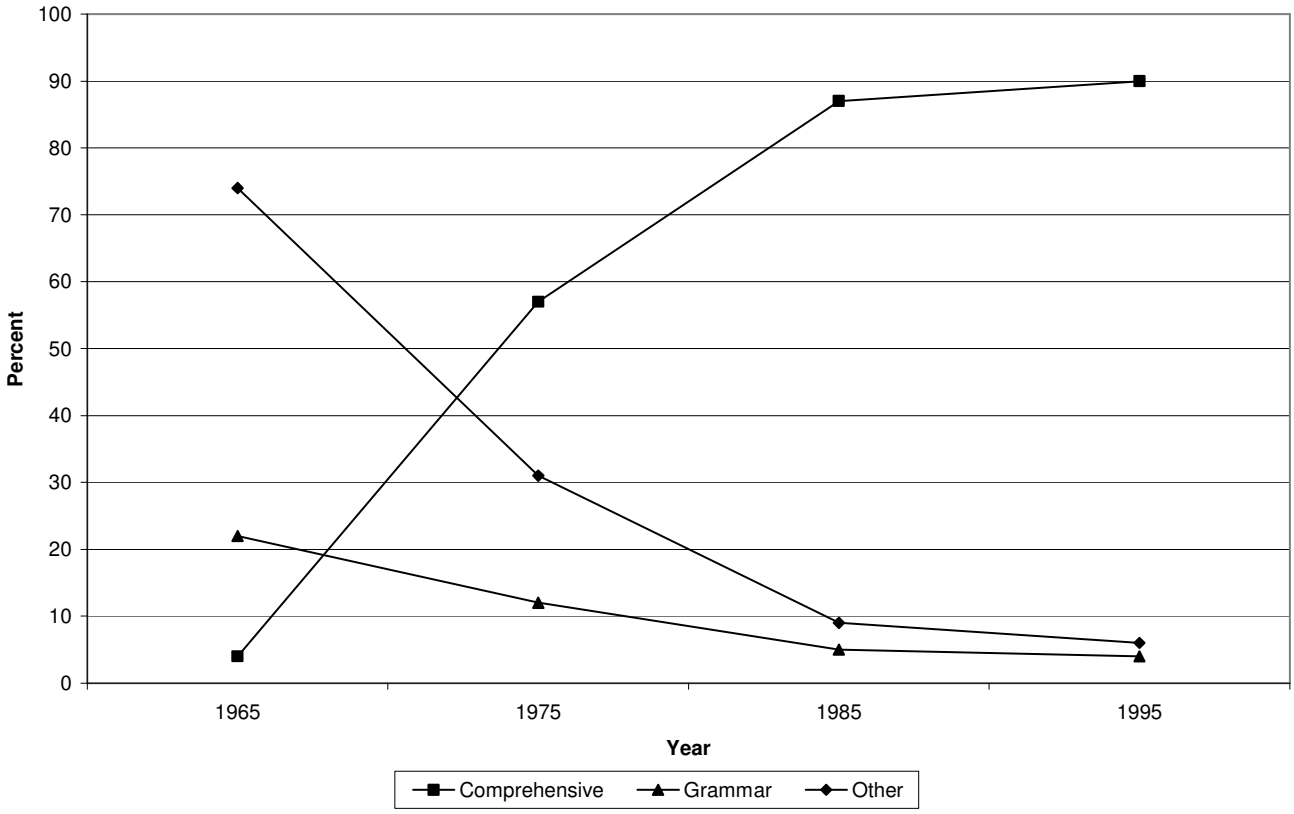

Table 1
Percent of LEA Maintained Students by Type of School Attended

	Definition of Treatment and Control Groups			
	Narrow Definition		Wide Definition	
Year	Selective LEAs	Comprehensive LEAs	Selective LEAs	Comprehensive LEAs
Percent of Students in Comprehensive Schools				
1967	1.3	45.4	1.3	31.9
1969	2.3	91.9	2.3	60.7
1971	4.2	94.9	4.2	95.7
1974	9.5	97.5	9.5	98.5
Percent of Students in Grammar Schools (incl. LEA places in Direct Grant Schools)				
1967	24.4	12.3	24.4	16.0
1969	24.3	3.1	24.3	9.4
1971	24.6	1.7	24.6	1.5
1974	25.5	0.0	25.5	0.0
Percent of Students in Secondary Modern and Other Schools				
1967	74.2	42.2	74.2	52.1
1969	73.4	5.1	73.4	29.9
1971	71.3	3.5	71.3	2.8
1974	65.0	2.5	65.0	1.5

Sources: Education Statistics, various years
Notes: Percentages are weigthed by LEA enrollment

Table 2
Means of Selected Characteristics by Treatment Status, Age 7, Wide Treatment Definition

	Variable	Mean selective LEAs	Difference comprehensive LEAs	t-statistic	Number of observations
*	Female	0.509	-0.052	-2.39	1787
*	2 or more siblings	0.592	0.064	2.52	1741
*	Twin	0.016	0.013	1.45	1787
*	No father figure	0.024	0.015	1.41	1744
	Father social class 1 (highest)	0.056	-0.023	-1.99	1732
	Father social class 2	0.157	-0.054	-2.92	1732
	Father social class 3	0.540	0.02	0.79	1732
	Father social class 4	0.165	-0.002	-0.07	1732
	Father social class 5 (lowest)	0.037	0.038	2.68	1732
	Mother worked before child in school	0.465	-0.013	-0.35	1715
	Housing rented from council	0.358	0.062	1.64	1735
	Accommodation has own bath	0.903	-0.052	-2.29	1737
	Accomm. has own indoor lavatory	0.843	-0.049	-1.75	1730
	Parent's hardly ever read to child	0.107	0.019	1.09	1737
	Child talked at age 2	0.942	-0.011	-0.88	1723
	Age started school	4.609	0.058	1.13	1733
	Attends regular school	0.944	0.037	3.33	1787
	Attends independent school	0.047	-0.034	-2.89	1787
	Attends special needs school	0.009	-0.003	-0.72	1787
	School size	260.5	-3.8	-0.32	1745
	Age started reading in school	5.435	0.000	-0.00	1784
	Age started math in school	5.559	-0.062	-1.07	1785
	30%+ of dads in class in non-man. occ.	0.291	-0.108	-2.97	1787
	Dads in non-manual occ. Missing	0.274	-0.062	-2.10	1787
	Gets extra help in school for problems	0.046	-0.009	-0.99	1773
	Needs extra help or referred to agency	0.118	-0.010	-0.57	1784
	Mother shows no interest in school	0.127	0.060	2.29	1787
	Teacher doesn't know mother's interest	0.076	0.013	0.85	1787
	Father shows no interest in school	0.119	0.067	2.84	1783
	Teacher doesn't know father's interest	0.368	0.028	1.10	1783
	Child settled at school within 1 month	0.695	0.031	1.61	1786
	Class size	35.9	0.6	0.71	1784
	Ability tracking, child in upper ability track	0.051	0.008	0.49	1784
	Ability tracking, child in middle track	0.022	0.004	0.34	1784
	Ability tracking, child in lower ability track	0.026	-0.008	-0.91	1784
	No ability tracking	0.901	-0.005	-0.18	1784
	Teacher ratings: scale 1 – 5 (5 is best)				
*	Teacher rating: oral expression	3.233	-0.189	-3.66	1787
*	Teacher rating: awareness of world	3.004	-0.114	-2.07	1783
*	Teacher rating: reading ability	3.163	-0.125	-2.27	1782
*	Teacher rating: creativity	2.927	-0.095	-2.02	1784
*	Teacher rating: math ability	2.881	-0.057	-1.16	1787
*	Math test score	51.598	-0.54	-0.41	1770
*	Reading test score	80.034	-2.994	-1.85	1778
*	Draw a man score	41.783	-1.807	-2.39	1761

Note: Starred variables are used as control variables in the regressions below.

Table 3
Means of Selected Characteristics by Treatment Status, Age 11, Wide Treatment Definition

	Variable	Mean selective LEAs	Difference comprehensive LEAs	t-statistic	Number of observations
*	No mother figure	0.003	0.003	1.04	1787
*	No father figure	0.036	0.017	1.55	1787
*	Father social class 1 (highest)	0.061	-0.026	-2.05	1766
*	Father social class 2	0.226	-0.090	-4.49	1766
*	Father social class 3	0.479	0.056	2.73	1766
*	Father social class 4	0.151	0.014	0.79	1766
	Father social class 5 (lowest)	0.046	0.028	1.61	1766
*	Father unemployed	0.020	0.017	1.89	1787
*	Mother worked since child was 7	0.628	-0.015	-0.52	1768
*	Housing rented from council	0.363	0.082	1.91	1782
*	Child has own bedroom	0.470	-0.059	-2.08	1765
*	Accommodation has own bath	0.956	-0.033	-2.22	1784
*	Accomm. has own indoor lavatory	0.912	-0.064	-2.60	1777
*	Moved two or more times since birth	0.373	-0.026	-1.30	1774
*	Receives free meal at school	0.082	0.029	1.40	1787
*	Family had fin. difficulties last 12 months	0.110	0.003	0.13	1759
*	Parents expect child to leave at min. age	0.044	0.025	1.59	1783
*	Child goes to public library often	0.255	-0.004	-0.22	1773
*	Contact with criminal justice since age 7	0.019	0.007	0.88	1787
	Did not miss school due to health last year	0.619	0.015	0.66	1779
*	Missed one week to one month last year	0.328	-0.018	-0.88	1779
*	Missed more than one month last year	0.053	0.003	0.22	1779
*	Mother had health problem after age 7	0.037	-0.002	-0.23	1787
	English not usual language at home	0.023	0.012	1.11	1764
	Attends secondary school	0.002	0.009	1.55	1787
*	Attends independent school	0.061	-0.041	-3.58	1787
*	Attends special needs school	0.016	-0.002	-0.42	1787
*	Has attended two or more schools	0.523	-0.086	-1.83	1768
*	School size	336.6	-0.013	-0.00	1728
	Gets extra help in school for problems	0.076	0.015	1.02	1784
	Gets extra help in school for superiority	0.019	-0.007	-1.22	1783
	Needs extra help or referred to agency	0.009	0.010	1.58	1784
	Teacher discussed child at parent initiative	0.609	-0.076	-2.46	1758
	Teacher discussed child at own initiative	0.473	-0.116	-2.93	1718
	Mother shows no interest in school	0.105	0.046	2.29	1749
	Teacher doesn't know mother's interest	0.097	0.042	2.23	1749
	Father shows no interest in school	0.144	0.044	2.04	1682
	Teacher doesn't know father's interest	0.234	0.065	2.51	1682
*	Class size	34.3	0.495	0.70	1738
*	Ability tracking, child in upper ability track	0.170	-0.021	-0.71	1761
*	Ability tracking, child in middle track	0.113	0.024	0.91	1761
*	Ability tracking, child in lower ability track	0.088	0.021	0.93	1761
	No ability tracking	0.629	-0.024	-0.37	1761

Table 3
Means of Selected Characteristics by Treatment Status, Age 11, Wide Treatment Definition (continued)

	Variable	Mean selective LEAs	Difference comprehensive LEAs	t-statistic	Number of observations
*	Teacher ratings: scale 1 – 5 (5 is best)				
*	Teacher rating: general ability	3.027	-0.124	-3.16	1779
*	Teacher rating: math ability	2.923	-0.099	-2.06	1773
*	Teacher rating: use of books	3.157	-0.137	-3.34	1778
*	Teacher rating: oral ability	3.054	-0.078	-1.79	1779
*	Child plans to get job after school	0.193	0.024	1.33	1751
*	Child plans to study after school	0.291	0.006	0.27	1751
*	Child watches TV most of the day	0.836	0.015	0.83	1787
*	Math test score	46.0	-7.8	-4.08	1781
*	Reading test score	47.8	-3.9	-3.24	1782
*	Verbal score on general ability test	59.7	-9.0	-5.17	1782
*	Non-verbal score on general ability test	54.9	-5.6	-4.13	1782
	Sum of math, reading, and general ability	52.1	-6.6	-4.55	1781
*	Design copy score	70.6	-2.1	-3.36	1776
	County level variables (1971)				
*	Proportion households in council housing	0.260	0.066	2.24	1779
*	Proportion working in manufacturing	0.378	0.032	0.98	1779
*	Proportion single parent families	0.090	0.009	0.97	1779
*	Proportion econ. active men in population	0.609	-0.001	-0.20	1779
*	Proportion men born in the UK	0.946	0.007	0.64	1779
*	Wales	0.000	0.226	2.34	1788
*	County borough	0.521	0.195	1.05	1787

Note: Starred variables are used as control variables in the regressions below.

Table 3
Means of Selected Characteristics by Treatment Status, Age 11, Wide Treatment Definition

Variable	Mean selective LEAs	Difference comprehensive LEAs	t-statistic	Number of observations
Attends comprehensive school	0.118	0.796	26.99	1787
Attends selective school	0.779	-0.729	-24.11	1787
Attends grammar school	0.255	-0.230	-9.71	1787
Attends secondary modern school	0.514	-0.489	-11.98	1787
Attends independent school	0.052	-0.030	-2.74	1787
Attends technical school	0.010	-0.010	-2.03	1787
Attends other school	0.046	-0.033	-1.58	1787
Years comprehensive (if in comprehensive)	1.73	1.98	6.45	847
Comprehensive in 69 (if in comprehensive)	0.23	0.38	3.78	848
Students class tracked at age 12	0.928	-0.325	-8.94	1547
Tracking in English at age 16	0.971	-0.116	-5.46	1772
Tracking in math at age 16	0.993	-0.074	-6.19	1780
Number of schools attended since age 11	1.167	0.199	2.75	1515
School is mixed gender	0.609	0.282	4.78	1782
School size	740.6	338.1	6.45	1768
Pupil teacher ratio	17.2	-0.2	-0.69	1650
Class size in English	25.5	0.5	1.44	1708
Class size in Math	24.9	0.6	1.72	1710
30%+ of dads in class in non-manual occupations	0.591	-0.146	-2.17	1524
Fraction in school studying for O-levels	0.358	0.069	2.97	1210
Fraction in school passing 2+ A-levels	0.061	0.064	2.74	1048
Fraction in school going on to degree prog.	0.034	0.042	2.94	1018
Hours of English	3.6	-0.2	-1.97	1719
Hours of Math	3.4	-0.1	-1.65	1717
Percent expelled for disciplinary reasons	0.037	-0.015	-2.02	1698
School uses severe disciplinary measures (suspension, corporal punishment)	0.314	0.130	2.36	1754
School uses extra work, detention, etc.	0.882	-0.016	-0.64	1759
School uses reports to parents, etc.	0.926	0.040	2.71	1767
Fraction of school days missed	0.096	0.044	4.39	1665
Child has been in trouble with police	0.066	0.061	3.81	1600
Parents have disc.child with school staff	0.575	-0.047	-1.11	1656
Discussion at initiative of parents	0.121	0.003	0.16	1637
Discussion at initiative of school	0.494	-0.048	-0.92	1637
Parents satisfied with school	0.676	-0.032	-0.84	1508
Student's attitudes towards school:				
School is largely a waste of time	0.091	0.039	2.38	1675
Homework is a bore	0.513	0.094	4.21	1666
Difficult to keep mind on work	0.351	0.019	0.84	1667
In never take work seriously	0.136	0.031	1.68	1662
I don't like school	0.284	0.071	2.85	1666
No point in planning for the future	0.149	0.050	2.87	1669
Student works while in school	0.522	-0.063	-1.76	1684
Weekly hours worked	4.0	-0.4	-1.06	1677
Earnings per week (£)	1.33	-0.10	-0.82	1673

Table 5
Test Scores at 16, Wide Treatment Definition
OLS Regressions

	Dependent Variable			
	Math Test Score		Reading Test Score	
Regressor	(1)	(2)	(3)	(4)
Comprehensive LEA	-5.58 (1.42)	-2.05 (1.09)	-3.26 (1.39)	0.05 (0.61)
Control variables	no	yes	no	yes

Note: Number of observations is 1281. Control variables are the starred variables in Tables 2 and 3.

Table 6
First Stage Regressions

	Dependent Variable: Comprehensive LEA according to				
	Wide Treatment Definition				Narrow
Regressor	(1)	(2)	(3)	(4)	(5)
Conservative control 1961	-0.67 (0.13)	---	-0.56 (0.17)	-0.69 (0.16)	-0.48 (0.16)
Labour control 1961	---	0.47 (0.17)	0.18 (0.18)	0.43 (0.17)	---
Conservative control 1967	---	---	-0.06 (0.15)	-0.02 (0.13)	---
Labour control 1967	---	---	0.09 (0.20)	0.73 (0.37)	---
Labour control both 1961 and 1967	---	---	---	-0.88 (0.37	---
P-value for excluded instruments	0.000	0.006	0.000	0.000	0.006
Observations	1281	1281	1281	1281	1058

Note: All regression include the starred variables in Tables 2 and 3 as control variables.

Table 7
Math Test Score at 16, Wide Treatment Definition
OLS versus IV Regressions

	OLS	Instrumental Variables			
Regressor	(1)	(2)	(3)	(4)	(5)
Comprehensive LEA	-2.05 (1.09)	-3.77 (2.15)	-6.17 (3.86)	-4.78 (2.25)	-6.13 (1.79)
Instruments:					
Conservative control 1961		✓		✓	✓
Labour control 1961			✓	✓	✓
Conservative control 1967				✓	✓
Labour control 1967				✓	✓
Labour control both 1961 and 1967					✓

Note: Number of observations is 1281. All regression include the starred variables in Tables 2 and 3 as control variables.

Table 8
Math Test Score at 16, Alternative Treatment Definitions

	OLS		IV
Regressor	(1)	(2)	(3)
Narrow Treatment Definition			
Comprehensive LEA	-6.28 (1.63)	-2.47 (1.41)	-6.36 (3.53
Observations	1058	1058	1058
Wide Treatment Definition			
Comprehensive LEA	-5.58 (1.42)	-2.05 (1.09)	-3.77 (2.15)
Observations	1281	1281	1281
Control Variables	no	yes	yes

Note: Control variables are the starred variables in Tables 2 and 3. Instrument for comprehensive LEAs is Conservative control in 1961.

Table 9
Math Test Score at 16
Using School Attended as Treatment Variable

	OLS		IV	OLS	
Regressor	(1)	(2)	(3)	(4)	(5)
Student attends comprehensive school	-7.18 (0.74)	-1.22 (0.43)	-4.15 (3.59)	---	---
Student attends grammar school	---	---	---	27.21 (0.85)	6.05 (0.61)
Student attends secondary modern school	---	---	---	-5.97 (0.77)	-1.46 (0.53)
Control Variables	no	yes	yes	no	yes

Note: Number of observations is 5161. Control variables are the starred variables in Tables 2 and 3. Instrument for comprehensive LEAs is Conservative control in 1961.

Table 10
Exams and Highest Education Level, Wide Treatment Definition

	OLS		IV
Regressor	(1)	(2)	(3)
Dependent Variable: Number of O-Levels taken anywhere			
Comprehensive LEA	-0.65 (0.20)	-0.16 (0.14)	-0.09 (0.28)
Observations	1154	1154	1154
Dependent Variable: Number of A-Levels taken anywhere			
Comprehensive LEA	-0.16 (0.05)	-0.06 (0.06)	0.03 (0.10)
Observations	1154	1154	1154
Dependent Variable: Post-secondary degree			
Comprehensive LEA	-0.04 (0.02)	-0.01 (0.02)	0.01 (0.04)
Observations	952	952	952
Dependent Variable: Years of Post-16 Schooling (truncated at 7)			
Comprehensive LEA	-0.39 (0.17)	-0.16 (0.12)	-0.39 (0.23)
Observations	900	900	900
Control Variables	no	yes	yes

Note: Control variables are the starred variables in Tables 2 and 3. Instrument for comprehensive LEAs is Conservative control in 1961.

Table 11
Labor Market Outcomes, Wide Treatment Definition

	OLS		IV
Regressor	(1)	(2)	(3)
Dependent Variable: Employed full time at age 33			
Comprehensive LEA	0.037 (0.036)	-0.017 (0.031)	0.032 (0.048)
Observations	961	961	961
Dependent Variable: Log wage at age 33			
Comprehensive LEA	0.028 (0.037)	0.048 (0.049)	0.105 (0.078)
Observations	733	733	733
Control Variables	no	yes	yes

Note: Control variables are the starred variables in Tables 2 and 3. Instrument for comprehensive LEAs is Conservative control in 1961.

Table 12
Math Test Score at 16, Interactions of Treatment Status
Wide Treatment Definition

Regressor	(1)	(2)	(3)	(4)
Comprehensive LEA	0.22 (2.56)	-2.41 (1.53)	-2.14 (1.45)	-1.76 (1.59)
Comrehensive LEA*ability	-0.05 (0.05)	---	---	---
Comprehensive LEA*high ability	---	-0.89 (2.46)	-1.03 (2.40)	-1.96 (2.80)
Comprehensive LEA*medium ability	---	1.56 (1.90)	1.49 (1.85)	1.38 (1.88)
Comprehensive LEA*low SES	---	---	-0.77 (1.43)	-2.01 (1.69)
Comprehensive LEA*high ability*low SES	---	---	---	4.95 (4.21)

Note: Number of observations is 1281. Control variables are the starred variables in Tables 2 and 3. Ability is the average of reading, math, verbal, and non-verbal score. High ability is ability of 60 or greater. Medium ability is ability of 40 or greater but less than 60. Low SES is father in social class 4 or 5 or father's social class missing. Control variables are the starred variables in Tables 2 and 3. Column (2) also includes a high and medium ability main effect. Column (3) includes a low SES main effect in addition to the controls in column (2). Column (4) includes a high ability*low SES interaction in addition to the controls in column (3).

Table 13
Math Test Score at 16, Interactions of Treatment Status
IV, Wide Treatment Definition

Regressor	(1)	(2)	(3)	(4)
Comprehensive LEA	1.01 (4.04)	-3.19 (2.43)	-4.40 (2.67)	-4.05 (2.82)
Comrehensive LEA*ability	-0.10 (0.08)	---	---	---
Comprehensive LEA*high ability		-3.68 (3.56)	-3.23 (3.40)	-5.04 (4.05)
Comprehensive LEA*medium ability		1.68 (2.52)	1.92 (2.47)	1.60 (2.54)
Comprehensive LEA*low SES			2.28 (2.30)	-0.13 (2.66)
Comprehensive LEA*high ability*low SES	---	---	---	9.35 (5.84)

Note: Number of observations is 1281. Control variables are the starred variables in Tables 2 and 3. Ability is the average of reading, math, verbal, and non-verbal score. High ability is ability of 60 or greater. Medium ability is ability of 40 or greater but less than 60. Low SES is father in social class 4 or 5 or father's social class missing. Control variables are the starred variables in Tables 2 and 3. Column (2) also includes a high and medium ability main effect. Column (3) includes a low SES main effect in addition to the controls in column (2). Column (4) includes a high ability*low SES interaction in addition to the controls in column (3). Instrument for comprehensive LEAs is Conservative control in 1961. Instruments for the various interactions are interactions of Conservative control in 1961 with the respective variable.

Table 14
Exams, Education and Labor Market Outcomes, Interactions of Treatment Status
Wide Treatment Definition

	Dependent Variable											
	O-levels		A-levels		Postsecondary degree		Post 16 years of school (up to 7)		Employed full time at 33		Log Wage at 33	
	OLS	IV	OLS	IV	OLS	IV	OLS	IV	OLS	IV	OLS	IV
Regressor	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
Comprehensive LEA	-0.16	0.08	-0.06	0.09	0.01	0.05	-0.12	-0.00	-0.016	-0.031	0.143	0.131
	(0.17)	(0.33)	(0.06)	(0.11)	(0.03)	(0.05)	(0.17)	(0.28)	(0.054)	(0.090)	(0.090)	(0.127)
Comprehensive LEA*high ability	-0.55	-1.02	-0.28	-0.27	-0.11	-0.14	-0.36	-1.13	-0.015	0.106	-0.198	-0.083
	(0.31)	(0.47)	(0.11)	(0.17)	(0.05)	(0.08)	(0.40)	(0.62)	(0.087)	(0.148)	(0.110)	(0.149)
Comprehensive LEA*medium ability	0.15	-0.11	0.13	0.05	0.00	0.00	0.20	0.07	0.016	0.037	-0.040	0.060
	(0.24)	(0.34)	(0.06)	(0.10)	(0.04)	(0.06)	(0.23)	(0.35)	(0.066)	(0.098)	(0.067)	(0.084)
Comprehensive LEA*low SES	-0.04	0.11	-0.01	-0.10	0.01	-0.09	-0.23	-0.76	0.010	0.062	-0.085	-0.106
	(0.23)	(0.35)	(0.06)	(0.09)	(0.04)	(0.06)	(0.29)	(0.44)	(0.072)	(0.105)	(0.109)	(0.155)
Comprehensive LEA*high ability*low SES	2.11	2.17	0.62	0.47	0.16	0.36	0.67	1.42	0.055	0.074	0.165	0.232
	(0.94)	(1.28)	(0.30)	(0.39)	(0.15)	(0.21)	(0.80)	(1.17)	(0.143)	(0.201)	(0.139)	(0.196)
Observations	1154	1154	1154	1154	952	952	900	900	961	961	733	733

Note: Ability is the average of reading, math, verbal, and non-verbal score. High ability is ability of 60 or greater. Medium ability is ability of 40 or greater but less than 60. Low SES is father in social class 4 or 5 or father's social class missing. Control variables are the starred variables in Tables 2 and 3 plus main effects for high and medium ability, low SES, and an interaction of high ability*low SES. Instrument for comprehensive LEAs is Conservative control in 1961. Instruments for the various interactions are interactions of Conservative control in 1961 with the respective variable.