DISCUSSION PAPER SERIES

No. 7977

THE EFFECT OF TRACKING STUDENTS BY ABILITY INTO DIFFERENT SCHOOLS: A NATURAL EXPERIMENT

Nina Guyon, Eric Maurin and Sandra McNally

LABOUR ECONOMICS

Centre for Economic Policy Research

www.cepr.org

THE EFFECT OF TRACKING STUDENTS BY ABILITY INTO DIFFERENT SCHOOLS: A NATURAL EXPERIMENT

Nina Guyon, Paris School of Economics
Eric Maurin, CEPR and Paris School of Economics
Sandra McNally, Centre for Economic Performance (CEP), LSE

ABSTRACT

The Effect of Tracking Students by Ability into Different Schools: a Natural Experiment*

The tracking of pupils by ability into elite and non-elite schools represents a common, but highly controversial policy in many countries. In particular, there is no consensus on how large the elite track should be and, consequently, little agreement on the potential effects of any further increase in its size. This paper presents a natural experiment where the increase in the relative size of the elite track was followed by a very significant improvement in average educational outcomes. The experiment under consideration provides a rare opportunity to isolate the overall contextual effect of allowing entry to the elite track for a group that was previously only at the margin of being admitted.

Nina Guyon
Paris School of Economics
48 Boulevard Jourdan
Paris 75014
FRANCE

Email: nina.guyon@gmail.com

For further Discussion Papers by this author see:
www.cepr.org/pubs/new-dps/dplist.asp?authorid=168060

Eric Maurin
Paris School of Economics
Bureau 105, Bat B
48 Bd Jourdan
75014 Paris
FRANCE

Email: eric.maurin@ens.fr

For further Discussion Papers by this author see:
www.cepr.org/pubs/new-dps/dplist.asp?authorid=131502

Sandra McNally
Deputy Director
Centre for Economic Performance
London School of Economics
Houghton Street
London WC2A 2AE

Email: s.mcnally1@lse.ac.uk

I   Introduction

The tracking of students by ability into different school types is a widespread, but highly controversial policy, with some countries starting to track as early as age 10 (Germany, Austria) whereas other countries start tracking much later, after the years of compulsory schooling (US, UK, France). The selection of a fraction of high ability students into a subset of elite schools modifies the peer groups and school context for all students. The net impact of such a strategy is extremely difficult to identify, as is the net effect of any education expansion policy relying on increased access to the more elite track. An opposing view is that increases in the size of the elite sector dilutes the value of education received by high ability students, while at the same time negatively affecting the school context of the low and middle ability students who remain in the non-elite sector. It might be argued that such negative contextual effects offset the potentially positive effect of the reform on the group of students who are allowed entry to elite schools and who were only at the margin of being admitted before the expansion policy. In fact, it is even debated whether these marginal students actually benefit from the reform and whether, beyond a certain point, education expansion initiatives generate any positive effect at all. Even in countries where there is no tracking at school-level, this becomes an issue when considering how many people should attend university (at public expense).

It is very difficult to shed light on these issues. One basic problem is that more selective areas (or countries) differ in many respects to those which are less selective. Hence, a comparison of average outcomes in more or less selective education systems does not provide a credible strategy for evaluating the true effect of educational tracking. Indeed, there is little convincing evidence about how variation in the relative size of the elite and non-elite tracks affects average educational outcomes (see for example Manning and Pischke, 2006,

Figlio and Page, 2002, Betts and Shkolnik, 1999). This is the substantive question that we address in this paper.

We make use of a unique natural experiment where the distribution of students by ability across secondary schools was modified within Northern Ireland at a particular point in time (1989). The secondary school system in Northern Ireland involves the distribution of students across a small set of elite schools and a much larger set of non-elite schools, where elite schools select about a third of students who obtain the best results at a national ability test taken at the end of primary school (at age 11). In 1989, elite schools were required to accept pupils up to a new (larger) admission number determined only by 'physical capacity', where 'physical capacity' was defined on a school-by-school basis by the Northern Ireland Education Department.

This reform led to a significant increase in the overall proportion of pupils in the elite track ('grammar schools') at the beginning of their secondary school education. Furthermore, the impact was very significant in some areas of Northern Ireland, but almost negligible in other areas (plausibly those where elite schools were considered already near 'full capacity' before the reform). This natural experiment allows identification of the effect of an increase in the share of pupils selected into elite schools on average educational attainment, by comparing average outcomes just before and after the reform as well as the distribution of average outcomes across local areas just before and after the reform. The attractiveness of this experiment is that the de-tracking reform is the only change that occurred during the period of interest. Most educational expansion reforms have several very different components whose effects cannot be separately identified. To the best of our knowledge, the reform in Northern Ireland is the first where it is possible to isolate the net effect of an increase in the relative size of the elite track.

We use administrative data covering the entire relevant population to examine the impact of the reform on entry flows to elite schools and the outcomes of affected cohorts. There is a clear discontinuity in the overall inflow to elite schools just after the reform – the number of students entering elite schools increased by about 15% between the 1978 and 1979 birth cohorts whereas it was reasonably stable for the three preceding and three subsequent cohorts. This discontinuity is reflected in outcome measures. For example, the number of students obtaining 3 or more A-levels at age 18 (i.e. a typical entry qualification for university) increased by about 10% over the same period whereas it followed the same stable trend as the number attending grammar school in the three preceding and subsequent cohorts. The increase is also reflected in the national examination taken by all pupils at age 16 (prior to the end of compulsory schooling).[2] The reform has been accompanied by a clear discontinuous improvement in average educational outcomes which provides the first piece of evidence for a positive effect of increasing the proportion of pupils in the elite track. We show that this is also reflected in university entry rates.

As expected, our administrative data also reveal significant heterogeneity in the effect of the reform within Northern Ireland across local areas. In some areas, the reform was followed by a very significant shift in the proportion of pupils selected into elite schools. In other areas, the reform produced only very small changes. We find that the reform produced shifts in educational achievement at age 16 or 18 which are much more significant in areas where the initial shift in elite school attendance was stronger..

Thus, the reform makes it possible to provide Instrumental Variable estimates of the effect of school segregation by ability using several different sources of identification. One can make use of the discontinuity across birth cohorts in the average proportion of pupils entering into elite schools. One can also rely on variation in the difference in the proportion of

---

[2] GCSE examinations (General Certificate of Secondary Education) are taken by all students at the end of compulsory education.

pupils attending elite schools across strongly and weakly affected local areas. Both strategies give estimates of the effect of expanding the elite track which are significant and similar, despite relying on very different identifying assumptions.

The net effect of the reform on average educational outcomes can be interpreted as the combination of three basic factors: the effect of attending an elite school on the group of students who would otherwise have entered a non-elite school; the effect of losing more able peers on the group of students entering non-elite schools after the reform; the effect of having less able peers on the group of students who would have entered the elite school even in the absence of the reform. Separately identifying these effects would amount to identifying the effect of changes in school type (or school context) for different ability groups, which is notoriously difficult. As shown in the last part of the paper, it is nonetheless possible to provide lower bound estimates of these effects by analysing the effect of the reform separately on elite and non-elite school outcomes. Interestingly, we find that the reform had a negative effect on average performance in non-elite schools, but not in elite schools, in spite of a decline in the average ability of their students. Hence, elite students do not seem to suffer from attending more heterogeneous schools with additional, relatively less able, peers. Also, students at the margin of being selected to elite schools seem to perform as well as top ability students when they are actually selected into these schools and benefit from a 'high ability' school context. Thus, increasing the share of the elite sector seems to generate positive externalities for mid-ability students, but no negative externalities for top ability students. This is a plausible reason for why this policy has such a strongly positive net effect on average outcomes.

The remainder of the paper is structured as follows. In Section II we briefly discuss the relevant literature. In Section III, we describe the institutional context and the reform. In Section IV, we present our administrative data as well the construction of the panel of local

areas in Northern Ireland that is used in the econometric analysis. In Section V we provide several sets of estimates of the elasticity of the number of students passing national examinations at age 16 or 18 to the proportion selected into elite schools at age 11. Section VI provides a discussion of our basic results, building on a separate analysis of the effect of the reform on elite and non elite schools. Section VII concludes.

II   Literature

There are several recent strands of the UK and international literature on school segregation by ability which are of relevance to our study. Using a panel of about 20 countries, Hanushek and Wößmann (2006) identify the effect of tracked secondary school systems by comparing performance differences between primary and secondary schools across tracked and non-tracked systems, where each country's own primary school outcome is included as a control. They find that tracked systems tend to increase educational inequality and to reduce average performance to some extent, although this effect is only marginally significant.   These findings have been challenged by Waldinger (2006) who finds that results are not stable to using different tracking measures and to restricting the sample to OECD countries.

In a UK context, several studies have compared the outcomes of students living in areas where students are tracked by ability into different schools to those where there is no tracking. Within Great Britain, regional variation in the exposure to a tracked system existed at a time when the system was being transformed (in the 1960s and 1970s) because the abolition of the tracked system in Great Britain only occurred gradually (whereas it did not happen in Northern Ireland). Galindo-Rueda and Vignoles (2004) and Kerkhoff et al. (1996) use variation within Great Britain to estimate the effect of exposure to a tracked system on educational outcomes (regardless of the school type actually attended by an individual).

Atkinson et al. (2004) use more recent administrative data to perform a similar analysis in a contemporary setting (the 'selective school' system was retained in a small number of areas in Great Britain). Manning and Pischke (2006) use the same data as that used by Galindo-Rueda and Vignoles (2004) and Kerkhoff et al. (1996), but show that the abolition of the grammar school system was not random across areas. They find that strategies relying on local variation in the degree of selectivity of the school system produce the same results regardless of whether the dependent variable is after the 'treatment' (i.e. age 16 test scores) or before the 'treatment' (age 11 scores). They conclude that caution is required in drawing strong conclusions from studies that rely on the timing chosen by local areas to abolish the tracked system.

Our paper is also related to the literature that investigates the effect of within school ability segregation (see, for example, Betts and Shkolnik, 1999, Figlio and Page, 2002, Duflo, Dupas and Kremer, 2008). Using a randomized evaluation applied to primary schools in Kenya, Duflo et al. (2008) find that schools with (maximum) segregation in two equal-sized ability groupings do better than schools with no segregation at all. Also they find that segregation was beneficial to students at all points in the ability distribution. Segregation within primary schools in a developing country is of course not equivalent to segregation across secondary schools in a developed country. For example, the potential negative effect of being assigned to a non-elite group is likely to depend a lot on the age of the students and on the importance placed on educational success in society. Also, it should be emphasised that education expansion reforms (such as that in Northern Ireland) typically involve an increase in the homogeneity of peers for low ability pupils, but a decrease in homogeneity for high ability pupils. It is unlikely to be possible to infer the effects of such policies from experiments where all pupils are affected by the same increase in the extent of homogeneity within the school (in terms of pupil ability).

Finally, our research is also related to the literature[3] on the impact of the educational expansion reforms that took place in Europe after World War II since de-tracking was often part of these reforms. However the reforms had typically several very different components, including increases in school leaving age. Hence, outcomes cannot be attributed to the specific effect of de-tracking. A distinguishing feature of our study is that the natural experiment under consideration has not modified the nature of the school system but only modified the relative size of the elite sector. To identify the effect of widening access to the academic track on average outcomes, we rely on comparisons between children who go to school in the same educational system, where marginal reforms are made to that system rather than involving conversion to a different type of system. To the best of our knowledge, this experiment is the first to isolate the overall contextual effect of allowing entry to the elite track for a group that was previously only at the margin of being admitted.

III   Institutions and reform

In a number of key respects, the education system is the same in Northern Ireland as that in England and Wales. Pupils spent six years in primary school, from age 5 to age 11, and then five additional years in secondary school, until age 16, the minimum school-leaving age. At the end of compulsory education (age 16), all students take GCSE examinations. It is usual for students to take 8 to 10 subjects, including English and Math. There is an externally set and marked exam for each subject (pass grades are A^*, A, B, C....G and then a fail). Anything from grade A^* to grade C is regarded as 'good' and the standard outcome measure for a student is whether he/she achieves 5 or more grades at A^*-C[4]. The National Qualifications

---

[3] See e.g. Meghir and Palme (2005) for the Sweden , Pekkarinen, Uusitalo and Pekkala (2009), for the Finland, Aavik, Salvanes and Vaage (forthcoming) for Norway or Gurgand and Maurin (2006) for France.
[4] Students might not be allowed to continue in a subject to A-level if they had not managed to get a C in it for GCSE.

Framework (NQF) used by UK employers consider grades *D-G* as a level 1 qualification; grades *A*-C* as level 2 (*A*-level being at level 3). The proportion of students achieving 5 or more grades at *A*-C* is also the key national indicator to measure performance at the end of compulsory schooling (and applies to England, Wales and Northern Ireland). In the UK, many studies find that qualifications which mark the end of compulsory education have a very large impact on labour market outcomes. In terms of data and methodology, one of the most convincing studies is by Blundell et al. (2005) who found a wage return of 18% for those entering the labour market with these qualifications versus stopping at age 16 without qualifications (see also McIntosh, 2006).

If the student decides to pursue academic education beyond GCSE, this involves studying for A-level exams which normally requires an extra two years of study. These examinations are externally set and graded and are the usual entry route to university. Compared to leaving school without qualifications, Blundell et al. (2005) finds an average wage return of 24% for those completing A-levels only, which rises to 48% for those completing higher education.

The education system in England, Wales and Northern Ireland is also similar in that they operate under a similar legislative framework and have a similar National Curriculum[5]. However, in Northern Ireland, there is still a selective system of secondary education whereas England and Wales largely converted to the comprehensive model in the 1960s and 1970s.[6] This change almost happened in Northern Ireland as well but plans were halted following the election of the Conservative government in 1979.

---

[5] Important Acts are the 1944 Education Act for England and Wales and the 1947 Act for Northern Ireland; the 1988 Education Reform Act in England and Wales and the Education Reform (Northern Ireland) Order 1989.
[6] Other important differences are religious segregation in the education system of Northern Ireland: most Catholics attend schools under Catholic management ('maintained') whereas most Protestants attend other state schools. Also, there are many more single sex schools in Northern Ireland – 25% compared to 16% in England. Of single sex schools, about 45% are grammar schools (i.e. those that select the more academically able).

A   Tracking of students by ability in Northern Ireland

Unlike the comprehensive system (where schools are not allowed to select on the basis of academic ability), the selective system in Northern Ireland involves a test at age 11 which determines the type of secondary school a child will attend: grammar schools (for the more academically able) or other secondary schools. Between 1981 and 1994 (i.e. cohorts born in 1970 and 1983), the transfer test was based on two tests of the verbal reasoning type with some questions designed to test specific aspects of English and mathematics (Sutherland, 1993).[7] Within this framework, the key difference between grammar and other secondary schools is in their pupil composition in terms of ability – along with the consequences this has for the teaching environment and the ethos of the school. Gallagher and Smith (2000) suggest that the 'grammar school effect' is explained by a combination of the clear academic mission of schools, high expectations for academic success on the part of teachers and the learning environment created by a pupil peer group which is selected on academic grounds. All of these factors combine to make the education experience very different in grammar schools than in other secondary schools, even though they operate under the same National Curriculum and implement the same public examinations. In contrast, there is no suggestion in the literature that this effect could be explained by differences in funding between sectors[8].

All schools are expected to apply the same National Curriculum which prescribes, in detail, the range of subjects which must be taught at all levels of compulsory education; the relative time allocation to different areas of the curriculum; and the actual course content for the various subjects (see Morgan, 1993). While grammar schools and other secondary schools operate under this same framework, in practice, there is some evidence of heterogeneity in the curricula actually implemented by schools, with pupils in a sample of grammar schools

---

[7] In 1993/94, the transfer tests were changed from a verbal reasoning to a curriculum orientated format. This affects cohorts born from 1983 onwards.

[8] Funding to schools in both sectors is based on formula funding and is largely determined by pupil numbers.

spending more time at academic subjects (particularly languages) than their counterparts in a sample of other secondary schools (Harland *et al.*, 2002).

The same public examinations are taken in both school types (GCSE at age 16, A-levels at age 18). In all grammar schools and in many other secondary schools, it is possible to stay on for 2 extra years.[9] Although school type is highly correlated with the probability of obtaining A-levels (reflecting the selection process as well as any genuine 'school' effect), there is no automatic relationship between entering grammar school and achieving A-levels or entering other secondary school and failing to achieve them.

B   The 1989 Reform

As explained above, it was a political accident that Northern Ireland did not abolish 'selective schooling' at the same time as the rest of the UK in the 1960s and 1970s. As a consequence, the system of very early tracking (i.e. at age 11) has been maintained in Northern Ireland up to the present day, whereas in other respects the education system has remained similar to that in other parts of the UK. However, an important reform to grammar school admission was implemented in Northern Ireland in the late 1980s. This involved a rise in the level of quotas applied to grammar school intakes. Following the Education Reform (Northern Ireland) Order 1989 (implemented from 1990 and affecting cohorts born from 1979), grammar schools were required to accept pupils, on parental request, up to a new (larger) admission number determined by the Department of Education and based only on the physical capacity of the school. This 'open enrolment' reform was in the spirit of making the education system more amenable to parental choice. Between 1985 and 1989 (before the reform), about 8,100 pupils (31% of the cohort) entered grammar schools each year, whereas this increased to about 9,400

---

[9] It is also possible to study for A-levels in colleges of further education. However, the majority of students in Northern Ireland who obtain A-levels do so when at school.

pupils (35% of a cohort) just after the reform, between 1989 and 1992 (i.e., between cohorts born from 1979 to 1982, see Figures 1 and 2).

The reform generated a 15% increase in the number of students attending grammar school, for a time period in which cohort size was relatively stable. This corresponds to an 11% increase in the probability of attending grammar school between the 1978 and 1979 cohorts, whereas this probability was fairly stable immediately before the policy (1976-78) and immediately afterwards (1979-81).

The raising of quotas on grammar school intakes was controversial because of the fear that grammar schools would 'cream-skim' the highest ability students from other secondary schools and that all would suffer as a result. A concern voiced by the Northern Ireland Economic Council (1995) was that the reform could undermine the selective system: 'The educational impact of allowing the grammar school sector to expand needs to be questioned. The fundamental point of such a system is that educating the more academically able is seen as being of benefit to both the more and least able. By definition, it would seem that allowing students who previously would have entered a secondary environment to attend a grammar school must inevitably dilute the perceived value of selective education...' Our evidence allows us to consider what reducing selectivity did to educational credentials in the overall population.

IV   Data and variables

We use two administrative data sets that were obtained from the Department of Education in Northern Ireland. The first one provides annual school-level information on the number of pupils entering each grade. The second data set provides school-level data on all school

leavers by grade and year.[10] Also, this data set contains information on national examination outcomes and all qualifications attained. Both data sets contain information on the name, religious affiliation (Catholic or Protestant), location and type of school (grammar or non-grammar). Note that these datasets cover the entire population of secondary schools, except independent schools. In Northern Ireland only a small percentage of pupils attend independent schools (less than 1 per cent) and this has not changed over the time period of interest to us. We use these administrative datasets to build a panel of 23 local areas with information on the proportion of pupils attending grammar schools and average examination outcomes for each local area and each cohort born between 1974 and 1982.[11] We created these local areas on a geographic and religious basis: first we divide the set of all schools in Northern Ireland by religious denomination[12] (for the most part, Catholics attend either Catholic grammar or non-grammar schools; Protestants attend Protestant grammar or non-grammar schools). Second, we match each non-grammar school to the grammar schools of its local administrative district (LAD). Education at a local level in Northern Ireland is administered by five "Education and Library Boards" (ELB) covering different geographical zones (Belfast, North Eastern, South Eastern, Southern, Western) and these ELB are divided in 26 LAD. Whenever a LAD does not contain any grammar school of a given religious denomination we match the corresponding non-grammar schools of this LAD to an adjacent LAD[13]. Finally, we merge some additional adjacent LADs in order to eliminate small areas with erratic size. Overall, we obtain a total of 23 areas (10 Catholic and 13 Protestant) such that the proportion of pupils found in each area is very stable across cohorts. There is, for example, no significant **difference in the average number of pupils in each area before and after the reform,**

---

[10] This is called the School Leavers Survey and is actually a census of all school leavers. It contains details of all their qualifications.

[11] Since grade repetition is not a feature of the school system in the UK, it is possible to derive birth cohort using available information on grade and date (i.e., cohort = date - grade).

[12] There are 113 Catholic schools (31 grammar and 82 non grammar) and 143 Protestant schools (40 grammar and 103 non grammar) in Northern-Ireland.

[13] We observe 4 LAD without any Protestant grammar school, 11 LAD without any Catholic grammar school.

which is consistent with the assumption that the reform has mostly affected the allocation of students across schools within areas and not across areas.

With respect to religion and size, our procedure yields one large Protestant area (with 11 grammar schools) and one large Catholic area (with 7 grammar schools) in the Belfast region, plus 12 smaller Protestant and 9 smaller Catholic areas outside Belfast (with on average 2.6 grammar schools in each of these smaller areas). Each large Belfast area represents about 12% of the population of pupils whereas each smaller area represent on average 3.6% of the population (see descriptive statistics in Appendix A).

Within this framework, our basic research question is whether the reform to grammar school admission had any influence on the number of students achieving 5 or more GCSEs at grades A*-C at age 16 or achieving A-levels at age 18. Data are available for cohorts born between 1974 and 1982, for which there were no major reforms to A-levels, or to the age 16 examinations or to the transfer tests determining entry to grammar school. As it happens, reforms to the A-level system have taken place in 1987/88 (affecting cohorts from 1972 onwards) and in 2000 (affecting cohorts from 1984 onwards) whereas reforms to the examination taken at age 16 by all pupils (GCSE – formerly O-levels) took place in 1988 (affecting cohorts from 1972 onwards), but no reforms took place for cohorts born between 1972 and 1988[14]. To illustrate this, Figure 3 shows the change in our measures of educational success in England[15] for the cohorts born before and after the reform under consideration (i.e., before and after 1978). We do not find any significant shift at the time of the reform. We observe the same smooth increase in the proportion of successful students across cohorts born before and after the reform (about a 1 percentage point increase per year). Given that the

---

[14] As discussed above, reform to the transfer test affected cohort born form 1983 onward. The Universities and Colleges Admission Service (UCAS) provide a detailed account of these reforms and what the examinations consist of.

[15] Pre-reform information is not available for exactly the same cohorts in England and Northern Ireland. With regard to GCSEs in England, we have used school-level information from the School Performance Tables that is available from 1992 onwards i.e., cohorts from 1976 onwards. With regard to A levels, we have used pupil level information, which gives comprehensive coverage of the results of all students taking A-levels in England and which is available from 1993 onwards (enabling us to consider outcomes from the 1975 cohort).

examination system at age 16 and 18 is exactly the same in England and Northern Ireland, this figure provides further support to the assumption that examination procedures and the overall ability to pass examinations did not undergo any discontinuous change in Northern Ireland at the time of the reform. In the next section, we build on this assumption to provide several estimates of the effect of early de-tracking on subsequent average educational outcomes.

IV   Educational Effects of the Reform

In this Section, we estimate the educational effects of the reform using different identifying assumptions. We use a simple model where the number of students who pass their exams at the end of secondary education (i.e. at age 16 or 18) in area i and cohort c depends on (a) the total number of students who enter secondary education in area i and cohort c and (b) the distribution of students across elite and non-elite schools in area i and cohort c. Specifically, we assume the following model of education production:

(1) $Y_{i,c} = \alpha + \beta G_{i,c} + \gamma S_{i,c} + \theta_0(c) + u_i + \varepsilon_{i,c}$

where $Y_{i,c}$ represents the number of students who pass their exams at age 16 (or 18) in area i and cohort c, $S_{i,c}$ the total number of pupils who enter into secondary education in area i and cohort c and $G_{i,c}$ the proportion of pupils selected into elite schools at age 11 in area i and cohort c. Variables $Y_{i,c}$, $G_{i,c}$ and $S_{i,c}$ are specified in log format so that parameter β can be directly interpreted as the educational effect of a 1% increase in admission numbers in elite schools, holding cohort size constant. In Appendix B, we report the full set of regression results using alternative specifications and we obtain very similar results (see Table B1). For example, the results are qualitatively unchanged when $Y_{i,c}$ and $S_{i,c}$ only are in log format. Also, the conclusions remain the same when we use the proportion of successful students as

the dependant variable (rather than the log number) as well as when we constraint parameter γ to $\gamma=1$. We prefer our more flexible specification since it is not obvious *ex ante* that cohort size, as such, has no effect on the quality of education and probability of success in an area.

Variable $\theta_0(c)$ captures any continuous cohort trends that may affect the proportion of successful students either before or after the reform: we use a spline function with a knot at the reform date[16]. The variable u_i represents fixed effects which capture permanent differences in outcomes across areas. Finally $\varepsilon_{i,c}$ represents cohort-specific shocks to pupils' ability to pass examinations at age 16 (or 18) in area i. Within this framework, the parameter of interest is β which captures the effect of school segregation by ability on educational outcomes. The basic identification issue comes from the fact that cohort-specific shocks to student ability $\varepsilon_{i,c}$ may be correlated with the cohort-specific shocks to the proportion of students selected into elite schools[17]. In such a case, the OLS regression of $Y_{i,c}$ on $G_{i,c}$ provides a biased estimate of β, even after de-trending and purging out fixed effects. To address this issue, we first make use of the discontinuous shift affecting the average level of elite school attendance as a consequence of the reform.

A   Change in Average Elite School Attendance After the Reform

Assuming that there is no discontinuity in average ability to pass exams $E(\varepsilon_{i,c} \mid c)$ at the time of the reform in Northern Ireland, parameter β is identified as the ratio of the shift in the proportion of successful students and the shift in the proportion of pupils in elite schools observed just after the reform. It can be estimated in Model (1) using a 'reform on' dummy $1(c>c_0)$ as an instrumental variable (where c_0. is the last unaffected cohort).

---

[16] $\theta_0(c)$ is written $\theta_{01} c+\theta_{02}(c-c_0)1(c>c_0$) where parameter θ_{01} captures pre-reform cohort trend whereas parameter θ_{02} represents the change in cohort trend after the last unaffected cohort c_0 .
[17] Suppose for example that the proportion of students selected into grammar school in area i tends to be larger for cohorts who happen to have a larger proportion of very good students in area i (in an absolute sense). In such a case, $Y_{i,c}$ and $G_{i,c}$ will be correlated even if there is no causal effect of $G_{i,c}$ on $Y_{i,c}$.

Before moving on to the estimation results, it is of interest to consider Figures 4 to 6, which uses the area-level data to show variation across cohorts in the average proportion of grammar school students and average number of successful students at age 16 (or age 18). Interestingly, they reveal a significant discontinuity in both variables at the reform date[18], which is consistent with the hypothesis that variation in the proportion of students selected into elite schools at age 11 affected the number of successful students at age 16. For example, the reform generated an increase of about 14% in our measure of success at age 18, whereas it was only weakly increasing in the pre-reform period and it is stable in the period immediately post policy.

Table 1 provides the result of the corresponding regression analysis. Column 1 shows the results of the first-stage regression,

$$(2) \qquad G_{i,c} = \delta + \pi 1(c > c_0) + \gamma_1 S_{i,c} + \theta_1(c) + v_i + \upsilon_{i,c}.$$

where $1(c>c_0)$ is a dummy indicating that the reform is on whereas $\theta_1(c)$ is a spline function with a knot at c_0 (i.e., $\theta_1(c)=\theta_{11}c+\theta_{12}(c-c_0)1(c>c_0)$). It confirms a significant discontinuous increase in $G_{i,c}$ at the date of the reform. The estimate of π is positive and significant at standard levels. Columns 2 and 3 show the results of reduced form regressions. These confirm that there was a shift in the number of successful students (either at GCSE or A-level) at the date of the reform which is parallel to that observed for the proportion of students selected into elite schools for the relevant cohort. Columns 4 and 5 show results for the corresponding second stage regressions, which suggest that a 10% increase in the proportion of students selected into elite schools generates a 4.1% increase in the number of successful students at age 16 and a 7.5% increase at age 18. These estimates are actually quite close to the basic OLS estimates[19] (see Columns 6 and 7). In Table B2 of Appendix B, we replicate the same

---

[18] Note that these shifts cannot be interpreted as reflecting changes happening at one point in time (changes in evaluation practises for instance) since they correspond to the same cohort shift observed at different ages.

[19] One plausible reason for the similarity of OLS and IV estimates is that they use the same basic source of identification. As it happens, putting aside the year of the reform, the probability of selection into grammar

analysis using the difference in average outcomes between areas in Northern Ireland and England as the dependant variable (i.e., using $Y_{i,c}$-$Y_{0,c}$ rather than $Y_{i,c}$ as the dependant variable, where $Y_{0,c}$ represents English outcomes) and adding English cohort size $S_{0,c}$ in the set of control variables. It amounts using England as a control group and purging the common factors that may have affected educational outcomes in both countries during the period under consideration[20]. This approach yields estimated impacts that are very similar to those in Table 1 for age 18 outcomes and even larger for age 16 outcomes (although the difference is not significant at standard levels). The same results hold true regardless of whether we control for pre- and post-reform trends in the difference in outcomes between Northern Ireland and England.

Before moving on to the next identifying strategy, let us emphasize that our regression tables provide robust estimates of standard errors.[21] In particular, these estimates account for potential within-cohort correlation of residuals (which can be a major source of imprecision for estimates of changes affecting all areas simultaneously). In fact, we have checked that estimates of standard errors are very similar when using standard uncorrected estimates or robust estimates. In our specific case, the within-cohort correlation of residuals is weak and does not lead to significant bias in uncorrected estimates of standard errors.

B   Change in the Distribution of Elite School Attendance Across Areas

---

schools is the ratio between a quasi constant number of places and a more fluctuating cohort size (see Figure 2). Thus, the unobserved shocks to the selection probability in fact coincide with shocks to cohort size. Given that these shocks are likely to be absorbed in our regressions by the control variable S_i, the only remaining source of identification in an OLS regression of Y_{ic} on G_{ic} is the shift in G_{ic} at the time of the reform, i.e., exactly the same source of identification as the IV.

[20]When using this estimation strategy, we no longer have to assume that the unobserved determinants of average educational outcomes did not change discontinuously at the time of the reform (as in the previous model). We only exclude that such shifts (if any) affected Northern Ireland and England in a different way.

[21] Note that one key difference between having disaggregated panel data on 22 x 9 area-level observations and having only 9 national-level observations lies precisely in the possibility of estimating the variance-covariance of residuals $v_{i,c}$ and the precision of estimates of national-level shifts.

The previous analysis provides an estimate of the effect of the reform under the assumption that other national-level determinants of educational outcomes did not undergo a discontinuous shift in Northern Ireland at the time of the reform.[22] In this sub-section, we provide an evaluation relying on a completely different assumption, using a feature of the reform that we have not yet exploited. Specifically, we make use of the fact that the reform did not have the same impact on the proportion of pupils in elite schools in different areas of Northern Ireland. As discussed above, the effect of the reform in a given area was determined only by local capacity constraints. Hence, the effect of the reform on grammar school entry was determined in each local area by parameters that had plausibly nothing to do with the variation in pupils' ability to pass exams across cohorts. In such a case, the educational effect of increasing the proportion of pupils entering elite schools in an area can be identified by evaluating whether the most affected areas are also those which experienced the largest improvement in educational outcomes after the reform. Specifically, under the maintained assumption that the area-specific changes in $v_{i,c}$ between post-reform and pre-reform cohorts are uncorrelated with the area-specific changes in $\varepsilon_{i,c}$ across the same periods (i.e., $E(v_{i,c} \mid i,c>c_0)$- $E(v_{i,c} \mid i,c\leq c_0)$ uncorrelated with $E(\varepsilon_{i,c} \mid i,c>c_0)$- $E(\varepsilon_{i,c} \mid i,c\leq c_0)$), we can evaluate parameter β by estimating Model (1) after taking long-differences between post-reform and pre-reform period,

(3) $Y_{i,after} - Y_{i,before} = \delta + \beta (G_{i,after} - G_{i,before}) + \gamma (S_{i,after} - S_{i,before}) + (\varepsilon_{i,after} - \varepsilon_{i,before})$

where, for each variable x_{ic}, $x_{i,after}$ represents the mean of x_{ic} in area i across post-reform cohorts and $x_{i,before}$ represents the mean of x_{ic} in area i across pre-reform cohorts. Note that this second strategy provides an estimate of β even in the case where there is a nation-level

---

[22] If $E(\varepsilon_{i,c} \mid i, c>c_0)$-$E(\varepsilon_{i,c} \mid i,c\leq c_0)$) denotes the difference between mean unobserved ability in area i across post-reform cohorts and mean unobserved ability in area i across pre-reform cohorts, the identifying assumption used in the previous sub-section is that the mean of $E(\varepsilon_{i,c} \mid i, c>c_0)$-$E(\varepsilon_{i,c} \mid i,c\leq c_0)$) across areas is zero.

discontinuity in pupils' average ability at the time of the reform,[23] i.e. even when our first identification strategy provides a biased estimate of β. Also this second strategy does not necessarily coincide with the fixed effect OLS estimate of model (1) since it relies on the sole change observed at the time of the reform whereas the fixed-effect OLS evaluation uses all observed fluctuations for identification. Table 2 shows the result of estimating Model (3). Panel A uses the full set of available cohorts (i.e. 1974-1982) and provides estimates using the difference in mean educational outcomes between the four post-reform cohorts and the five pre-reform ones as the dependant variable. By contrast, Panel B focuses only on the two pre-reform and two post-reform cohorts (i.e., 1977-1980) and provides estimates using the difference in mean educational outcomes between the two post-reform and the two pre-reform cohortsas the dependant variable. The results are very similar across the two specifications. This analysis suggests that a 10% increase in the proportion of grammar school entrants generates an increase of about 4% in the number of students obtaining 5 or more GCSEs at grades A*-C and an increase of about 7% in the number of students with 3 A-levels or more at age 18. Most interestingly, this estimated elasticity is very close to the estimates obtained in the previous sub-section even though the source of identification is completely different. The first strategy used the nation-level discontinuity in the relationship between entry to elite schools and cohort of birth whereas the second strategy uses the differential impact across areas as a source of identification. Figures 7 and 8 show graphically that there is a very clear correlation between area-level variation in the proportion of successful students at age 16 (5 or more GCSEs at grades A*-C) or at age 18 (3 A-levels or more) and area-level variation in the proportion of students selected into grammar schools.

---

[23] Formally, denoting $\Delta v_i = (v_{i,after} - v_{i,before})$ and $\Delta \varepsilon_i = (\varepsilon_{i,after} - \varepsilon_{i,before})$ we can have $E(\Delta v_i \Delta \varepsilon_i) = 0$ even when $E(\Delta \varepsilon_i) \neq 0$ (and conversely, $E(\Delta \varepsilon_i) = 0$ even when $E(\Delta v_i \Delta \varepsilon_i) \neq 0$). As it happens, the two strategies rely on two different sources of identification: the change in nation-level elite school attendance (first strategy) vs. the change in the distribution of elite school attendance across areas (second strategy).

In substance, the identifying assumption used in this sub-section is that the change in students' average ability after the reform is not particularly strong (nor weak) in areas where the reform implied a strong increase in grammar school capacity. One potential issue is that some families may have moved into these areas after the reform in order to benefit from the increase in enrolment to elite schools. Consequently, the number and average ability of pupils may have changed at the same time as the enrolment capacity of elite schools in these areas, which could create a bias in the OLS estimates of Model (3). In such a case, however, we should observe a positive correlation between the change in the size of the elite sector in an area and the change in the total number of students in this area after the reform. As shown by the last column of Table 2, this is not the case: there is no positive association between the change in the size of elite schools and the change in the total number of students after the reform. When we focus on the two pre-reform and two post-reform cohorts, this also confirms that the reform has not been followed by any significant reallocation of students from weakly affected to strongly affected areas.

C   Discontinuity in the Difference in Elite School Attendance Across Areas

To further explore the robustness of our results, we have divided the set of areas into two groups according to the magnitude of the impact of the reform ("strongly" versus "weakly" impacted) and we have analysed whether a change in the difference in educational achievement between these two groups occurred precisely at the time of the reform. To conduct this analysis, we have assumed an extended version of model (1),

$$(4) \qquad Y_{i,c} = \alpha + \beta G_{i,c} + \gamma S_{i,c} + \theta_2(c) \mathrm{x} T_i + \tau_c + u_i + \varepsilon_{i,c}$$

where τ_c represents cohort fixed effects, $\theta_2(c)$ a spline function with a knot at c_0 and T_i a dummy indicating that area i is a strongly-treated one. The interaction $\theta(c) \mathrm{x} T_i$ captures

potentially diverging cohort trends between $T=1$ and $T=0$ group before and after the reform. In this model, the identifying assumption is that the difference in pupil ability between $T=1$ and $T=0$ areas does not undergo a shift at the reform date c_0,

$$(5) \qquad E(\varepsilon_{i,c} \mid c > c_0, T_i=1) - E(\varepsilon_{i,c} \mid c > c_0, T_i=0) = E(\varepsilon_{i,c} \mid c \leq c_0, T_i=1) - E(\varepsilon_{i,c} \mid c \leq c_0, T_i=0).$$

In such a case, parameter β is identified even when there is a discontinuity in average ability at c_0 or when there are diverging trends across areas in elite school attendance and student achievement[24]. Specifically, β is identified as the ratio between the shift in the difference in student achievement at the cut-off date and the shift in the difference in grammar school attendance at the same cut-off date.

Given the institutional set-up, the simplest way to define our control group ($T=0$) is to focus on areas with the smallest variation in grammar school attendance at the time of the reform. The grammar schools in these areas were plausibly near full capacity at the time of the reform and have been impacted only marginally. Table 3 shows the regression results when $T=0$ corresponds to areas below the first quartile of the distribution of changes in grammar school entry at the time of the reform[25]. Specifically, column (1) of Table 3 shows the corresponding first-stage regression,

$$(6) \qquad G_{i,c} = \delta + \pi 1(c > c_0) \times T_i + \gamma S_{i,c} + \theta_3(c) \times T_i + \tau_c + u_i + \varepsilon_{i,c}$$

Unsurprisingly, the estimate of π is significantly positive and suggests that the reform was followed by a 15% increase in the relative proportion of students attending grammar schools in $T=1$ areas. Also the regression confirms that there is no significant difference in pre-reform or in post-reform cohort trends across areas, so that the shift in relative attendance occurs precisely at the date of the reform. Columns (2) and (3) show the reduced form

---

[24] By focusing on the discontinuity in the difference in attendance, this strategy provides an evaluation of β even when there are common trends in grammar school attendance and student achievement that existed prior to the reform, i.e. even when $\Delta\varepsilon_i = \varepsilon_{i,after} - \varepsilon_{i,before}$ and $\Delta v_i = v_{i,after} - v_{i,before}$ are affected by common factors and OLS estimation of Model (3) is biased.

[25] We have 5 areas in the weakly treated group (with changes in grammar school entry below 10%). We have 17 areas in the strongly treated group (with changes above 10%). We have checked that our results are robust to changes in the specification. For example, they are almost unchanged when we define "weakly treated" as the group of areas with changes in grammar school entry below 15%.

regressions which reveal that this shift was accompanied by an increase of about 11% in the relative number of successful students in 'strongly affected' areas. Columns (4) and (5) show the corresponding IV estimates. The estimated effect on the number of successful students at age 18 is similar to previous estimates, even though it is less precisely estimated (significant at the 7% level only). The estimated effect at age 16 is even larger than previous estimates, although the difference between this estimate and previous estimates is not statistically significant. Finally, the last column of the Table shows that there is no discontinuity in the relative size of areas at the time of the reform. The reform has not generated a significant reallocation of students and families across areas.

Another way to define our "control" and "treated" groups is to rely on pre-reform characteristics only, i.e., religious affiliation and pre-reform grammar school attendance. It turns out that the effect of the reform is significantly stronger in areas where the pre-reform entry rates were the weakest, especially in Protestant areas. The next question is whether we observe an improvement in educational outcomes in low-entry Protestant areas compared to the other areas at the time of the reform. We have checked that this is actually the case and that re-estimation of model (4) with this alternative definition of T generates similar estimates for β (analysis available on request).

D   Effect on Entry into Higher Education

The School Leavers Survey (SLS) consists of a questionnaire sent to all secondary schools where they are asked to provide information on the secondary qualifications obtained by school leavers (GCSEs, A levels) and also on the post-secondary destination of these students (higher education, employment, unemployment, training, unknown). By construction, the information on destinations is more speculative and less precise than the information on

qualifications obtained before leaving school[26]. As mentioned in the guidance notes of the SLS, schools often have difficulties in coding the destinations of students who change residence or students who start to work during the summer after leaving school, but who may nonetheless enter into university at the beginning of the next academic year. With all these data limitations in mind, for each area and each cohort, we have constructed a measure of the number of students who have attended higher education after secondary school[27] and we have analysed this destination outcome using exactly the same methods as those used previously to analyse secondary qualifications. As shown in Table 4, all three strategies suggest a positive effect of the reform on university attendance, even though the effect is less well estimated than the effect on qualifications. For example, model (3) shows that the increase in university attendance is stronger in areas where the increase in grammar school attendance is greater, suggesting that a 10% difference across areas in the increase in grammar school attendance between periods 1977-1978 and 1979-1980 generates a 5% difference across areas in the increase in university attendance between the same period (.53 elasticity significant at the 1% level).

VI.   Interpretation and Discussion.

The interpretation of the overall improvement in exam performance in Northern Ireland is that it is the combination of three basic effects: the effect of attending grammar school on pupils who would otherwise have attended another secondary school; the effect of losing more able peers on students still entering non-grammar schools after the reform; the effect of having less

---

[26] The guidance notes ask schools to use the "unknown" code in not more than 5% of the cases. Thus it is not possible to have an idea of the true "unknown" rate (i.e., the one that would emerge without coding constraints).
[27] Note that, in contrast, the SLS data cannot be used to construct for each cohort a measure of unemployment at entry into the labour market or unemployment at a given age.

able peers on students who would have entered a grammar school even in the absence of the reform. It is not possible to point identify the specific contribution of each of these effects. Nonetheless, it is possible to provide plausible lower bounds by examining the impact of the reform separately for elite and non-elite schools.

A   Bounds to Contextual Effects

To be specific, the reform defines three different ability groups *(g=A, B* and *C)*. Firstly, there is a group of relatively high ability pupils (*g=A*) who would have entered grammar school even in the absence of the reform. The impact of the reform on this group amounts to the effect of having a group of peers with relatively low average ability compared to what would have been the case in the absence of the reform. Secondly, there is a group of mid-ability pupils (*g=B*) who attend grammar school after the reform, but who would have attended another secondary school had the reform not taken place. The effect of the reform on these pupils is potentially very important since such pupils are exposed to a radically different school context than what they would have faced in the absence of the reform. Finally, there is a group of relatively low ability pupils (*g=C*) who attend other secondary schools both before and after the reform. They are affected by the change in the composition of these schools. Specifically, they have lost their best peers (group *B*) because of the reform.

Using these notations, elite schools include group A only before the reform, but are composed of groups A + B after the reform. In such a case, the variation in elite schools' average outcomes after the reform reflects (1) the fact that the average ability of pupils has declined in these school (because of the inclusion of group B) (2) the fact that the performance of group A may itself have been affected by this new group of peers. In other

words, the change in elite schools' average outcomes is a mix between a potentially negative composition effect and more ambiguous peer effects on high ability pupils.

Hence, the impact of the reform on the average outcomes in elite schools does not point identify peers' effect on group *A*, but provides a lower bound for this contextual effect. A more formal presentation of this argument is given in Appendix C.

Similarly, the variation in average outcomes in non-elite schools after the reform is a mix between (1) the effect of the change in peers' composition on the group *C* of low-ability pupils and (2) the potentially negative composition effect due to the loss of group *B*, i.e., a group of pupils with higher ability than group *C*. Hence, the effect of the reform on the average outcomes in non-elite schools does not point identify peers' effect on group *C*, but provides a plausible lower bound for this effect. In the next sub-section, we provide a separate empirical evaluation of the effects of the reform on grammar and non-grammar schools which we interpret as lower bounds for the contextual effects that have affected top and bottom ability students after the reform.

B Separate Effects for Grammar and Non Grammar Schools

Table 5 shows regressions of the number of successful students in each school type (and of the total number of students in each school type) on the same set of explanatory variables as those used in Table 1: cohort size ($S_{i,c}$) and a spline function of cohort with a knot at the reform date. Column (1) confirms that the size of elite schools increased by about 12% just after cohort 1978. This timing corresponds to the inflow of relatively low ability students generated by the reform in these schools. Interestingly, Column (2) reveals that the reform was followed by an even larger shift (+14%) in the number of successful students at age 16 in these schools. Overall, success has increased at about the same rate as entry to grammar

schools which is consistent with the assumption that new students in elite schools have not generated negative externalities (in spite of their relatively low ability) and have in fact strongly benefited from their new high-ability peers. We are in a situation where the reform generates unambiguously non-negative contextual effects in elite schools.

The picture is somewhat different in non-elite schools. Column (3) confirms that they underwent a significant negative shift in size just after the reform (-4%). But column (4) reveals that it was accompanied by an even more negative shift in the number of successful students at age 16 (-11%), although the difference between the two estimates is not statistically different. Overall, success seems to have declined more rapidly than attendance in non-elite schools. Hence, we are a situation where the sign of the contextual effect of the reform on low ability students is ambiguous. The decline in average outcomes observed in non-grammar schools may simply reflect the decline in the average level of ability of students after the reform in these schools. However, it may also partly reflect the fact that students in these schools have lost their best peers after the reform.

VII. Conclusion

The tracking of students by ability into different schools is a common phenomenon in developed countries. Also, reforms increasing the size of the more selective tracks have occurred in many countries over recent decades. The effects of such 'de-tracking' policies are difficult to identify because they often happen at the same time as other educational reforms. Thus, there is little reliable evidence with which to debate the consequences of such controversial reforms. In this context, the reform examined in this paper is particularly interesting: there was a large increase in the number of pupils admitted to the elite track whereas, in other respects, the educational system remained unchanged. Analysing the

discontinuity in the distribution of educational outcomes across cohorts and local areas, we show that the net effect of the 'de-tracking' reform was a very significant increase in examination results at the end of compulsory schooling (i.e. GCSEs, age 16) and 'high school' (i.e. A-levels, age 18). According to our basic estimates, a 10% increase in the proportion of students selected in elite school at age 11 in an area is followed by an increase of about 4% in the number of students who pass national examinations at age 16 and an increase of about 7% in the number of students who pass national examinations at age 18. These effects encompass not only the direct effect of attending grammar school for the marginal entrants, but also the indirect effect arising from the change in school context in both elite and non-elite schools. Overall, this paper provides an unambiguous piece of evidence that widening access to the more academic track can generate very positive net effects.

References

Aavik, Arild, Kjell Salvanes and Kjell Vaage, "Measuring Heterogeneity in the Returns to Education in Norway Using An Educational Reform", *European Economic Review*, forthcoming.

Atkinson, A., P.Gregg, and B.McConnell, The Result of 11 Plus Selection: An Investigation into Equity and Efficiency of Outcomes for Pupils in Selective LEAs, Mimeo, Leverhulme Centre for Market and Public Organisation, University of Bristol.

Betts, J.R., and J.L. Shkolnik, (1999), The Effects of Ability Grouping on Student Math Achievement and Resource Allocation in Secondary Schools. *Economics of Education Review*, 19: 1-15

Blundell, R., L. Dearden and B. Sianesi, (2005), Evaluating the Effect of Education on Earnings: Models, Methods and Results from the National Child Development Survey. *Journal of the Royal Statistical Society*. A. 168(3): 473-512.

Duflo, Esther, Dupas Pascaline and Michael Kremer, (2008), Peer Effects, Teachers Incentives and the Impact of Tracking: Evidence from a Randomized Evaluation in Kenya, NBER Working Paper No. 14475, *American Economic Review*, forthcoming.

Figlio, D.N., and M.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, F., and A.Vignoles, (2004), The Heterogeneous Effect of Selection in Secondary Schools: Understanding the Changing Role of Ability, Institute for the Study of Labor (IZA) Discussion Paper 1245. Bonn, Germany.

Gallagher, T., and A. Smith, (2000), The Effects of the Selective System of Secondary Education in Northern Ireland, *Report to the Department of Education*, Northern Ireland.

Gurgand, Marc and Eric Maurin (2006), "Démocratisation de l'enseignement secondaire et inégalités salariales en France », *Annales, Histoire, Sciences Sociales*, 4, 845-859.

Hanushek, E.A. and L. Wößmann (2006), Does Educational Tracking Affect Performance and Inequality? Differences-in-Differences Evidence Across Countries, *Economic Journal* 116, C363-C76.

Harland, J., H. Moor, K. Kinder and M. Ashworth (2002), Is the Curriculum Working? The Key Stage 3 Phase of the Northern Ireland Curriculum Cohort Study, National Foundation for Educational Research, Slough.

Kerckhoff, A.C., K.Fogelman, D.Crook and D.Reeder, (1996), Going Comprehensive in England and Wales. The Woburn Press.

Manning A. and S. Pischke, (2006), 'Comprehensive versus Selective Schooling in England and Wales: What Do We Know?' *IZA DP* No. 2072

Machin, S., and S. McNally, (2005), 'Gender and Student Achievement in English Schools,' *Oxford Review of Economic Policy*, Vol. 21, No. 3, pp. 357-372.

McIntosh, S., (2006), Further Analysis of the Returns to Academic and Vocational Qualifications. *Oxford Bulletin of Economics and Statistics*. Vol 68: 225-251.

Meghir, C., and M. Palme, (2005), Educational Reform, Ability and Family Backround, *American Economic Review*, Vol. 95, No. 1, pp. 414-424.

Morgan, V., (1993), 'Gender and the Common Curriculum, 'in R. Osborne, R. Cormack and A. Gallagher (eds)., After the Reforms: Education and Policy in Northern Ireland. Avebury.

Northern Ireland Economic Council (1995), Reforming the Educational System in Northern Ireland, *Occasional Paper 1*, January.

Pekkarinen, Tuomas, Roope Uusitalo and Sari Pekkala, (2009), "Education Policy and Intergenerational Income Mobility : Evidence from the Finnish Comprehensive School Reform", *Journal of Public Economics*, 93, (965-973).

Sutherland, A., (1993), 'The Transfer Procedure Reformed?' in R. Osborne, R. Cormack and A. Gallagher (eds)., After the Reforms: Education and Policy in Northern Ireland. Avebury.

Waldinger, Fabian (2006), 'Does Tracking Affect the Importance of Family Background on Students' Test Score,' *Unpublished Manuscript*, London School of Economics.

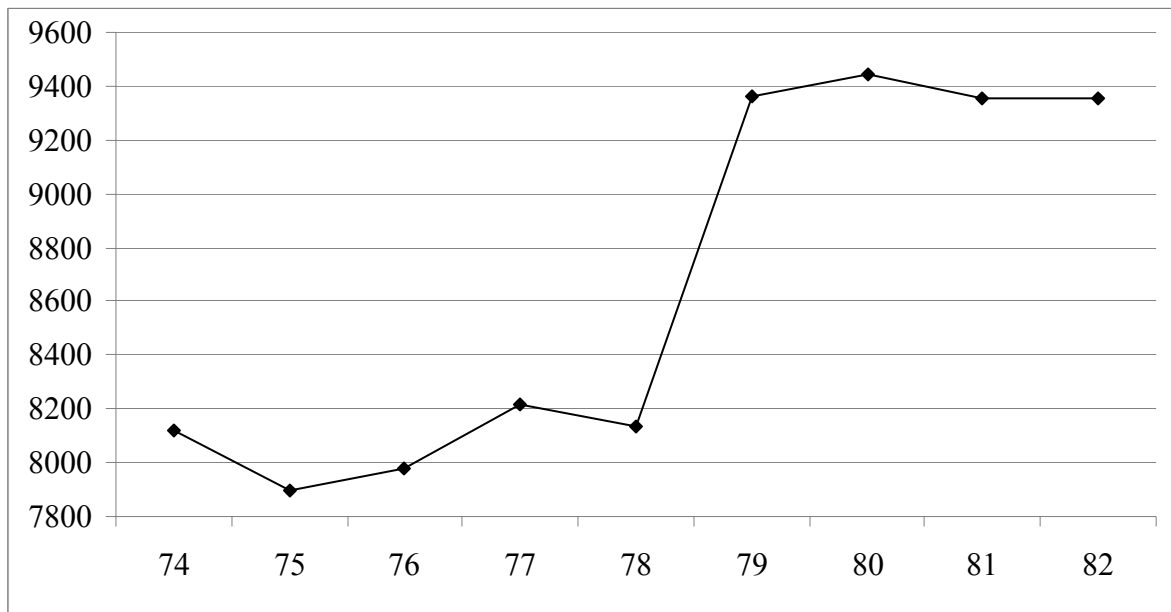Figure 1: Number of Entrants to Grammar School in Northern-Ireland, by Year of Birth.

Figure 2: Evolution of Cohort Size and Number of Entrants to Grammar School, by Year of Birth (1974=1).
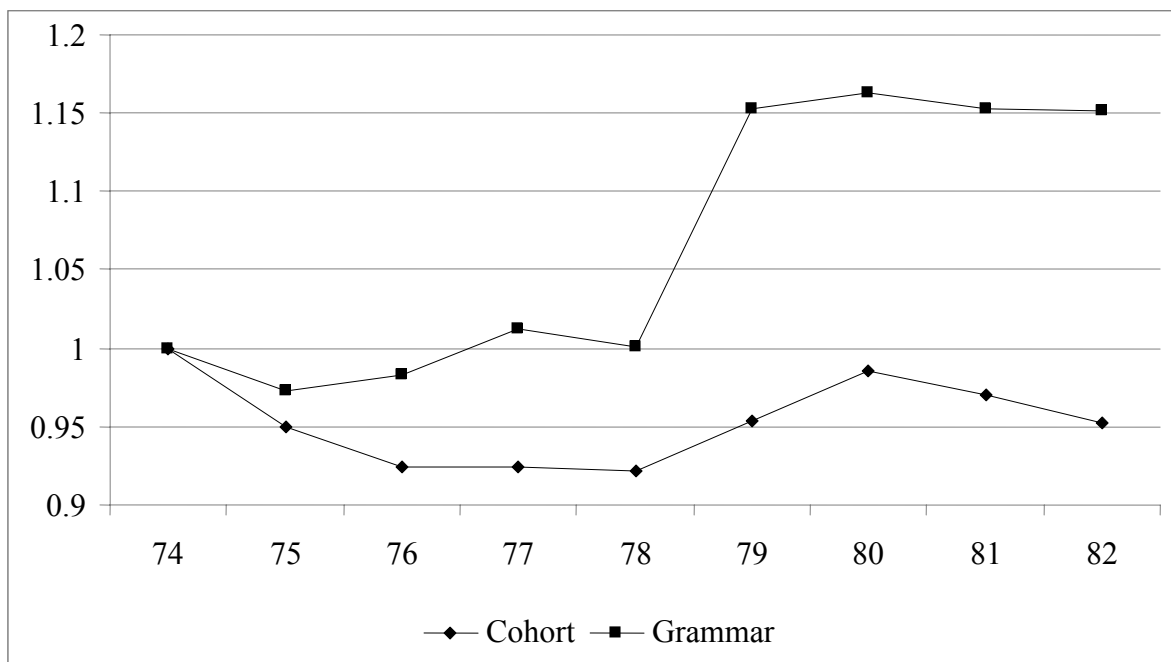
Figure 3: Educational Outcomes in England, by Year of Birth.
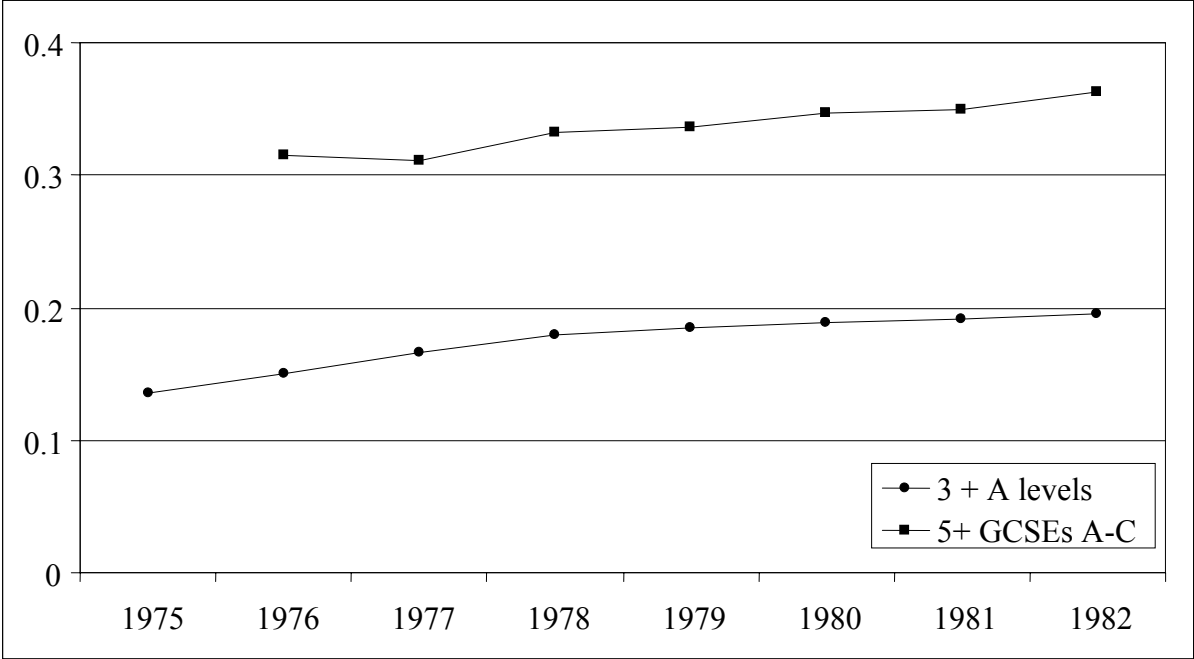
Figure 4: Variation across Cohorts in the (log) Number of Students Attending Elite Schools in Northern Ireland.

Note: Using the area-level data, the graph shows the change across cohorts in the (log) number of students attending grammar schools (cohort 1974 taken as a reference). The average number of students attending elite schools is 15% higher in cohort 1979 than in cohort 1978. Dotted lines show confidence intervals.

Figure 5: Variation across Cohorts in the (log) Number of Successful Students at Age 16.

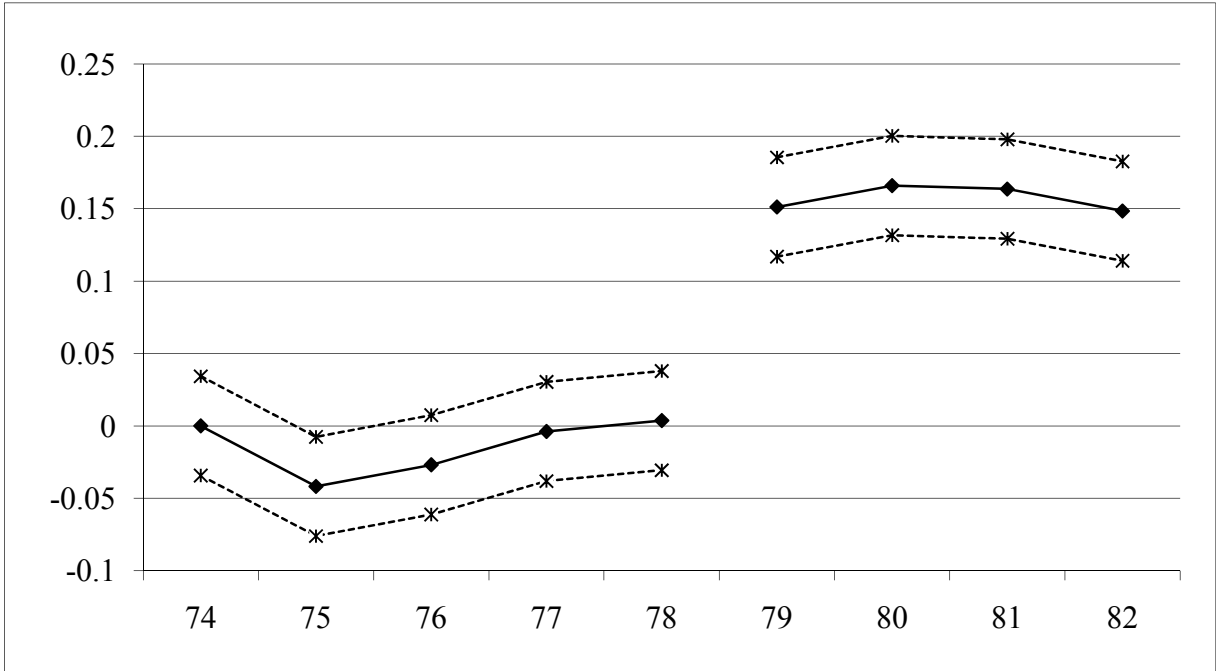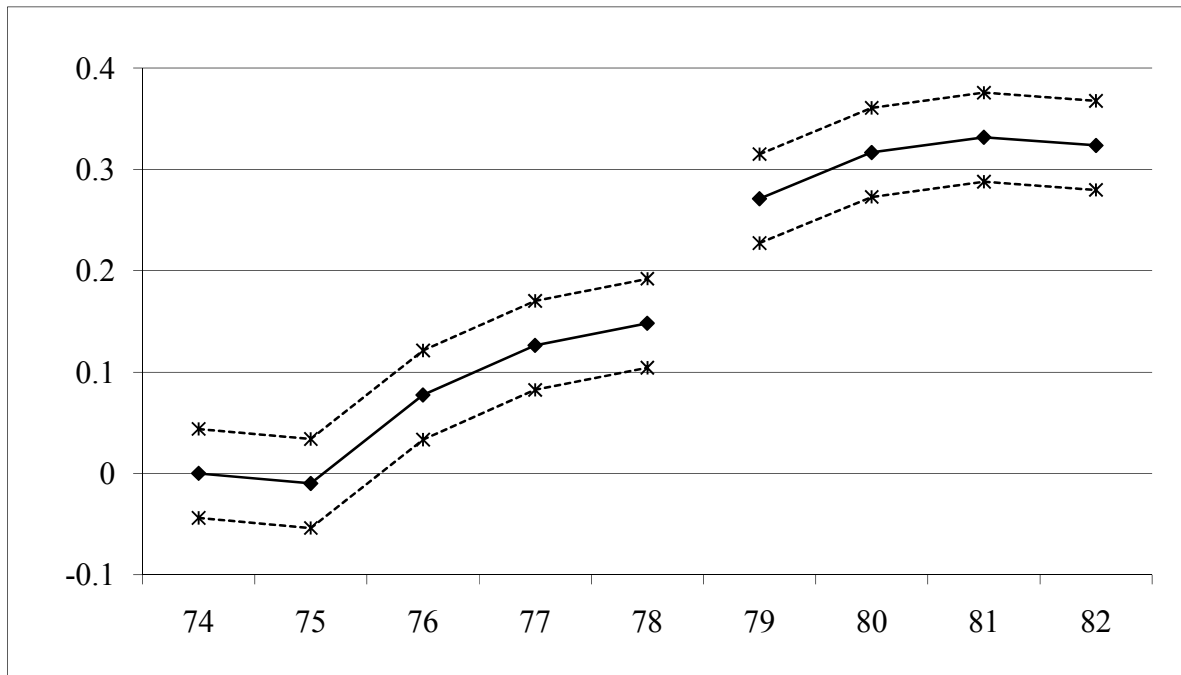Note: Using area-level data, the graph shows the change across cohorts in the average of the (log) number of students obtaining 5 or more GCSEs at grades A*-C (cohort 1974 taken as a reference). Dotted lines show confidence intervals.

Figure 6: Variation across Cohorts in the (log) Number of Successful Students at age 18.
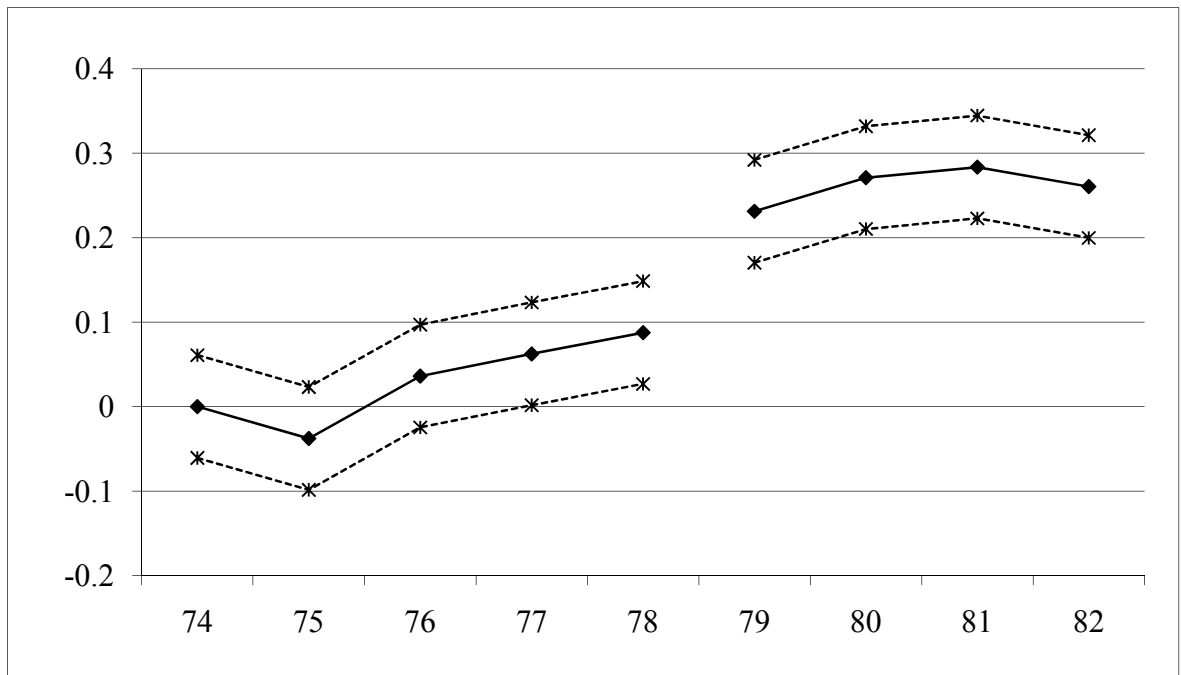
Note: Using the area-level data, the graph shows the change across cohorts in the average of the (log) number of students obtaining 3 or more A levels (cohort 1974 taken as a reference). Dotted lines show confidence intervals.

Figure 7: Variation in the Proportion of Successful Students at Age 16 and Variation in Elite School Attendance Between Pre-Reform and Post-Reform Cohorts.
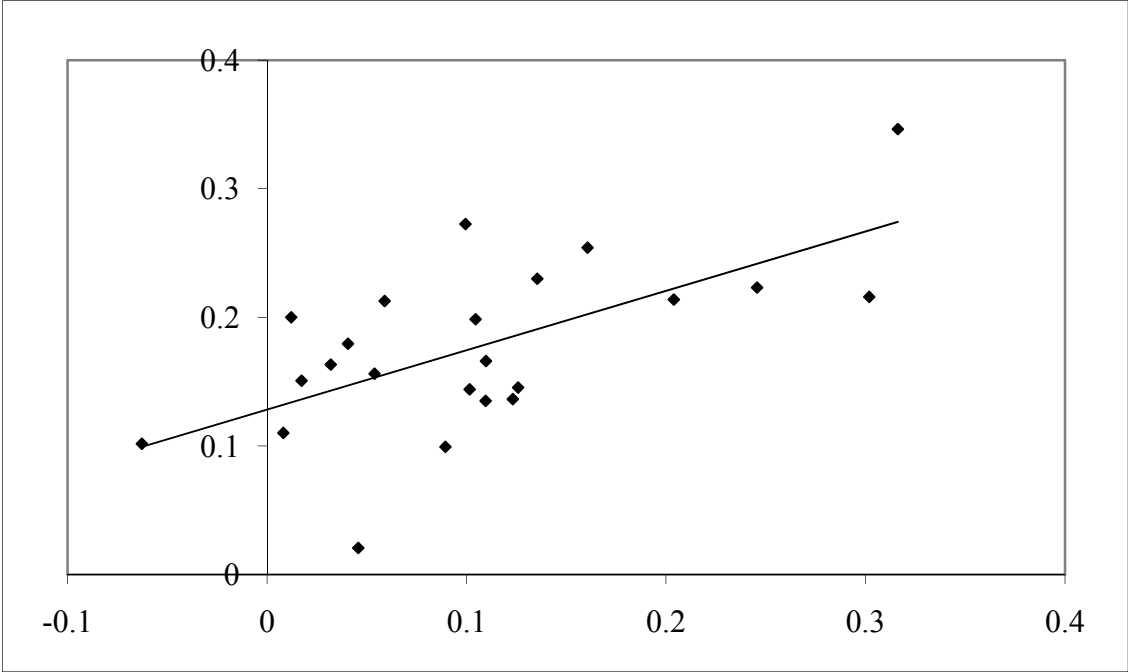
Note : for each local area, the X-axis corresponds to variation in the log proportion attending elite schools between cohorts 1974-1978 and cohorts 1979-1982, whereas the Y-axis corresponds to variation in the log proportion of successful students at age 16.

Figure 8: Variation in the Proportion of Successful Students at Age 18 and Variation in Elite School Attendance Between Pre-Reform and Post-Reform Cohorts.
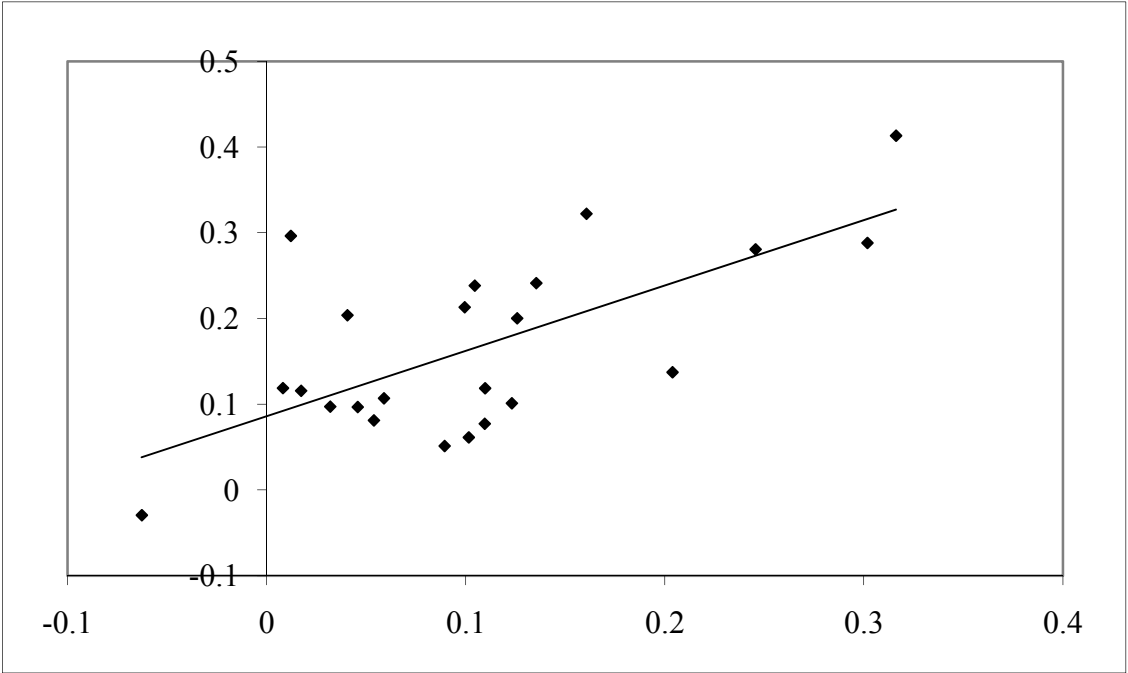
Note : for each local area, the X-axis corresponds to variation in the log proportion attending elite schools between cohorts 1974-1978 and cohorts 1979-1982, whereas the Y-axis corresponds to variation in the log proportion of successful students at age 18.

Table 1: Effect of the Proportion Attending Elite School at Age 11 on Educational Outcomes at Age 16 and 18: An Evaluation Using the Discontinuity in Grammar School Attendance at the Reform Date.

	First-stage	Reduced form		IV		OLS	
	Prop. Elite (G_{ic})	Nb Success. Age 16	Nb Success. Age 18	Nb Success. Age 16	Nb Success. Age 18	Nb Success. Age 16	Nb Success. Age 18
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Prop. Elite (G_{ic})	-	-	-	.405 (.148)	.752 (.217)	.405 (.061)	.715 (.090)
Reform on (c>1978)	.124 (.020)	.050 (.020)	.093 (.031)	-	-	-	-
Year of birth (c)	.011 (.004)	.054 (.004)	.037 (.006)	.049 (.005)	.029 (.007)	.049 (.004)	.030 (.005)
(c-1978) x (c>1978)	-.014 (.007)	-.038 (.007)	-.029 (.011)	-.033 (.006)	-.019 (.009)	-.033 (.006)	-.019 (.009)
Total Nb students(S_{ic})	-.539 (.081)	.712 (.080)	.664 (.121)	.930 (.087)	1.069 (.128)	.930 (.072)	1.055 (.106)
N	207	207	207	207	207	207	207

Note: Column (1) shows estimation of Equation (2); columns (4) and (5) show estimation of Equation (1) for age 16 or age 18 outcomes using "reform on" as an instrumental variable. Columns (2) and (3) show the corresponding reduced-form regressions. Columns (6) and (7) show the corresponding OLS regressions. The outcome at age 16 is the (log) number of students obtaining 5 or more grades at A*-C in the GCSE examination. At age 18, it is the (log) number of students obtaining 3 or more A-levels. All regressions include 23 area fixed effects. Standard errors are in parentheses.

Table 2: The Effect of the Variation in Elite School Attendance at Age 11 on the Variation in Average Educational Outcomes Across Pre-Reform and Post-Reform Cohorts.

	Dependant variable : $Y_{i,after} - Y_{i,before}$		$(S_{i,after} - S_{i,before})$
	at age 16 (5 or + GCSEs A*-C) (1)	at age 18 (3 or + A Levels) (2)	(3)
Panel A: 1974-1982			
$(G_{i,after} - G_{i,before})$	.461 (.141)	.752 (.212)	-.269 (.121)
$(S_{i,after} - S_{i,before})$	1.001 (.230)	.962 (.345)	-
N	23	23	23
Panel B: 1977-1980			
$(G_{i,after} - G_{i,before})$	.367 (.143)	.615 (.152)	-.047 (.108)
$(S_{i,after} - S_{i,before})$	1.304 (.286)	.829 (.305)	-
N	23	23	23

Notes: Columns (1) and (2) show estimation of Equation (3) using age 16 or age 18 outcomes as the dependant variable. Column (3) shows the result of using changes in an area's size as the dependant variable instead of educational outcomes. In Panel A, pre-reform cohorts=1974-1978 and post-reform cohorts=1979-1982. In panel B, pre-reform cohorts=1977-1978 and post-reform cohorts=1979-1980. Standard errors are in parentheses.

Table 3: Effect of the Proportion Attending Elite School at Age 11 on Educational Outcomes at Age 16 and 18: An Evaluation Using the Variation of the Difference in Grammar School Attendance across Areas at the Reform Date

	First-stage	Reduced-form		IV		Total Nb students in the area ($S_{i,c}$)
	($G_{i,c}$)	Nb Success. Age 16	Nb Success. Age 18	Nb Success. Age 16	Nb Success. Age 18	
	(1)	(2)	(3)	(4)	(5)	(6)
Prop. Elite ($G_{i,c}$)	-	-	-	.764 (.297)	.758 (.416)	-
Reform on x T	.150 (.043)	.115 (.045)	.114 (.070)	-	-	.035 (.037)
c x T	-.003 (.009)	-.001 (.010)	.002 (.015)	.001 (.009)	.005 (.013)	-.000 (.008)
$(c\text{-}78)$ x $(c\text{>}78)$ x T	-.008 (.016)	-.025 (.017)	-.026 (.026)	-.018 (.016)	-.020 (.022)	.006 (.014)
Total Nb students in the area ($S_{i,c}$)	-.624 (.088)	.731 (.091)	.610 (.143)	1.208 (.201)	1.083 (.281)	-
N	207	207	207	207	207	207

Note: Column (1) shows estimation of Equation (6). Columns (4) and (5) show estimation of Equation (4) for age 16 or age 18 outcomes using the interaction between a "reform on" dummy and "strongly treated" dummy as an instrumental variable. Columns (2) and (3) show corresponding reduced-form regressions. All regressions include 23 area fixed effects and 9 cohort fixed effects.

Table 4: Effect of the Reform on Entry into Higher Education.

	Strategy 1 (equation 1)			Strategy 3 (equation 4)		
	Red.Form (1)	IV (2)	OLS (3)	Red.Form (4)	IV (5)	OLS (6)
Prop. Elite (G_{ic})	-	.368 (.246)	.627 (.100)	-	.636 (.460)	.665 (.118)
Reform on x T	-	-	-	.095 (.075)	-	-
Reform on	.046 (.033)	-	-	-	-	-
N	207	207	207	207	207	207

	Strategy 2 (equation 3)	
	Panel A *(cohorts 1974-1982)* (7)	*Panel B* *(cohorts 1977-1980)* (8)
($G_{i,after}$ -$G_{i,before}$)	.776 (.242)	.534 (.152)
($S_{i,after}$ -$S_{i,before}$)	1.401 (.393)	1.242 (.304)
N	23	23

Note: Regressions (1) to (3) include cohort size and pre-reform and post-reform trends as control variables (i.e., *c* and *(c-1978)* x *(c>1978)*). Regressions (4) to (6) also include the interaction between these trends and *T* as additional controls. Standard errors are in parentheses.

Table 5: Effect of the Reform by School Types.

	Grammar		Non-grammar	
	Nb students (S_{ict}) (1)	Nb Successful at age 16 (2)	Nb students (S_{ict}) (3)	Nb Successful at age 16 (4)
Reform on (*c>1978*)	.124 (.020)	.135 (.023)	-.044 (.012)	-.113 (.053)
Year of Birth (*c*)	.011 (.004)	.020 (.005)	-.008 (.003)	.141 (.011)
(c-1978) x *(c>1978)*	-.014 (.007)	-.031 (.008)	.002 (.004)	-.078 (.019)
Total Nb students (S_{ic})	.461 (.081)	.657 (.091)	1.178 (.049)	.766 (.209)
N	207	207	207	207

Appendix A

Table A1: Descriptive statistics on the 23 local areas.

	Local Administrative Districts	Nb.Elite Schools	Pre-Reform (cohorts 1974-1978)		Post-Reform (cohorts 1979-1982)	
			Prop. Elite	Weight	Prop. Elite	Weight
Catholic	Antrim, Belfast, Carrickfergus, Castelreagh, Lisburn, Newtonabbey, North Down	7	29.2	11.9	33.0	12.4
	Ards, Down	2	25.0	3.3	27.9	3.4
	Armagh, Cookstown, Craigavon, Dungannon	5	28.0	7.2	28.5	7.8
	Ballymena, Larne, Magherafelt	3	34.8	4.3	36.2	4.3
	Banbridge, Newry & Mourne	5	33.5	6.1	38.4	5.8
	Coleraine	2	33.7	2.1	35.6	1.9
	Derry, Limavady	2	24.9	7.0	27.5	6.9
	Fermanagh	2	32.2	2.5	32.6	2.5
	Omagh	2	47.9	2.0	48.3	2.1
	Strabane	1	15.3	1.8	16.1	1.7
Protestant	Antrim	1	24.0	1.6	29.5	1.5
	Ards, North Down	4	42.7	6.1	45.0	5.6
	Armagh, Banbridge, Craigavon, Newry&Mourne	4	38.3	5.5	36.1	6.4
	Ballymena, Larne	4	42.1	3.9	47.6	3.8
	Belfast, Castelreagh	10	41.7	12.4	46.5	12.4
	Carrickfergus, Newtonabbey	3	27.1	5.7	31.8	5.3
	Coleraine	3	38.3	3.7	42.4	3.5
	Cookstown, Dungannon	1	17.2	2.1	21.9	1.9
	Derry, Limavady, Omagh, Strabane	4	33.5	4.5	36.6	4.4
	Down	1	27.8	1.1	38.0	1.1
	Fermanagh	2	31.8	1.4	42.8	1.3
	Lisburn	2	37.7	3.2	38.9	3.0
	Magherafelt	1	40.4	0.9	44.6	0.9
	Mean outside Belfast:	2.6	32.2	3.6	35.5	3.6
	Std outside Belfast:	1.3	8.3	2.0	8.4	2.1

Appendix B

Table B1: Re-estimation of Equations (1) and (4): a comparison of different specifications.

	First-stage	Reduced-form		IV	
	Prop. Elite (G_{ic})	Nb Success. Age 16	Nb Success. Age 18	Nb Success. Age 16	Nb Success. Age 18
			Equation (1)		
Specif. 1					
Prop. Elite (G_{ic})	-	-	-	.405 (.148)	.752 (.217)
Reform on *(c>1978)*	.124 (.020)	.050 (.020)	.093 (.031)	-	-
Specif. 2					
Prop. Elite (G_{ic})	-	-	-	1.300 (.481)	2.413 (.707)
Reform on *(c>1978)*	.039 (.007)	.050 (.020)	.093 (.031)	-	-
Specif. 3					
Prop. Elite (G_{ic})	-	-	-	.678 (.230)	.547 (.173)
Reform on *(c>1978)*	.039 (.007)	.026 (.010)	.021 (.007)	-	-
Specif. 4					
Prop. Elite (G_{ic})	-	-	-	.708 (.360)	.691 (.281)
Reform on *(c>1978)*	.023 (.007)	.016 (.009)	.016 (.007)	-	-
			Equation (4)		
Specif. 1					
Prop. Elite (G_{ic})	-	-	-	.770 (.307)	.693 (.423)
Reform-on x *T*	.147 (.043)	.113 (.045)	.102 (.070)	-	-
Specif. 2					
Prop. Elite (G_{ic})	-	-	-	2.18 (.89)	1.96 (1.21)
Reform-on x *T*	.052 (.014)	.113 (.045)	.102 (.070)	-	-
Specif. 3					
Prop. Elite (G_{ic})	-	-	-	1.03 (.41)	.330 (.295)
Reform-on x *T*	.052 (.014)	.053 (.022)	.017 (.017)	-	-
Specif. 4					
Prop. Elite (G_{ic})	-	-	-	1.09 (.48)	.332 (.337)
Reform-on x *T*	.045 (.017)	.049 (.022)	.015 (.017)	-	-

Note: Specification 1: Log(Y_{ic}) regressed on Log(G_{ic}) and Log(S_{ic}). Specification 2: Log(Y_{ic}) regressed on G_{ic} and Log(S_{ic}). Specification 3: Y_{ic}/S_{ic} regressed on G_{ic} and Log(S_{ic}). Specification 4: Y_{ic}/S_{ic} regressed on G_{ic}.

Table B2: Effect of the Proportion Attending Elite School at Age 11 on Educational Outcomes at Age 16 and 18: A Re-Evaluation Using Deviation from England as the Dependant Variable.

	First-stage Prop.Elite (G_{ic}) (1)	Reduced form Nb Success. Age 16 (2)	Reduced form Nb Success. Age 18 (3)	IV (1) Nb Success. Age 16 (4)	IV (1) Nb Success. Age 18 (5)	IV (2) Nb Success. Age 16 (6)	IV (2) Nb Success. Age 18 (7)
			Dependant variable : Y_{ic} -Y_{0c}				
Prop. Elite (G_{ic})	-	-	-	.861 (.166)	.680 (.138)	1.133 (.323)	.723 (.382)
Reform on *(c>1978)*	.168 (.019)	.165 (.031)	.114 (.026)	-	-	-	-
N	184	161	184	161	184	161	184

Note: Models (1) to (5) include area fixed effects, cohort size in Northern Ireland (S_{ic}) and cohort size in England (S_{0c}) as control variables. Models (6) and model (7) use pre-reform and post-reform cohort trends as additional control variables (i.e., *c* and *(c-1978)* x *(c>1978)*). Standard errors are in parentheses.

Appendix C

The combination and re-combination of the three ability groups *A*, *B* and *C* define two school contexts before the reform (*s=A* for grammar schools and *s=B+C* for non grammar) and two new school contexts after the reform (*s=A+B* for grammar, *s=C* for non grammar). If we denote $y_s(g)$ the average outcome of ability group *g* in school context *s*, the average outcome in grammar school is $y_A(A)$ before the reform and $q_A\, y_{A+B}(A) + q_B\, y_{A+B}(B)$ after the reform, where q_A represents the weight of group A in grammar school after the reform (and $q_B =1-q_A$). Using this notation, the effect of the reform on the average outcomes of grammar schools is,

$$\Delta(G) \equiv (q_A\, y_{A+B}(A) + q_B\, y_{A+B}(B)) - y_A(A),$$

which can be rewritten,

$$\Delta(G) = q_A\, [y_{A+B}(A) - y_A(A)] + q_B\, [y_{A+B}(B) - y_A(A)].$$

This expression shows that the effect of the reform on average outcomes in grammar schools is a weighted average of an ability effect (i.e., $y_{A+B}(B) - y_A(A)$) and a contextual effect (i.e., $(y_{A+B}(A) - y_A(A))$. This contextual effect is precisely the effect on top ability pupils (*A* pupils) of having new peers, with relatively lower ability (B pupils).

Given this fact, it is clear that $\Delta(G)$ does not point identify the contextual effect of the reform on top ability students. However, under the assumption that pupils who are top ability at age 11 perform better at age 16 or 18 than pupils who are only mid-ability (i.e., $y_{A+B}(B)<y_{A+B}(A)$), it is easy to check that it provides a lower bound for this specific contextual effect. Specifically, under the simple assumption that $y_{A+B}(B)<y_{A+B}(A)$, we have,

$$\Delta(G)<y_{A+B}(A) - y_A(A).$$

Hence, $\Delta(G)$ provides us with a plausible lower bound for the potentially depressing contextual effect of the reform on top ability pupils. With respect to the effect of the reform on non-grammar schools, we have,

$$\Delta(NG) =y_C(C) - (p_B\, y_{B+C}(B) + p_C\, y_{B+C}(C)),$$

where p_B represents the weight of group B in non-grammar school before the reform (and $p_C =1- p_B$). Under the simple assumption that pupils who are mid-ability at age 11 perform better at age 16 or 18 than low ability pupils (i.e., $y_{B+C}(C)<y_{B+C}(B)$), it is again not very difficult to show that $\Delta(NG)$ provides an interesting lower bound for the contextual effect of the reform on low ability pupils, i.e. a lower bound for $(y_C(C) - y_{B+C}(C)$. Specifically, under the sole assumption that $y_{B+C}(C)<y_{B+C}(B)$ we have,

$$\Delta(NG)<y_C(C) - y_{B+C}(C).$$

Assuming that there is no negative externality on group A (i.e., $(y_{A+B}(A)=y_A(A))$, and using $\Delta(G) = (q_A\, y_{A+B}(A) + q_B\, y_{A+B}(B)) - y_A(A)$, our results that $\Delta(Grammar)=0$ implies that $y_{A+B}(B)= y_A(A)$ i.e., group B post-reform does as well as group A pre-reform.