



BANCA D'ITALIA  
EUROSISTEMA

## Temi di Discussione

(Working Papers)

STEM graduates and secondary school curriculum:  
does early exposure to science matter?

by Marta De Philippis

March 2017

Number

1107





BANCA D'ITALIA  
EUROSISTEMA

# Temi di discussione

(Working papers)

STEM graduates and secondary school curriculum:  
does early exposure to science matter?

by Marta De Philippis

Number 1107 - March 2017

*The purpose of the Temi di discussione series is to promote the circulation of working papers prepared within the Bank of Italy or presented in Bank seminars by outside economists with the aim of stimulating comments and suggestions.*

*The views expressed in the articles are those of the authors and do not involve the responsibility of the Bank.*

*Editorial Board:* INES BUONO, MARCO CASIRAGHI, VALENTINA APRIGLIANO, NICOLA BRANZOLI, FRANCESCO CAPRIOLI, EMANUELE CIANI, VINCENZO CUCINIELLO, DAVIDE DELLE MONACHE, GIUSEPPE ILARDI, ANDREA LINARELLO, JUHO TANELI MAKINEN, VALERIO NISPI LANDI, LUCIA PAOLA MARIA RIZZICA, MASSIMILIANO STACCHINI.

*Editorial Assistants:* ROBERTO MARANO, NICOLETTA OLIVANTI.

ISSN 1594-7939 (print)

ISSN 2281-3950 (online)

*Printed by the Printing and Publishing Division of the Bank of Italy*

# STEM GRADUATES AND SECONDARY SCHOOL CURRICULUM: DOES EARLY EXPOSURE TO SCIENCE MATTER?

by Marta De Philippis\*

This paper focuses on students at the very top of the ability distribution and explores whether strengthening high school science curricula affects their choice of enrolling in and completing a Science, Technology, Engineering and Maths (STEM) degree at university. The paper solves the standard endogeneity problems by exploiting the different timing in the implementation of a reform that encouraged secondary schools in the UK to offer more science to high ability 14- year-olds. Taking five more hours per week of science in secondary school increases the probability of enrolling in a STEM degree by 1.2 percentage points and the probability of graduating in these degrees by 3 percentage points. The results mask substantial gender heterogeneity: while girls are as willing as boys to take advanced science in secondary school - when offered -, the results on pure STEM degrees at university are entirely driven by boys. Girls are encouraged to choose more challenging subjects, but still opt for the most female-dominated ones.

**JEL Classification:** I23, J24, H52.

**Keywords:** STEM, high school curriculum, field of study, gender bias.

## Contents

1. Introduction.....	5
2. Data and institutional settings.....	8
2.1 The English school system .....	8
2.2 Science in secondary school .....	9
2.3 Data.....	11
3. Empirical strategy .....	12
4. Results.....	15
4.1 Compliers' characterization .....	19
4.2 Heterogeneity.....	20
4.3 Checks to the identification strategy .....	21
5. An alternative identification strategy.....	23
6. Potential alternative mechanisms .....	26
7. Conclusions .....	27
References .....	29
Figures and tables .....	32
Appendix .....	45

---

\* Bank of Italy, Directorate General for Economics, Statistics and Research and Centre of Economic Performance, LSE.



# 1 Introduction<sup>1</sup>

In the new heavily globalized and innovation driven economy, increasing the number of Science, Technology, Engineering and Math (STEM)<sup>2</sup> university graduates is considered crucial for long-term productivity and growth. Additionally, a STEM degree represents a very profitable private investment for college graduates themselves (Altonji et al. (2015) for a review). Fields of study are an important determinant of lifetime earnings, as they seem to matter more than the quality of the university attended (James et al. (1989); Kirkeboen et al. (2016); Arcidiacono et al. (2016)) and than the level of education achieved: nowadays intra-educational income differences are comparable to inter-educational differences (Altonji et al. (2012)). Finally, also non-monetary returns, e.g. the degree of temporal flexibility in the job,<sup>3</sup> seem to be high in STEM occupations (Goldin (2014)). Still, the general consensus among policy-makers is that the current supply of STEM skills is insufficient and it presents a potentially significant constraint on future economic activity.<sup>4</sup>

This paper explores how much of the lack of STEM graduates can be attributed to the quality of secondary schools, and in particular to the curriculum they offer. It investigates whether exposing students at the very top of the ability distribution to more science in secondary school increases by itself the supply of STEM graduates and whether, by reducing uncertainty, it shrinks the gender gap in enrollment in STEM degrees.

The identification of the effect of studying more science in secondary school is difficult to achieve because of the selection of students into different schools -based on the curriculum they offer- and the selection of students into different courses - within the school they chose. Thanks to a novel dataset that I obtained by combining different administrative sources from England, I exploit the different timing in the introduction of an advanced

---

<sup>1</sup>I thank Steve Pischke and Esteban Aucejo for very precious guidance, supervision and encouragement. I also thank Oriana Bandiera, Lorenzo Cappellari, Georg Graetz, Monica Langella, Alan Manning, Barbara Masi, Stephan Maurer, Sandra McNally, Guy Micheals, Sauro Mocetti, Michele Pellizzari, Lucia Rizzica, Jesse Rothstein, Paolo Sestito, Olmo Silva, Alessandro Vecchiato and Giulia Zane and participants at the LSE labour and education work in progress seminars, at the 2015 CEP conference, at the 5th FRDB workshop, at the 6th IWAAE workshop and at the XXX AIEL conference for providing me with very useful comments and information. The views expressed in this article are those of the author alone and do not necessarily reflect the official views of the Bank of Italy.

<sup>2</sup>Throughout the paper I define as “STEM” the following degrees: Physical science, Mathematical and Computer science and Engineering.

<sup>3</sup>How important it is to stay long or particular hours in the office.

<sup>4</sup>See for instance, (UK HM Treasury and BIS, 2010; The President’s Council of Advisor on Science and Technology, 2012; European Commission, 2010). Overall, STEM employment grew three times more than non-STEM employment over the last twelve years, and it is expected to grow twice as fast by 2018. According to a report by the Information Technology and Innovation Foundation (2010), the number of STEM graduates in the US will have to increase by 20-30% by 2016 to meet the projected growth of the economy.

science course in English secondary schools to solve the endogeneity issue. In 2004 the UK government introduced an entitlement to study advanced science for 14 year old high ability students, with the explicit aim of fostering enrollment in post-secondary science education. This resulted in a strong increase in the number of schools offering advanced science (from 20% in 2002 to 80% in 2011) and, therefore, in the share of students taking advanced science (from 4% in 2002 to 20% in 2011). The increase was almost entirely concentrated on high ability students<sup>5</sup> (Figure 1). In the spirit of [Joensen and Nielsen \(2009\)](#), I exploit the staggered introduction of the policy and the three year time lag between the moment when students choose their secondary school (age 11) and the moment when they choose their field courses (age 14). I therefore consider only the effect on students who were unexpectedly exposed to the advanced science course because their school started to offer it after their enrollment. Moreover, I evaluate the robustness of my results by adopting a second, alternative, identification strategy that exploits (unpredictable) over time variations in schools catchment areas, derived by the fact that schools in England, when oversubscribed, select students based on home-to-school distance.

My empirical findings can be summarized as follows: taking advanced science at age 14 (which implies five more hours of science per week) increases the probability of choosing science at age 16 by 5 percentage points and that of enrolling in STEM degrees at university by about 1.2 percentage points. Moreover, it increases the likelihood that students complete these degrees at university by 3 percentage points. This is important, given the very low persistence of students in this kind of degrees (for the US [Arcidiacono et al. \(2016\)](#); [Stinebrickner and Stinebrickner \(2014\)](#)).<sup>6</sup> Second, I find that the effect on STEM degrees (in its narrow definition) is very heterogeneous and it is concentrated on boys: the gender gap in STEM degrees enrollment widens as a consequence of this policy. This is not because less girls take advanced science at age 14 - boys and girls at this stage select into advanced science in the same proportion, when offered it - but because girls, when exposed to more science in secondary school, even if induced to take more challenging subjects<sup>7</sup>, still opt for the most female-dominated ones. Girls choose medicine or psychology, not engineering.

Taken together, my findings can inform ongoing debates over government intervention to address apparent mismatches and market frictions in the supply and demand of

---

<sup>5</sup>I define high ability students as those who were in the top 30 percentile of the primary school grades distribution. The increase for these students was around 35 percentage points, from 15% to about 50%.

<sup>6</sup>There is a problem of persistence in STEM majors also in England: in the cohort starting university in 2011, out of the 17% of students enrolled in a STEM major, only 17% graduated in the same STEM major within three years (this figure is 20% on average for the other majors).

<sup>7</sup>I define as challenging the subjects usually taken by students achieving very high grades in primary school.



post-secondary fields of study. My results suggest that, to reinvigorate STEM education and high-skilled STEM education in particular, governments should consider a policy aimed at offering more science courses to high ability students during secondary schools. I estimate that the policy I consider contributed to one third of the increase in the share of STEM graduates in England between 2005 and 2010.

This paper speaks to the growing literature that seeks to explain choices of university fields of study (Arcidiacono et al., 2012; Arcidiacono, 2004; Beffy et al., 2012; Stinebrickner and Stinebrickner, 2014; Wiswall and Zafar, 2014). Most of the evidence so far comes from surveys or informational experiments and the results are mixed. I explore in particular the role of schools and the curriculum they offer to boost students' preparation, a policy area which may potentially generate positive effects and over which is easy for governments to intervene. A recent research (Ellison and Swanson (2012)) pointed out that there is indeed a large heterogeneity in secondary schools effectiveness in developing talents in technical subjects like math, which does not depend on differences in schools' student composition. Moreover, not only every single government has to take at some point the decision about how to design its country secondary school curriculum, but also, differently from other policies like changes in peers, this is not a zero sum choice: everybody may potentially benefit from a well designed curriculum. Still, excluding some recent studies that evaluate the effects of secondary school curricula using quasi-experimental evidence (Joensen and Nielsen, 2009; Joensen and Nielsent, 2016; Cortes et al., 2015; Goodman, 2017), there is little quantitative work on the effects of secondary school courses (Altonji et al., 2012).

My paper improves on the existing literature in several ways.<sup>8</sup>

First, I address both layers of selection of students into courses. Most studies (Altonji, 1995; Levine and Zimmerman, 1995; Betts and Rose, 2004) use across school variation in the type of curriculum offered and do not fully address the possible selection of students into schools, based on the curriculum they offer. Since family background and individual motivation are important determinants of the choice both of university degrees and of secondary schools, the bias in estimates that do not take into account selection into schools could be important and could lead to an overestimation of the effects. I show that, even in my context, where the variation in curriculum is induced by a policy, the inclusion of school fixed effects and the presence of an instrument turn out to be crucial to correctly identify the effect of interest.

Second, the policy I consider does not intervene on other subjects, allowing me to identify the effect of offering more science courses only. Instead, changes in secondary

---

<sup>8</sup>I mention here papers that look at the effect both on earnings and on degrees, even if most of the literature looks at earnings without focusing on the effect on the choice of degree. Still, focusing on the impacts on degree choice helps understanding the mechanisms.

school curricula usually imply a restructuring of many different courses and it is difficult to isolate the effect of one single subject (Altonji, 1995; Joensen and Nielsen, 2009; Joensen and Nielsent, 2016; Gorlitz and Gravert, 2015; Jia, 2014). While also my treatment has multiple components, since taking advanced science implies also changes in classroom heterogeneity and composition,<sup>9</sup> I disentangle the curriculum from the peer channel by using an instrument for peers. I find that the effect of the advanced science course persists even after controlling for changes in peers' characteristics.

Third, the compliers for my instrument are extremely high ability students: I therefore look at the effect for those students of highest interest for policy-makers because they are characterized by potentially very high probability of succeeding in STEM degrees and of making important contributions to scientific and technological fields. Most of the existing empirical studies (Goodman (2017); Cortes et al. (2015)) analyze the effect of remedial policies or changes in minimum math requirements, which affect almost entirely low ability students, not likely to enroll at the university at all, or (Joensen and Nielsen (2009); Joensen and Nielsent (2016)) students for whom taking science is rather costly.<sup>10</sup> Finally, boys and girls attend the advance science course in the same proportion, therefore allowing to test the gender heterogeneity of the effect without worrying about differences in compliance with the instrument.

The remainder of this paper is organized as follows. In Section 2, I describe the data, the English school system and the reform of the advanced science program in UK secondary schools. Section 3 provides an overview of the main identification strategy. Section 4 presents the estimated impact of advanced science on post-16 educational outcomes and it checks the identifying assumptions and the robustness of the results. Section 5 shows results using an alternative identification strategy; Section 6 inspects the mechanisms behind the estimates. Finally, Section 7 concludes.

## 2 Data and institutional setting

### 2.1 The English school system

Compulsory education in England is organized in four phases (Key Stages KS). At the end of each stage students are evaluated in standardized national exams. Figure 2 shows a time-line of the English educational system. Pupils enter school at age 4, the Foundation

---

<sup>9</sup>Because the advanced science course provides the possibility of taking a course exclusively attended by other very high ability students.

<sup>10</sup>These studies exploit for instance changes in minimum math requirements across US states over time or compare students just below or just above the threshold for attending remedial classes in math and find modest effects on earnings, concentrated on low-SES students. In my setting, instead, compliers include also extremely high ability students, within the same school.

Stage, then they move to Key Stage 1 (KS1- spanning ages 5 and 6) and Key Stage 2 (KS2, from age 7 to age 11).<sup>11</sup> At the end of KS2 children leave primary school and go to secondary school. Admission to secondary school is based on criteria usually set by the school or by the local council. Usually students indicate up to five alternatives and schools give priority to children who live close to the school or whose siblings already attend the school. Selection cannot be based on previous grades or ability. During secondary schools students progress to Key Stage 3 (KS3, age 12-14) and Key Stage 4 (KS4, age 15-16).

My analysis focuses on subjects taken at age 14 (KS4), the first time when students have the possibility of choosing some subjects.<sup>12</sup> In particular, at that stage there are 6 compulsory subjects and students typically choose other 4 to 6 subjects.<sup>13</sup> At age 16 compulsory education ends and students may choose to further continue their studies. I will evaluate whether they decide to continue studying and which subjects they choose in the following education phases. In particular, students who want to go to university should spend at age 16 other two years at school before applying. This phase is called Key Stage 5 (age 17-18) and may take place in the same secondary school (about 60% of the schools also offer KS5 courses) or in a different school. Again, students can choose more vocational or more academic-oriented qualifications (the so-called A levels).<sup>14</sup> Finally, higher education usually begins at age 19 with a three-year bachelor's degree. Admission to university is usually based on which subjects were chosen and on the grades achieved at KS5.

## 2.2 Science in secondary school

While science is a core component of the National Curriculum at age 14 (KS4), there are several different ways to fulfill the requirement. All students are required to study the basics elements of all three natural sciences (physics, chemistry and biology) and should at least take the so-called 'single science' or core science course (which is worth one KS4 exam and provides the basic knowledge in all three subjects). Students can, moreover, choose to take the 'double science' course (worth two exams) which leads to more knowledge in all the three subjects or the 'triple science' course (which is called advanced science and

---

<sup>11</sup>KS1 corresponds to grade 1 and 2 in the US school system, KS2 to grades 3,4 and 5.

<sup>12</sup>A number of different qualification types are available to young people at KS4, varying in their level of difficulty. These include: GCSE (the most common qualification in England and the most academic oriented), and other more vocational qualifications. I will only consider GCSE qualifications or GCSE equivalent qualifications.

<sup>13</sup>The six compulsory subjects are: English, math, (single) science, information and communication, physical education and citizenship. Students in general take overall between 10 and 12 qualifications.

<sup>14</sup>Students who want to go to university usually choose the more academic-oriented exams (slightly less than half of each cohort undertakes at least one A-level exam at age 16 and usually students take three A level or equivalent qualifications) and are free to choose any subject.

is equivalent to take one full exam in each of the three natural science subjects). Finally students can also take more vocational science qualifications. Taking triple science implies both longer instruction time (5 hours more per week on average) and the coverage of more complex science topics.<sup>15</sup> Double science and, more recently, triple science provide the standard routes into the fulfillment of KS4 requirements.

In 2004 the UK Government published a ten-year investment framework for science and innovation (UK Government, 2004). The framework set out the Government's ambition for UK science and innovation over the next decade and emphasized in particular the need for more graduates in science. Taking triple science was considered extremely important, because "it gives students the necessary preparation and confidence to go on and study science" (Confederation of British Industry). The document established an entitlement to study triple science for very high ability students (the students on the top 40% of the grade distribution, i.e. those who achieved level 6 or above in the KS3 science exam). The document stated that by 2006, students who obtain a high enough grade in their science exam at age 12 (KS3) have the right to take triple science and their schools must arrange a way to offer the above mentioned course.<sup>16</sup> The explicit aim of the policy was to foster enrollment in STEM degrees. Many schools encountered difficulties in meeting this requirement by 2006 and this explains the staggered adoption of the policy. On average however the result was a very large increase in the number of schools offering triple science. While in 2002 less than 20% of schools offered triple science, by 2011 the share became more than 80% (see Figure 1). Between 2002 and 2011 the share of students choosing triple science increased from 4% to 20% and the increase was mostly concentrated among high ability students (for whom the share increased from 15% to 50%).

For my identification strategy, it is crucial to understand what drives the staggered adoption of the policy and whether it is mainly ascribable to supply or demand driven factors. There are several, mainly supply driven, reasons why the exact timing of the introduction of the triple science option differs across schools. First, the lack of specialized teachers. 50% of science and math students in English secondary schools are not taught by teachers specialized in the subject. For teachers teaching outside their expertise it requires more time for them to get familiar with the material. Second, the school size: for small schools it is difficult to offer a large number of subjects. With the ten-year investment framework, the government encouraged new collaborative arrangements with

---

<sup>15</sup>In this case students study more difficult topics such as electric current, transformers, some medical application, more quantitative topics in chemistry etc.

<sup>16</sup>In particular the government stated that "all pupils achieving at least level 6 [Level 6 or above is equivalent to the top 40% of students] at KS3 should be entitled to study triple science at KS4, for example through collaborative arrangements with other schools".

other schools (to jointly provide triple science). However, setting these agreements up takes time and many schools need the support of their Local Education Authority (LEA). The exact timing of the conclusion of these agreements is therefore uncertain. Finally, support and pressure on schools to fulfill the entitlement to triple science was provided at the LEA level.<sup>17</sup> Some LEAs were not as supportive as others regarding the introduction of triple science: the increase in the share of schools offering triple science was indeed very heterogeneous across different LEAs.

## 2.3 Data

By combining different administrative sources, my final dataset follows all students in public (so-called “maintained”) schools in the England,<sup>18</sup> from primary school till the end of their university career.

I obtain information on students demographic characteristics from The Pupil Level Annual School Census (PLASC) that collects information on students’ gender, ethnicity, Free School Meal Eligibility (FSM), Special Education Needs (SEN), language group as well as their postcodes. The National Pupil Database (NPD) provides instead information on students’ attainments in all their Key Stage exams (from KS1 till KS5) as well as on every single subject chosen (and the corresponding grade) in KS4 and KS5 and on school characteristics (peer groups, type of school, teachers’ hirings, school location etc.). From the NPD dataset I obtain also the information about which courses are offered by each school. In particular, I follow the official methodology used by the English Department of Education and I infer that a school offers a course if at least one pupil at the school took an assessment in that specific course and year.<sup>19</sup> I then link the NPD to the universe of UK university students, the Higher Education Statistical Agency (HESA) dataset. The HESA dataset provides information on whether pupils progress to university, on their degree, on the institution they attend and on whether they graduate and in which degree. I combine these two data sources to create a dataset following the entire population of five cohorts of English school children. My sample includes pupils who finished compulsory education (took KS4 examinations, at age 16) between the academic years 2004/2005 and 2009/2010<sup>20</sup> Using information on the secondary school attended by each individual,

---

<sup>17</sup>LEAs organize courses both on how to organize the time schedule to fit the new curriculum and on the new material covered and encourage school-to-school learning. There is large heterogeneity on how actively different LEAs promoted and pushed the introduction of the Triple Science option in schools. In total there are 152 local authorities in England.

<sup>18</sup>The dataset refers only to England and it excludes private schools, that however educate a small share (7%) of British children.

<sup>19</sup>My results are robust to different definitions (at least 5 pupils, at least 5% of the students, for at least two consecutive years etc.) and all different definitions are extremely highly correlated.

<sup>20</sup>After 2010, there would be no information on university outcomes, because I only have data on

I match the individual record with school level data on whether the school was offering triple science when the student applied and three years later, when she had to choose her KS4 subjects.

Finally, I impose a set of standard restrictions on the data. First, I exclude special schools, hospital schools, schools where there is a three tier system instead of a two tier system. Second, I only use students who can be tracked from KS2 to KS4.<sup>21</sup> This leaves me with approximately 530,000 students per cohort.

### 3 Empirical strategy

The main aim of the paper is to estimate the effect of being exposed to more science at age 14 on the choices of subjects at age 16 and at university and on the probability of continuing to study. In particular the equation I would like to estimate is of the following type:

$$Y_{ist} = \beta_1 TS_{ist} + \beta_2 X_{ist} + v_{ist} \quad (1)$$

where  $TS_{ist}$  is the dummy equal to 1 if student  $i$  in secondary school  $s$ , in cohort  $t$  takes triple science and 0 otherwise;  $X_{ist}$  are some controls (in my case, I include controls at the school and student level, school fixed effects  $\delta_s$  and time/cohort fixed effects  $\delta_t$ ).  $Y_{ist}$  is the outcome variable, usually a dummy indicating whether the student takes science at age 16 or at the university (and 0 if she does not take science or does not continue studying). Finally,  $v_{ist}$  is the error term.

The main identification challenge when studying the effects of secondary school courses on post-secondary school outcomes, is to correct for selection bias.

To fix ideas, consider the case in which students choose between taking more science in secondary school ( $TS = 1$ ) or not ( $TS = 0$ ). The observed choice of university degree ( $Y$ ) can be linked to potential degrees ( $Y_j$  where  $j = 1, 0$ ) and the type of science in secondary school ( $TS$ ) as:

$$Y = Y_0 + TS(Y_1 - Y_0) \quad (2)$$

The OLS estimates of the effect of choosing more science in secondary school, can be written as follows:

$$E(Y|D = 1) - E(Y|TS = 0) = E(Y_1|TS = 1) - E(Y_0|TS = 0) \quad (3)$$

---

university results till 2013. Before 2005, there is no information on whether the school was offering triple science when the student applied to the school, because the data collection starts in 2002 and there are three years of lag.

<sup>21</sup>I checked whether this selection generates any bias (i.e. is correlated with the instrument) and this is not the case. The results are available upon request.

The main challenge is that students selecting into certain secondary school courses would have different potential outcomes in any case, meaning that a simple OLS comparison does not provide the right counterfactual ( $E(Y_0|TS = 0) \neq E(Y_0|TS = 1)$ ). In practice there are two layers of selection: selection of students into schools offering triple science and selection of students into triple science, for a given school.

Let's call  $S$  a dummy equal to 1 if the school attended by student  $i$  offers triple science and 0 otherwise. Then, the OLS estimates can be written as follows:

$$\begin{aligned}
E(Y|TS = 1) - E(Y|TS = 0) &= \underbrace{E(Y_1 - Y_0|TS = 1, S = 1)}_{\text{ATT}} + \\
&P(S = 1|TS = 0) \underbrace{[E(Y_0|TS = 1, S = 1) - E(Y_0|TS = 0, S = 1)]}_{\text{selection into courses within school}} + \\
&P(S = 0|TS = 0) \underbrace{[E(Y_0|TS = 1, S = 1) - E(Y_0|TS = 0, S = 0)]}_{\text{selection into schools+courses}}
\end{aligned}$$

I address the selection problem by tackling both the first and the second layer of selection. Selection of students into courses within the same schools is addressed by collapsing the analysis at the school level. I do this indirectly by choosing an instrument ( $Z_{st}$ ) that varies at the school-cohort level only, which implies I only use this type of variation for my estimates.<sup>22</sup> Most papers (in the spirit of [Altonji \(1995\)](#)) use school average curriculum as instrument and therefore address this first type of selection only. This leaves space, however, to endogeneity due to selection of students into schools offering different curricula. I address this other layer of selection exploiting the panel structure of my data and the specific features of the English educational system.

The school fixed effects in equation 4 take care of time invariant school heterogeneity, therefore address the concern that the overall quality of the school, of the students usually attending it or of the teachers' pool is different across schools and may be correlated with the probability of offering triple science.

The time fixed effects absorb cohort effects or the presence of policies that uniformly affect the entire English school system, like changes in the exam structure or syllabus.

Still, there may exist time varying factors, like yearly changes in cohort quality, that may bias my estimates because they may be correlated both with the timing of the introduction of triple science course and with the students' willingness to take science subjects. I use, as instrument for  $D_{ist}$ , the dummy  $Z_{st}$  which equals one if the cohort  $t$  in school  $s$  was unexpectedly exposed to the triple science option.

---

<sup>22</sup>I do not collapse the entire analysis at the school-cohort level just to increase precision of the estimates by adding individual level controls.

My identification strategy is based on the following first stage equation:

$$TS_{ist} = \gamma_1 Z_{st} + \gamma_2 X_{ist} + \zeta_s + \zeta_t + e_{ist} \quad (4)$$

where  $TS_{ist}$  is the dummy equal to 1 if student  $i$  in secondary school  $s$ , in cohort  $t$  takes triple science and 0 otherwise;  $X_{ist}$  are school and student controls;  $\delta_s$  are school fixed effects and  $\delta_t$  are year fixed effects. Finally,  $e_{ist}$  is the error term.

For my instrument, I rely on the time span between the time when students choose secondary schools (age 11) and the time when they choose their optional subjects (age 14).  $Z_{st}$  is a dummy equal to 1 if school  $s$  was not offering triple science when students from cohort  $t$  applied to secondary schools but starts to offer triple science by the time they choose their KS4 subject, three years later, and I only include schools not offering triple science when students applied. I compare two types of students, a priori identical because they all selected schools not offering triple science at age 11: those whose schools unexpectedly started to offer triple science by the time they turned 14 (my treatment group) and those whose school did not offer triple science when they chose subjects at age 14 (my control group).<sup>23</sup>

This strategy mainly relies on two assumptions.

First, the assumption that the information set of both students in the treatment and in the control group at age 11 is the same and does not include the information on whether the school is going to offer triple science in the next three years. This is very likely, given the large time lapse and uncertainty on when exactly teachers/classrooms and time schedules would be ready. Moreover, students are not totally free to choose the school they want: there are exogenous geographical constraints in choosing schools in England. In Section 4.3, I show first that students in my control and treatment group are observationally identical. Second, I show that there is no sign of strategic selection of schools based on whether the schools offer the advanced science course, even if the information is available to parents and students at age 11: students who enrolled at age 11 in schools offering triple science are observationally identical to students who enrolled at age 11 in schools not offering triple science.

Second, the assumption that schools' decisions on when exactly to start offering triple science are related to supply-driven rather than demand-driven factors: schools must decide when to start offering triple science not based on the quality of the current cohort

---

<sup>23</sup>A similar idea, with only one year lag, has been used in [Joensen and Nielsen \(2009\)](#); [Joensen and Nielsent \(2016\)](#), to evaluate the effects of increasing secondary school curriculum flexibility, that induced students to take more math at secondary school in Denmark. I study a different policy that affects very high ability students and identifies the effect of more science only. Thanks to the availability of data on previous test scores and of many cohorts, I am able to use within school variation and to explore more in details the effect on choices of university degrees.



attending the school. In Section 2.2 I described some supply driven reasons why schools may delay the introduction of triple science. In Section 4.3 I show that the timing of the introduction of the triple science option is not correlated with (observable) characteristics of current students in the school and that schools in my treatment and control group were on the same trend before the introduction of the triple science course.

## 4 Results

Table 2 presents the main estimates of the effect of taking triple science at age 14 on the probability of choosing at least one natural science subject at age 16 (KS5) and a STEM degree at the university.<sup>24</sup> The Table proceeds by estimating the effect of interest under different specifications. Column 1 displays results from a simple OLS regression; in column 2 I add school fixed effects; column 3 follows Altonji (1995) and uses as instrument for triple science the share of students taking triple science in school  $s$  and year  $t$ ; column 4 uses my first instrument ( $Z_{st}$ ) and some school time varying controls<sup>25</sup>, but does not include school fixed effects; column 5 shows results from my preferred specification that uses my instrument and exploits within school over time variation only; finally column 6 adds a school-specific trend. Reassuringly, the coefficients of columns 5 and 6 are very similar, suggesting that schools offering triple science are on a similar trend. Column 7 estimates the specification of equation 4, but it eliminates controls ( $X_{ist}$ ). The coefficients of columns 5 and 7 are again very similar, suggesting that -conditional on my fixed effects- the instrument is quasi randomly assigned. As expected the bias in the OLS estimates is upward: the coefficient indeed gets smaller as I correct for all different layers of selection. The Table shows that, if a student strengthens her science preparation at age 14, she is 5 percentage points more likely to take science at age 16 and 1.5 percentage points more likely to choose a STEM degree at the university.

Table 3 shows the coefficients obtained from estimating equation 4 on other outcomes at age 14 (KS4), age 16 (KS5) and university. The top panel shows results on KS4 grades and on the number of exams taken in KS4 and KS5. Since triple science is more difficult, taking it reduces the average science grade at KS4. Columns 2 and 3 show that there are no spillovers on other subjects' grades. Columns 4 and 5 investigate whether the total number of qualifications taken at age 14 and 16 changes, as a consequence of the new course offered. The results show that the number of exams taken at age 14 slightly increases.

---

<sup>24</sup>The dependent variables in all cases are dummies equal to one if students attend a certain course and equal to 0 if they do not attend those courses or do not continue studying.

<sup>25</sup>In particular, the share of girls attending school  $s$  in year  $t$  and the share of FSME (Free School Meal Eligible). In the spirit of Joensen and Nielsen (2009); Joensen and Nielsent (2016).

The second panel refers to outcomes at age 18, the results of KS5 exams. Column 1 shows that the policy does not have any effect on the probability of continuing to study at age 16, probably because the instrument mainly affects high ability students, who would have continued to study in any case. Since a change in the probability of enrolling in science subjects at age 16 may be driven both by a change in the likelihood of continuing to study after age 16 and by a change in the likelihood of choosing science subjects - conditional on continuing-, column 1 shows that the coefficient estimated on KS5 subjects comes entirely from an increase in the second component, because the first is not affected by the policy. The result displayed in column 2 shows that the effect of studying triple science is not limited to the pure natural science subjects but it has spillovers on math, for instance. The third panel refers to university outcomes. Column 1 shows again that the policy does not have any effect on the probability of continuing to study at the university.<sup>26</sup> The other columns show the effect on the choice of degree and on the quality of the institution attended. Students taking triple science are more likely to attend institutions belonging to the Russell group.<sup>27</sup> Moreover studying more science in secondary school also increases the probability of graduating on time in STEM degrees.<sup>28</sup> This is extremely relevant given the large debate that is taking place in many countries, the US in particular, about the low persistence of students in scientific fields (Arcidiacono et al., 2016; Stinebrickner and Stinebrickner, 2014).

Table 4 shows that the effect masks substantial gender heterogeneity<sup>29</sup>: while girls are affected by the policy- for instance they are induced to take more medicine or biological sciences, the effect on pure STEM degrees is entirely driven by boys. Some studies claims girls may shy away from STEM degrees because of fear for competition or lack of confidence about their ability (Buser et al., 2014; Niederle and Vesterlund, 2010). One may therefore think that increasing preparation and fostering scientific culture in secondary schools may shrink the gender gap in STEM degrees. My results suggest instead that strengthening the science curriculum at age 14 is not helpful. It may increase the share of girls taking science at age 14 and age 16, but it does not affect the share of girls choosing STEM subjects at the university. This is in line with the

---

<sup>26</sup>Note that even if the magnitude of the coefficient is similar to the other coefficients, the baseline in this case is much larger: the average is 36% in this case.

<sup>27</sup>The Russell group represents 24 leading UK universities in terms of research and teaching.

<sup>28</sup>The results on university outcomes are estimated on students taking the final KS4 exam in the years 2005-2007 only, otherwise there is no information on whether the students graduated from university.

<sup>29</sup>As shown in Table A1 of the Appendix, there are other interesting source of heterogeneity. The group mostly affected by the policy are the middle-high ability students. The very high ability students would probably be very well prepared in any case and are less likely to be at the margin, the low ability students are instead less likely to be affected by the policy at all. Moreover the effect on science at age 16 is slightly stronger for low SES students, the effect on university outcomes is instead more difficult to estimate with enough precision for low SES students because of the small share (20%) of low SES students attending university.

findings of some recent studies (Gemici and Wiswall, 2014; Zafar, 2013) which shows that differences in preferences are the main driver behind the gender gap in college degrees; and preferences are difficult to be shaped by secondary school courses. Moreover, my results are complementary to what is found in Joensen and Nielsent (2016) for Denmark. Differently from the results obtained in Table 4, Joensen and Nielsent (2016) estimate very positive effects both for boys and for girls on the probability of choosing technical subjects at the university for students taking advanced math in secondary school. A first reason behind this difference in the results may be that Joensen and Nielsent (2016) find a rather large effect on the probability of attending university as well, because their instrument affects slightly lower ability students than in my case. Their effect may therefore be the combination of changes in the pool of students attending university and changes in the willingness to choose STEM subjects, conditional on going to university; my effect instead comes exclusively from the second component. A second reason is related to differences in the type of compliers. As also pointed out by Joensen and Nielsent (2016) and extensively addressed for the regressions on earnings, the policy they analyze affects girls much more than boys and compliers for the two groups of students are likely to be very different. This makes the coefficients of the IV difficult to compare across genders. As I will address more extensively in Subsection 4.1, my instrument affects boys and girls in a very similar way.

Tables A4 and A5 explore moreover the extent and the presence of subjects complementarity and substitutability. If one takes more science at age 14, which other (complement) subjects is she more likely to take and, more importantly, from which (substitute) subjects does she opt out? Table A3 in the Appendix shows the coefficients and standard errors obtained from estimating the second stage of equation 4 using each time a different KS4 subject as dependent variable. Tables A4 and A5 report the same type of estimates but they refer to KS5 subjects and university degrees. Students who take triple science at KS4 tend to drop more vocational subjects, foreign languages like German and other core subjects like history. In terms of KS5 courses, taking triple science induces students to choose more natural science subjects and math, and to drop more vocational subjects, like media and accounting. Finally, triple science increases the probability of choosing scientific subjects at the university, like physics, engineering and medicine, but also non scientific but more challenging subjects, like classical languages. It decreases, instead, the probability of enrolling in law and architecture. The effect are different for boys and girls, especially for what concerns university degrees.

It is difficult to draw general conclusions from the coefficients of Tables A3, A4 and A5: anecdotal evidence may suggest that a vocational course in music is very different from an advanced course in science at age 14 but, to evaluate each subject according to

some objective criteria, Table 5 uses a more formal procedure. I define courses along two dimensions: (i) ‘high achievers’ courses, characterized by a high average primary school grade of students choosing them in out-of-sample academic years; (ii) ‘female dominated’ courses, characterized by a high share of girls attending the courses in out-of-sample academic years (2002-2005). Figure 3 describes each subject, along these dimensions. In particular it shows three scatter plots where for each course it is displayed on the x-axis the share of girls usually enrolled in it and on the y-axis the average primary school grade of student attending it. Triple science stands out as the course at KS4 that is attended by the best students, followed by foreign languages, history and geography. With respect to KS5 options, math is the most challenging course, followed by physics, chemistry and foreign languages. For university degrees, medicine, languages and STEM subjects are attended by very good students while education, subjects allied to medicine and art are attended by the worst students on average. The correlation between the ability of students usually attending each course and the share of girls enrolled in those courses is negative. This is surprising, given that on average girls have higher grades than boys in primary school.

Table 5 shows whether students start choosing more ‘high achievers’ courses at age 18 (KS5) and at the university as a consequence of taking advanced science at KS4.<sup>30</sup> Taking advanced science at age 14 induces students to choose more challenging subjects later on. Students taking triple science are induced to choose at age 16 courses usually attended by students whose average grade in primary school is about 0.2 standard deviations higher. The same is true for university degrees, but the magnitude of the effect is smaller. Moreover, for KS5, I disentangle how much of the reported increase is automatically due to the higher probability of choosing natural science subjects and how much to the fact that students choose other (complement) more ‘high achievers’ subjects, different from the three natural sciences. I find that the increase is partly driven by an higher probability of choosing science courses (63%) and partly due to a higher willingness to enroll in other difficult subjects not strictly in the natural science field (37%).<sup>31</sup>

The other columns look at the sample of boys and girls separately. The first row shows that girls who take triple science are induced to choose more challenging subjects (i.e. more ‘high achievers’ subjects) in about the same proportion as boys, the second row shows that they still opt for female-dominated subjects (like medicine for instance). This is an interesting result and confirms what found in table 4: while at age 16 girls taking triple science still opt for more male-dominated subjects (physics or math for instance - even if to a lower extent than for boys), strengthening the science preparation

---

<sup>30</sup>To obtain these results I multiply the coefficients displayed in Tables A3, A4 and A5 by the numbers displayed in Figure 3 and I sum the series. Standard errors are computed through the Delta method.

<sup>31</sup>This result is available upon request.

in secondary school does not have any effect on the likelihood that girls opt for STEM (male-dominated) subjects at the university. This suggests that once the subject choice is actually related to the characteristics of their future jobs, girls still prefer the most female-dominated degrees.

## 4.1 Compliers' characterization

This Section analyses who decides to take triple science, when the school offers it. This helps to understand how students make decisions about which subject to take at age 14 and whether the heterogeneity in the  $\beta_1$  coefficient, especially along the gender dimension, is actually driven by differences in the treatment effect or by differences in compliance across genders. Even if teachers in England usually make recommendations about which field courses to choose, the actual choice of whether to take triple science or not is a free decision made by students.<sup>32</sup>

Pupils will choose to take triple science if their expected utility when  $TS = 1$  is higher than their expected utility when  $TS = 0$ . This may happen because triple science reduces their costs (or their perception of the cost) of graduating in certain degrees or of graduating at all or because triple science directly increases their productivity, and therefore wage. The contribution in terms of utility of taking triple science with respect to the second best option, will not be the same for all students: those already very good in science or with very strong preferences towards other subjects may not find it as beneficial to take triple science.<sup>33</sup> This means that the likelihood of taking triple science will not be the same for everybody: it will depend on preferences, on innate ability and on perceptions towards their ability.

The first row of Table 6 shows results from the first stage regression. Being unexpectedly exposed to the offer of taking triple science increases students' probability of enrolling in it by 15 percentage points. The F statistics is around 2800.

Table 6 then characterizes compliers for the entire population and for boys and girls separately (columns 2 and 3, respectively). I obtain information on compliers' characteristics looking at the first stage for several subgroups of the population. For instance the ratio between the instrument's coefficient of the first stage estimated on the sample of females only (0.149) and the coefficient of the first stage estimated on the entire sample

---

<sup>32</sup>One caveat should be considered when interpreting the results: sometimes supply of triple science is constrained since classes in England cannot be larger than 30. Since schools mainly prioritize based on previous science and math scores, any differences in the probability of taking triple science based on previous test scores may not be driven by students' willingness to take triple science, but by schools admission rules.

<sup>33</sup>Unless triple science has a positive effect also in reducing the cost of taking exams in other subjects, for instance through changes in self confidence.

(0.163) represents the relative likelihood that a complier is female.<sup>34</sup> The Table shows that compliers are more likely to be very good students in primary school: the relative likelihood a complier is in the top 20th percentile of test scores in primary school is more than two. Moreover compliers tend to be high income students and, interestingly, there does not seem to be any particular gender difference in compliance. The second and the third columns compare compliers for the subgroups of girls and boys respectively and show that compliers' characteristics are very similar between these two groups.

## 4.2 Heterogeneity

This section evaluates the heterogeneity of the effect of strengthening the science curriculum in secondary school for different subgroups of the population. In particular, I analyse the heterogeneity of the effect by gender, socio-economic status and previous grades in science.

The first panel of Table ?? looks at whether attending more science classes at secondary school has a different effect depending on students science grades in primary school. In particular the Table looks at the probability of enrolling in STEM degrees and of persisting in these studies. The group mostly affected by the policy are the middle-high ability students. The very high ability students would probably be very well prepared in any case and are less likely to be at the margin, the low ability students are instead less likely to be affected by the policy at all.

The second panel analyses heterogeneity by socio-economic status (SES).<sup>35</sup> The effect on science at age 16 is slightly stronger for low SES students, the effect on university outcomes is instead more difficult to estimate with enough precision for low SES students because of the small share (20%) of low SES students attending university.

The third panel analyses gender heterogeneity. The effect is positive for both boys and girls, but the effect on STEM degrees is entirely driven by boys. Still, girls are affected by the policy: they are induced to enroll in more scientific degrees, but tend to choose more female-dominated science degrees like medicine instead of engineering.

Table ?? summarizes the results on gender, following the same method adopted in Table 5 but it looks at the sample of boys and girls separately. While the first row shows that girls tend to choose more challenging subjects (i.e. more 'high achievers' subjects) in about the same proportion as boys, the second row shows that they still opt for female-dominated

---

<sup>34</sup>First stages in this case do not include any control a part from year and school fixed effects. This does not affect the effect of interest because controls are not correlated with the instrument.

<sup>35</sup>Two separate proxies of socio economic status are available in the NPD: Free School Meal eligibility (FSM), a dichotomous variable indicating whether the student is eligible for or in receipt of FSM (approximately 14% of students) and Income Deprivation Affecting Children Index (IDACI), that indicates the proportion of children under age 16 in the local area where the student lives who are living in low income households (the median is 16% of low income households in the area). Table ?? uses only FSM, but results are consistent using the other proxy.

subjects (like medicine for instance).

### 4.3 Checks to the identification strategy

As stated in Section 3, the instrument used in the analysis relies on some assumptions.

First, the assumption that the information set of both the treatment and the control groups of students at age 11 is the same and does not include the information on whether the schools not offering triple science when students apply are going to offer it in three years. To check this assumption, I test first the extreme assumption that, even when parents or students know the school is offering triple science when applying, they do not select schools accordingly. This may be consistent with the notion that students cannot freely choose their schools because schools, when oversubscribed, have to select students based on geographical distance. I estimate the following equation, including all schools in the sample (both offering and not offering triple science when student  $i$  applies):

$$W_{ist} = \alpha_1 Z_{st}^{11} + \alpha_2 Z_{st} + \alpha_3 X_{ist} + \xi_s + \xi_t + \eta_{ist} \quad (5)$$

where  $Z_{st}^{11}$  is a dummy equal to 1 if school  $s$  attended by student  $i$  in cohort  $t$  offered triple science when students were 11 and they chose their secondary school;  $Z_{st}$  is my usual instrumental variable and  $W_{ist}$  are several outcomes (like the dummy for whether student  $i$  chooses a STEM degree or whether she graduates in it) or pre-determined characteristics (like the average science grade in secondary school, her gender etc). Table 7 shows the results with (panel 1) and without (panel 2) school specific trends. The coefficient  $\alpha_1$  is not significant for most variables and in any case is usually extremely small. This means that students applying to schools already offering triple science or not offering it appear very similar- at least in terms of observable characteristics.

Second, the assumption that schools decide when to start offering triple science not based on the quality of the current cohort attending the school and not because the school is already on an increasing trend. Table 8 provides evidence that, when using my identification strategy, the timing of the introduction of the triple science option is not correlated with (observable) characteristics of current students in the school. The Table runs a set of placebo tests, where I estimate the reduced form of equation 4 (without controlling for  $X_{ist}$ ) and where the dependent variable is a pre-determined characteristic, the grade in the science course in primary school. The triple science dummy ( $TS$ ) in this case should not be significant, because the instrument should not be correlated with the grade at KS2, unless my specification does not take full care of selection. The Table has the same structure of Table 2 and it shows how different identification strategies may fail to address selection. Column 1 shows results from a simple OLS regression, column 2 adds

school fixed effects, column 3 replicates the specification used by [Altonji \(1995\)](#) and uses as instrument the share of students taking triple science in school  $s$  and year  $t$ , Column 4 uses my instrument but does not include schools fixed effects.<sup>36</sup> Column 5 includes also school fixed effects. Reassuringly, the effect in this case is 0. Finally column 6 adds school specific time trends, and the coefficient is again 0. Table [A2](#) in the Appendix shows results from a set of other balancing tests obtained estimating the same specifications as in columns 5 and 6 for a bunch of other predetermined observable characteristics. All balancing tests show that the treatment is not correlated with observable characteristics of the current students in the school.

Moreover, I check whether, before school  $s$  started to offer triple science, the trend was parallel to that of all other schools still not offering triple science. I augment my reduced form regression with leads and lags of the instrument (following [Autor \(2003\)](#)):

$$y_{ist} = \sum_{t=0}^m \gamma_{\tau-t} z_{s(\tau-t)} + \sum_{t=0}^q \gamma_{\tau+t} z_{s(\tau+t)} + \zeta_t + \zeta_s + u_{ist} \quad (6)$$

where  $z_{st}$  is my instrument,  $\tau$  is the year school  $s$  starts offering triple science,  $\zeta_s$  and  $\zeta_t$  are the usual school and year fixed effects and  $u_{ist}$  is the error term. I then check for the presence of parallel pre-treatment trends by evaluating whether all coefficients  $\gamma_{\tau-t}$  are close to 0, for every  $\tau$ . [Figure 4](#) shows that the trends are parallel before the introduction of the advanced science course and there is a jump in the outcomes and in the treatment correspondingly exactly to the year of the introduction of the new course.<sup>37</sup> This confirms the results obtained in [Table 7](#) and [8](#).

Another possible concern is that, once a school sets up all arrangements in terms of teaching qualifications and staff in order to offer triple science, it may start to offer more science courses at KS5 as well. In England about 60% of the schools offer both KS4 (age 14) and KS5 (age 16) exams. This would imply that part of the effect I find may be purely mechanical: students take more KS5 science courses because the set of options changes also at KS5. I address this concern in [Table 9](#). Columns 1 and 2 look at how the probability of offering science at KS5 evolves over time and whether it corresponds exactly to the cohort when the school starts offering triple science at KS4. The correlation is 0. Columns 3 and 4 look at whether the effect of studying triple science on the probability of choosing science at KS5 is larger for schools offering both KS4 and KS5 courses than for schools offering KS4 courses only. The effect is identical. If part of the effect I find in my results was mechanical, it would be stronger for schools offering both KS4 and KS5

<sup>36</sup>This column partly replicates, even if in a very different context, [Joensen and Nielsent \(2016\)](#)

<sup>37</sup>I also estimated the same graphs but using predetermined characteristics as dependent variables: in this case there is no jump at year 0, nor at year -3, that correspond to the time when students know, when applying, that the school offers triple science. These results are available upon request.



exams.

Moreover, one may worry that taking triple science could potentially directly affect the possibility of being admitted to STEM degrees at the university. However, while universities often require some KS5 subjects in order to admit students to certain degrees, in no case they require specific KS4 subjects. For instance, in 2013, a KS5 exam in math was required in 13% of the cases (i.e. of degree-university combinations) and at least one KS5 exam in science was required in 12% of the cases. In no case<sup>38</sup>, in 2013, there was a specific requirement for age 14 (KS4) subjects.

Finally, it may be that the simple fact of having the possibility of being enrolled in advanced science but having been excluded, for example because the class was oversubscribed and schools had to select students, may generate a direct effect on some students and may therefore violate the exclusion restriction assumption. This is impossible to test. Table A6 however exploits some of the institutional features of English school system to evaluate how problematic this may be. Figure 5 plots the distribution of the size of triple science courses in each school. From the Figure it is clear that class size bunches at multiples of 30. There is a discontinuity both corresponding to 30 students and corresponding to 60 students. Since class size in England is required to be lower than 30, this Figure suggests that in some cases the triple science course was oversubscribed, and schools had to select students. Unfortunately the exact admission rule is different for each school and is not publicly available. Table A6 exploits this feature of the system and runs the main specification (using equation 4) on the sample of schools where the triple science course was very likely not to be oversubscribed, because the number of enrolled students was not close to the maximum.<sup>39</sup> The results of this exercise are very similar to the main ones.

## 5 An alternative identification strategy

This Section proposes an alternative identification strategy that addresses the concern that, even if there is no evidence that schools decide when to offer triple science depending on observable characteristics of their current cohort, it may still be that unobservable characteristics matter. This is impossible to test. My second instrument however is not subject to this last concern because it exploits variation in available courses that existed even before current students started to attend their secondary schools. This excludes the possibility that the choice of offering triple science depends on specific characteristics to

---

<sup>38</sup>Data are taken from <http://www.thecompleteuniversityguide.co.uk/courses/search>

<sup>39</sup>Those schools where the number of students enrolled in the triple science classes was not between 28 and 32 or between 58 and 62.

the particular cohort in the school.

This instrument compares students living in the same neighbourhood but who are more or less likely to enroll in schools offering triple science, because of exogenous changes in schools' catchment areas.

I exploit the fact that when schools in England are oversubscribed, usually prioritize students based on geographical distance.<sup>40</sup> Therefore, in each year there will be a maximum distance between the school and the students' addresses above which students will not be accepted. I build my instrument in two steps: first, I compute the school catchment areas for each year as the area delimited by the circle whose centre is the school and ray is the maximum observed home-to-school distance,<sup>41</sup> and I define the set of 'reachable' schools for each student. Second, I compute the share of 'reachable' schools that offered triple science when student  $i$  applied. Student address refers to the lower level output area (LLOA)<sup>42</sup> where student  $i$  used to live at age 10. The instrument used in this section of the analysis counts how many schools, out of the set of schools reachable by students  $i$  in year  $t$ , offered triple science when  $i$  applied to secondary school (in this case the instrument in year  $t - 1$  was 1 and in year  $t$  was 0.5). The instrument varies both because of (unpredictable) variations in schools catchment areas and because of the overall increase in the number of schools offering triple science within the catchment area.

I estimate the following first stage equation:

$$TS_{ipt} = \theta_1 Z_{pt}^2 + \theta_3 X_{ipt} + \theta_p + \theta_t + v_{ipt} \quad (7)$$

where  $TS_{ipt}$  is the usual dummy indicating whether student  $i$  in year  $t$ , who used to live in neighbourhood  $p$  when she was 10 year old, takes triple science and 0 otherwise;  $Z_{pt}^2$  is the share of reachable schools offering triple science for students of cohort  $t$  residing in neighbourhood  $p$ ;  $X_{ipt}$  are individual controls and  $\theta_t$  and  $\theta_p$  are cohort and neighbourhood fixed effects respectively;  $v_{ipt}$  is the error term.

This instrument compares students attending schools that offer triple science with students attending schools not offering it, i.e. it uses across school-within neighbourhood-over time variation (instead of within school-over time variation). Offering triple science is however likely to be related to other school characteristics, like school quality, that may directly affect the choices of degree at the university. This issue may be more relevant

---

<sup>40</sup>With some exceptions for students with siblings attending the same school or for students with special education needs. Since I do not have the full set of information necessary to simulate the exact admission formula for each school, I can't adopt an RDD strategy.

<sup>41</sup>In order to exclude exceptions I eliminated outliers (the distances higher than the 5th percentile for every school).

<sup>42</sup>In total there are more than 30,000 LLOAs in England and Wales and each LLOA contains on average 1500 households.

when we use across school rather than within school variation because differences in quality across schools are likely to be much more sizable than differences within schools over time. I address this concern by including as control the average quality level of the set ‘reachable’ schools in each catchment area over time.

Table 10 shows the results obtained from my second identification strategy.<sup>43</sup> The first three columns refer to the probability of choosing a natural science subjects at Key Stage 5 (age 18), the last three columns refer to the probability of attending a STEM degree at the university.<sup>44</sup> The first and the fourth columns do not include neighbourhood fixed effects, but control for the lagged value of my instrument: they compare neighbourhoods which had the same share of reachable schools offering the triple science course the previous year and they exploit variation between  $t$  and  $t - 1$ . All other columns include neighbourhood fixed effects.

This instrument compares students living in the same neighbourhood but attending different schools which offer or do not offer triple science. However, the probability of offering triple science is likely to be related to other school characteristics, like school quality, that may directly affect the choices of degrees at the university. Since the variation in school quality may be much larger when using across school rather than within school over time variation, like with the previous instrument, in Columns 3 and 6 I include the average quality of the set of reachable schools in year  $t$  as a control. I proxy school quality using the school value added in the out of sample years (2002-2005).

The results confirm the robustness of the first identification strategy: the estimated effects are positive and significant and the effects on STEM degrees are stronger for boys than for girls<sup>45</sup>. The estimates obtained through this strategy are however slightly larger, this may be related to the different type of variation, and therefore of compliers, exploited. While compliers for the first instrument are all individuals who take triple science because their school unexpectedly starts to offer it, which also include very good students who happened to be enrolled in a school not offering triple science; compliers in the second instruments are students who take triple science because, thanks to a larger supply of triple science in the set of reachable schools in their neighbourhood, they manage to enroll in a school offering it. In this second case, very good students would probably have enrolled in a school offering triple science in any case. This suggests compliers for the second strategy exclude the extremely high ability students. Since, as shown in Table A1 in the Appendix, those mostly affected by the policy are middle-high ability students,

---

<sup>43</sup> Since there is no information on postcode in primary school for students who finished secondary school in the years before 2007, this section only refers to the years 2007-2010. For these cohorts, however, I have information on whether they graduated only for the students who took KS4 exams in the year 2007, so I only analyze effects on enrollment and on KS5 outcomes.

<sup>44</sup>The effect on the probability of attending university is 0, as for the previous instrument.

<sup>45</sup>results available upon request

this may explain the larger effect found in Table 10.

## 6 Potential alternative mechanisms

This Section explores the mechanisms that may generate the effect found in Section 4 and explores whether the effect obtained is actually generated by changes in curriculum or, since the treatment has multiple components, it is also driven by changes in the peer composition of the courses attended or in the type of teachers in the school.

First, I analyze the peers channel. In particular, I use the following measure of peer quality in science ( $Q_{ist}$ ) for student  $i$ , attending school  $s$  in year  $t$  who takes science courses  $D_{ist}$ :

$$Q_{ist} = \bar{X}_{(-i)st}^D \quad (8)$$

where  $\bar{X}_{(-i)st}^D$  is the average science grade in primary school of students taking age 14 science course  $D^{46}$ , in school  $s$  in year  $t$  (excluding  $i$ ).

The first panel of Figure 6 shows how peers' composition in the science course taken at age 14 changes for schools offering triple science or not. The dashed line plots the density of  $Q_{ist}$  in the age 14 science course for students attending schools not offering triple science. The solid line refers instead to schools offering triple science. The figure shows that when schools offer triple science there is a concentration of very high ability students able to attend the science class with peers of much higher quality than before. Column 1 of Table 11 confirms this finding: it shows how peers' quality in science courses changes after the school starts offering the advanced science course, depending on students' primary school grade in science. The quality of peers in the science class decreases for lower ability students and increases quite extensively for higher ability students.

To control for this dimension and check whether the effect found in Table 3 comes mostly from changes in the peer composition or from changes in the curriculum, I control for peer quality in equation 4. Since students self-select into different types of science course at age 14, peers' quality may be endogenous. I therefore instrument peer quality by using within-school over-time changes in peers' composition (following Hoxby (2000)). In particular, I use the fact that classes in England cannot be larger than 30 (as shown in Figure 5).<sup>47</sup> I therefore predict, based on predetermined characteristics like previous test scores and demographics,<sup>48</sup> the probability of being enrolled into triple science and I

<sup>46</sup>Since there is no information about the exact class but only about the type of science course, I use the average grade in primary school of students taking the same course.

<sup>47</sup>While for primary schools this requirement is compulsory, it is just recommended for secondary school.

<sup>48</sup>In particular, KS2 and KS3 science grades (both teacher assessed and from standardized exams) , gender, Free School Meal Eligibility.

take the average science grade in primary school of the 30 or 60 students (depending on the number of triple science classes offered) with the highest probability of being enrolled into triple science. I then exploit within school over time variation in the average quality of these students and of all other students in school  $s$  and year  $t$ , allowing the effect to be different depending on whether the school offers (unexpectedly) triple science or not. My first stage equation is:

$$Q_{ist} = \theta_1 z_{st} + \theta_2 \widehat{Q}_{st(-i)}^{top30} + \theta_3 \widehat{Q}_{st(-i)}^{others} + \theta_4 \widehat{Q}_{st(-i)}^{top30} * z_{st} + \theta_5 \widehat{Q}_{st(-i)}^{others} * z_{st} + \theta_5 X_{ist} + \theta_s + \theta_t + \eta_{ist} \quad (9)$$

where  $z_{st}$  is the first instrument - the dummy equal to 1 student  $i$  was unexpectedly exposed to the option of choosing triple science-  $\widehat{Q}_{st(-i)}^{top30}$  is the average science grade in primary school of the 30 (or 60) students with the highest predicted probability of being enrolled in triple science and  $\widehat{Q}_{st(-i)}^{others}$  is the average science grade in primary school of all other students;  $\theta_s$  and  $\theta_t$  are school and year fixed effects and  $\eta_{ist}$  is the error term. Panel b of Figure 6 shows how the instrument works. The solid line refers to the average science grade in primary school for students predicted to attend the triple science class, the dashed line refers to all other students.

Table 11 displays the results. Columns 2 to 6 show that the effect of triple science is very similar to what found before, even after controlling for changes in peers' quality. The joint F statistic is 35.

Second, I evaluate whether teachers may change in response to the introduction of the triple science course. Unfortunately, it is not possible in England to link data on individual teachers to administrative data on individual students. In this section I use the yearly number of teachers and of qualified teachers in each school. Table A7 in the Appendix shows that neither the overall number of teachers nor the number of qualified teachers in a school change significantly once the school introduces the triple science option. This suggests that teachers' quality and quantity do not increase as a result of the introduction of the advanced science course.

## 7 Conclusions

This paper uses a reform that increased the probability of taking an advanced science course in English secondary schools for students at the top of the ability distribution to analyze whether secondary school curriculum affects post-16 outcomes, and in particular the probability of enrolling and graduating in a STEM degree. Moreover, by separately investigating the effect on boys and girls, this paper seeks to understand whether strengthening

school preparation in science shrinks the gender gap in enrollment in STEM degrees.

Since the policy I consider affected very high ability students, who would have continued studying in any case, I find that a stronger science curriculum in secondary school has no effect on university enrollment. Still, my estimates suggest that offering more science in secondary school improves educational outcomes in many domains. It induces students to attend higher quality universities and significantly increases the probability of enrolling and, very importantly, of graduating from university with a STEM degree. This effect masks a substantial and interesting gender heterogeneity: at age 14 when exposed to the option of studying more science in secondary school, there is no gender difference in the take-up probability. However, the difference arises later on, at the university, when subject choices are likely to be correlated with occupations and jobs: both boys and girls are induced to take more challenging courses on average, but girls still choose more female-dominated subjects like medicine, instead of engineering and math. This seems to be in line with the recent literature relating preferences towards job attributes to choices of university degrees ([Wiswall and Zafar, 2016](#); [Reuben et al., 2015](#); [Zafar, 2013](#)) that shows that job characteristics play an important role in the choice of subjects at the university, with women and men displaying very different preferences, even if at the very top of the ability distribution.

My findings show that there is a certain degree of persistence between what is studied at secondary school and what is studied at the university. An optimal design of the secondary school curricula may be useful to improve the match between supply and demand of specific skills.

## References

- Joseph Altonji, Peter Arcidiacono, and Arnaud Maurel. *The Analysis of Field Choice in College and Graduate School: Determinants and Wage Effects*, volume 5. Handbook of the Economics of Education, 2015.
- Joseph G. Altonji. The Effects of High School Curriculum on Education and Labor Market Outcomes. *Journal of Human Resources*, 30(3):409–438, 1995.
- Joseph G Altonji, Erica Blom, Costas Meghir, et al. Heterogeneity in Human Capital Investments: High School Curriculum, College Major, and Careers. *Annual Review of Economics*, 4(1):185–223, 2012.
- Peter Arcidiacono. Ability sorting and the returns to college major. *Journal of Econometrics*, 121(1-2):343–375, 2004.
- Peter Arcidiacono, V. Joseph Hotz, and Songman Kang. Modeling College Major Choices using Elicited Measures of Expectations and Counterfactuals. *Journal of Econometrics*, 166(1):3–16, 2012.
- Peter Arcidiacono, Esteban Aucejo, and V. Joseph Hotz. University Differences in the Graduation of Minorities in STEM Fields: Evidence from California. *American Economic Review*, 106, 2016.
- David H. Autor. Outsourcing at Will: The Contribution of Unjust Dismissal Doctrine to the Growth of Employment Outsourcing. *Journal of Labor Economics*, 21(1):1–42, 2003.
- Magali Beffy, Denis Fougere, and Arnaud Maurel. Choosing the Field of Study in Postsecondary Education: Do expected Earnings Matter? *Review of Economics and Statistics*, 94(1):334–347, 2012.
- Julian R. Betts and Heather Rose. The Effect of High School Courses on Earnings. *Review of Economic Studies*, 86(2):497–513, 2004.
- Thomas Buser, Muriel Niederle, and Hessel Oosterbeek. Gender, Competitiveness, and Career Choices. *The Quarterly Journal of Economics*, 129(3):1409–1447, 2014.
- Kalena Cortes, Joshua Goodman, and Takako Nomi. Intensive Math Instruction and Educational Attainment: Long-Run Impacts of Double-Dose Algebra. *Journal of Human Resources*, 50:108–158, 2015.

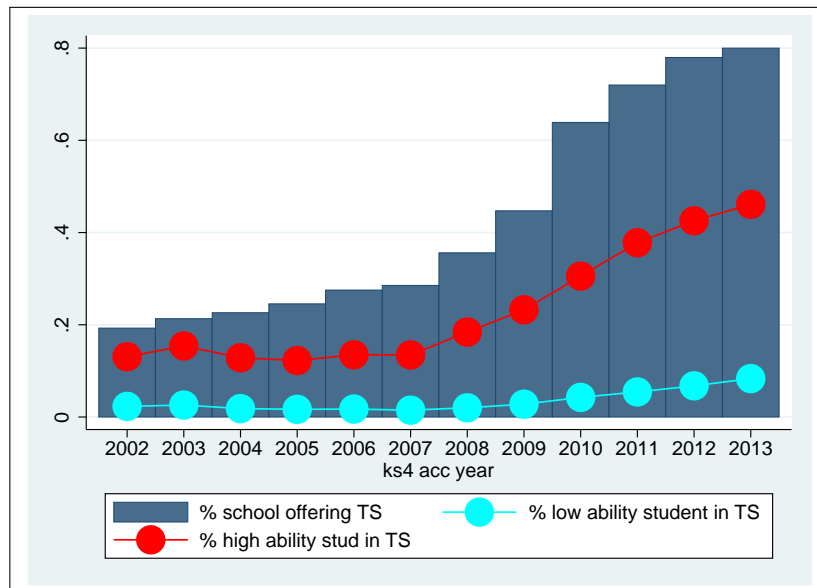
- Glenn Ellison and Ashley Swanson. Heterogeneity in High Math Achievement Across Schools: Evidence from the American Mathematics Competitions. NBER Working Papers 18277, National Bureau of Economic Research, Inc, 2012.
- European Commission. The European Commission: EUROPE 2020 Strategy. Technical report, 2010.
- Ahu Gemici and Matthew Wiswall. Evolution Of Gender Differences In Post-Secondary Human Capital Investments: College Majors. *International Economic Review*, 55: 23–56, 02 2014.
- Claudia Goldin. A Grand Gender Convergence: Its Last Chapter. *American Economic Review*, 104:1091–1119, 2014.
- Joshua Goodman. The Labor of Division: Returns to Compulsory Math Coursework. Technical Report 23063, NBER Working Paper, 2017.
- Katja Gurlitz and Christina Gravert. The Effects of a High School Curriculum Reform on University Enrollment and the Choice of College Major. Working Papers 8983, IZA, 2015.
- Caroline Hoxby. Peer Effects in the Classroom: Learning from Gender and Race Variation. NBER Working Papers 7867, National Bureau of Economic Research, Inc, 2000.
- Information Technology and Innovation Foundation. Refueling the U.S. Innovation Economy: Fresh Approaches to STEM Education. Technical report, 2010.
- Estelle James, A. Nabeel, and J. Conaty. College Quality and Future Earnings: Where Should You Send Your Child to College? *American Economic Review*, 79(2):247–52, 1989.
- Ning Jia. Do Stricter High School Math Requirements Raise College STEM Attainment? Working Papers 8983, Mimeo, 2014.
- Juanna Schroter Joensen and Helena Skyt Nielsen. Is there a Causal Effect of High School Math on Labor Market Outcomes? *Journal of Human Resources*, 44(1):171–198, 2009.
- Juanna Schroter Joensen and Helena Skyt Nielsent. Mathematics and Gender: Heterogeneity in Causes and Consequences. *Economic Journal*, 126:1129–63, 2016.
- Lars J. Kirkeboen, Edwin Leuven, and Magne Mogstad. Field of Study, Earnings and Self-selection. *Quarterly Journal of Economics*, 2016.



- Phillip B Levine and David J Zimmerman. The Benefit of Additional High-School Math and Science Classes for Young Men and Women. *Journal of Business & Economic Statistics*, 13(2):137–49, 1995.
- Muriel Niederle and Lise Vesterlund. Explaining the gender gap in math test scores: The role of competition. *Journal of Economic Perspectives*, 24(2):129–44, June 2010.
- Ernesto Reuben, Matthew Wiswall, and Basit Zafar. Preferences and biases in educational choices and labour market expectations: Shrinking the black box of gender. *The Economic Journal*, 2015.
- Ralph Stinebrickner and Todd R. Stinebrickner. A Major in Science? Initial Beliefs and Final Outcomes for College Major and Dropout. *Review of Economic Studies*, 81(1): 426–472, 2014.
- The President’s Council of Advisor on Science and Technology. Engage to Excel: Producing one Million Additional College Graduates with Degrees in Science, Technology, Engineering, and Mathematics. Report to the President. Technical report, 2012.
- UK Government. Science & Innovation Investment Framework 2004-2014. Technical report, Department of Education, HM treasury, Department of Health, Department of Trade and Industry, 2004.
- UK HM Treasury and BIS. The Plan For Growth. Technical report, 2010.
- Matthew Wiswall and Basit Zafar. Determinants of College Major Choice: Identification using an Information Experiment. *Review of Economic Studies*, 2014.
- Matthew Wiswall and Basit Zafar. Preference for the workplace, human capital, and gender. Technical report, National Bureau of Economic Research, 2016.
- Basit Zafar. College major choice and the gender gap. *Journal of Human Resources*, 48 (3):545–595, 2013.

# Figures

Figure 1: Take up in triple science



**Source:** NPD dataset. The bars represent the share of schools offering triple science; the red dots represent the share of high ability (based on English, math and science primary school grade, top 40 %) students taking triple science and the blue dots show the share of low ability (based on primary school grades, bottom 60 %) students taking triple science, by year.

Figure 2: Timeline of the English educational system

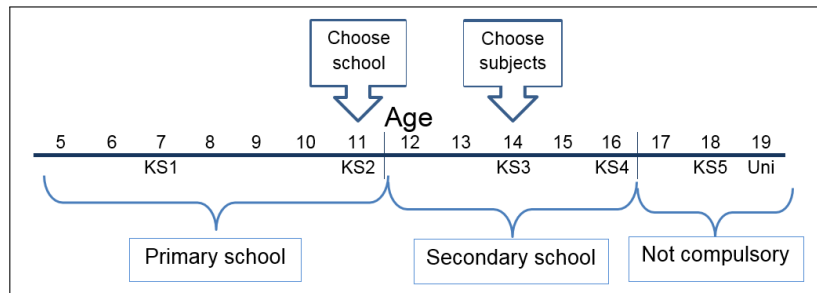
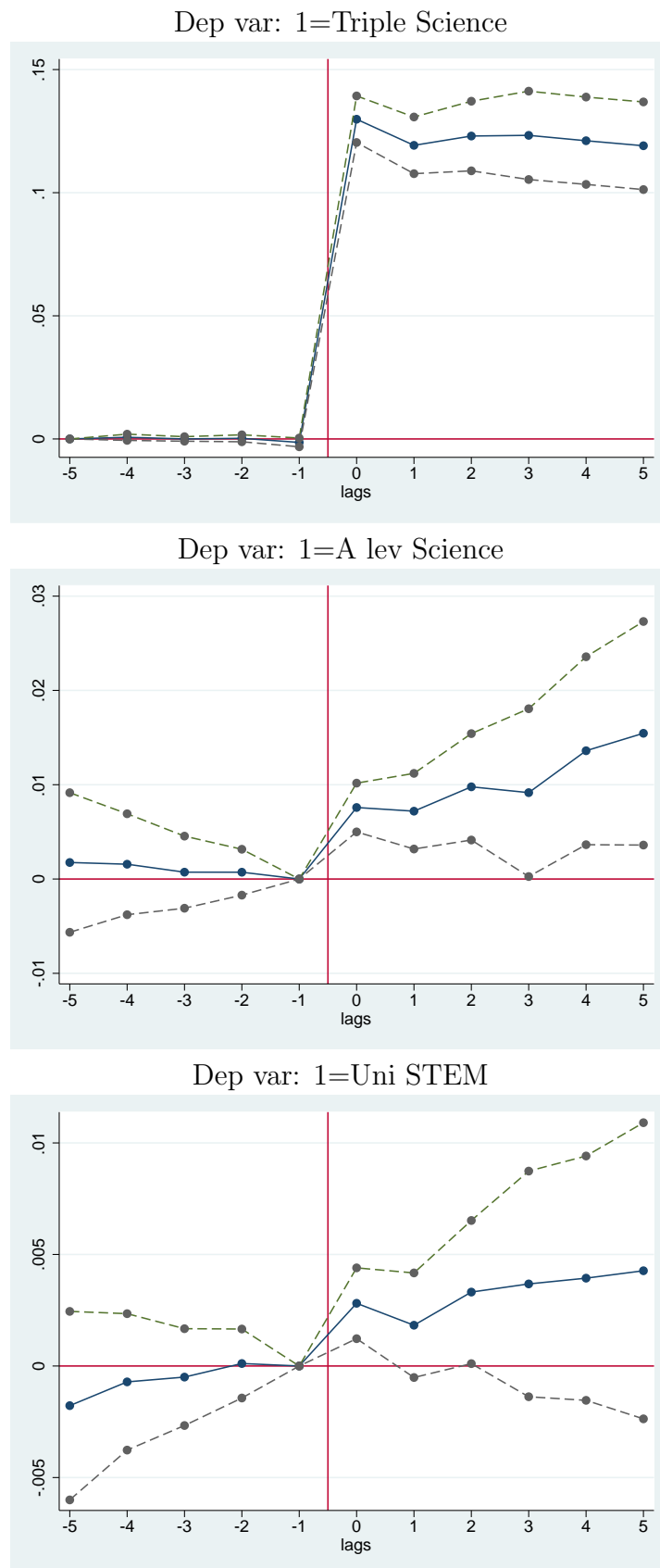


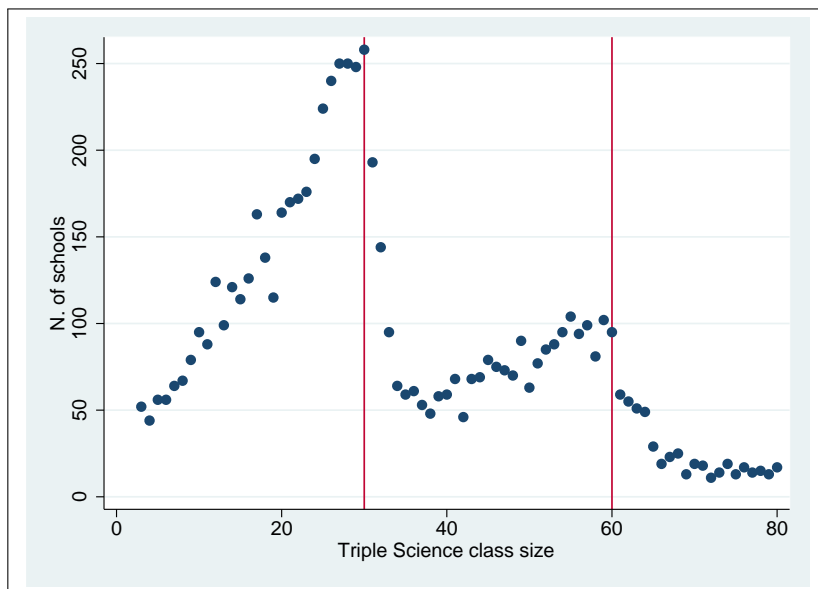


Figure 4: Parallel Trends: Leads and Lags of the instrument



**Source:** NPD dataset. The continuous line represent coefficients, the dashed lines the 5% confidence intervals, obtained from estimating equation 6. Omitted category: one year before the treatment.

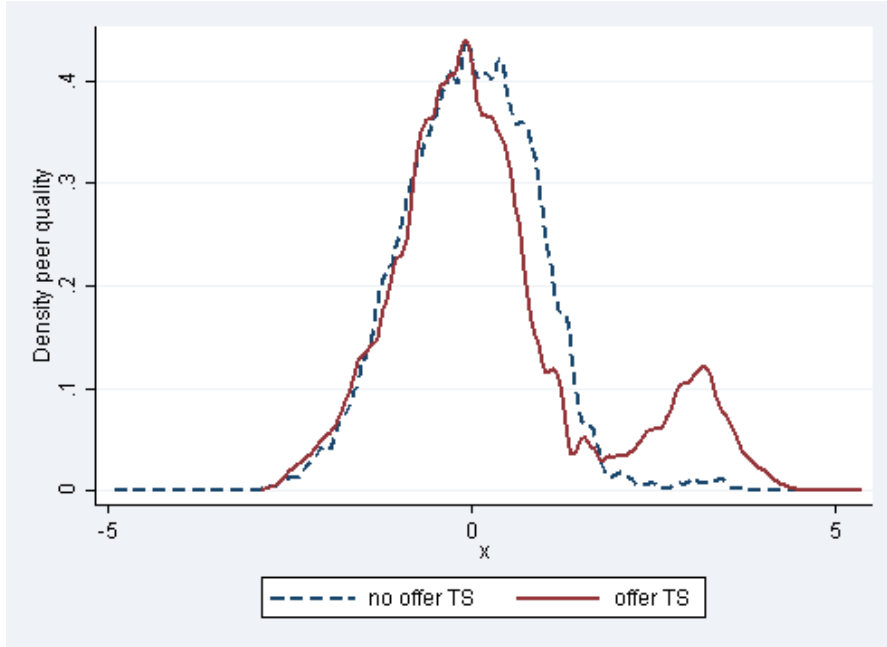
Figure 5: Class size and number of students in triple science



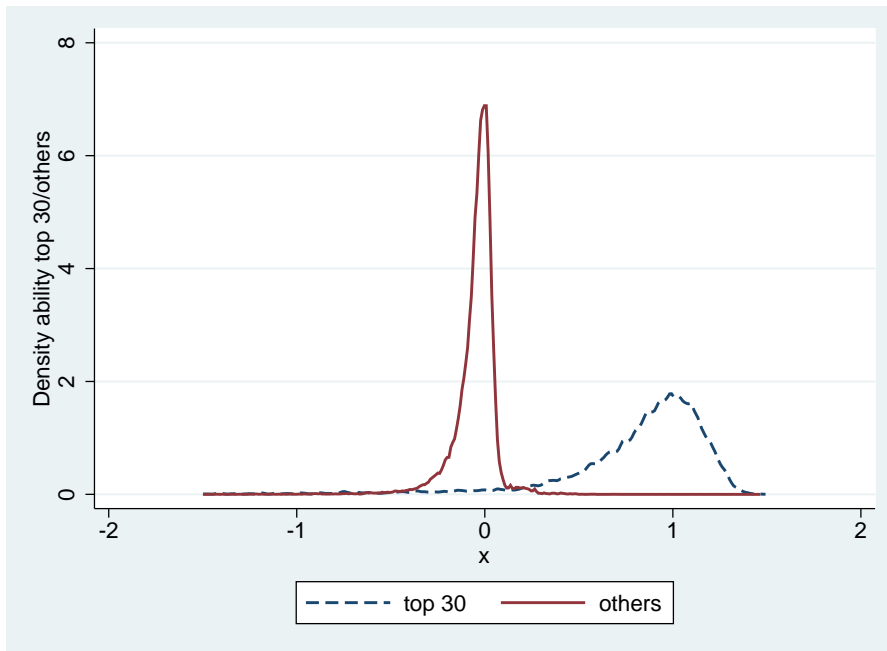
**Source:** NPD dataset. The dots are the number of schools, by triple science class size .

Figure 6: Peers

Actual peers' quality



Instrument



**Source:** NPD dataset. The first panel plots the distribution of science peers' quality, distinguishing whether the school offers triple science or not. The second panel plots the average peers quality for students predicted to take the TS class and students not predicted to take the TS class.

# Tables

Table 1: Summary statistics

Variable	Mean	Std. Dev.
<i>Key Stage 4 (age 14-16)</i>		
offer TS (unexpected)	0.196	0.397
1=Triple Sci	0.076	0.264
1=Double Sci	0.764	0.425
1=Single Sci	0.163	0.369
<i>Key Stage 5 (age 17-19)</i>		
1=KS5 science (if KS5)	0.198	0.282
1=KS5 math (if KS5)	0.142	0.252
<i>University</i>		
1=uni	0.348	0.470
1=STEM <sup>a</sup>	0.126	0.198
1=Russell	0.046	0.211
1=graduate <sup>a</sup>	0.481	0.361
<i>Demographics</i>		
1=female	0.497	0.500
1=FSM eligible <sup>b</sup>	0.144	0.356

The summary statistics reported in the Table refer to the entire sample of students taking their final KS4 exams (at age 16) between 2005 and 2010.

<sup>a</sup> Conditional on going to university.

<sup>b</sup> Free School Meal Eligible.

Table 2: Results for science at age 17 and 19

	OLS [1]	OLS-Fe [2]	Altonji [3]	IV [4]	IV-Fe [5]	IV-Fe tr [6]	IV-Fe [7]
Dep var:	1=KS5 Science						
1=TS	0.334*** (0.005)	0.257*** (0.005)	0.147*** (0.014)	0.072*** (0.010)	0.051*** (0.006)	0.048*** (0.008)	0.054*** (0.006)
1=female		-0.009*** (0.001)	-0.004*** (0.001)	-0.011*** (0.001)	-0.010*** (0.001)	-0.010*** (0.001)	
grade sci age 11		0.020*** (0.000)	0.019*** (0.001)	0.019*** (0.001)	0.021*** (0.001)	0.022*** (0.001)	
N	1690451	1690451	1690451	1690451	1690451	1690451	1690451
Fstat			559372	2234	2065	1742	2066
Dep var:	1=STEM university						
1=TS	0.104*** (0.002)	0.072*** (0.002)	0.039*** (0.005)	0.024*** (0.004)	0.014*** (0.004)	0.012** (0.006)	0.015*** (0.004)
1=female		-0.034*** (0.001)	-0.034*** (0.001)	-0.035*** (0.001)	-0.034*** (0.001)	-0.034*** (0.001)	
grade sci age 11		0.005*** (0.000)	0.005*** (0.000)	0.005*** (0.000)	0.006*** (0.000)	0.006*** (0.000)	
N	1690451	1690451	1690451	1690451	1690451	1690451	1690451
Fstat			559372	2234	2065	1742	2066
School Fe	No	Yes	No	No	Yes	Yes	Yes
School trends	No	No	No	No	No	Yes	No
School contr	No	No	Yes	Yes	No	No	No
Stud contr	No	Yes	Yes	Yes	Yes	Yes	No

Additional controls: year and school fixed effects; student controls: gender, Free School Meal Eligible, Special Education Needs, primary school grade in science, math and english; schools controls: school size. All dependent variables are set equal to 0 if students do not continue studying or if they do not take the considered subjects. Robust standard errors clustered by school in parentheses. \* denotes significance at 10%, \*\* denotes significance at 5%, \*\*\* denotes significance at 1%.



Table 3: Results for other outcomes

	[1]	[2]	[3]	[4]	[5]
<i>Panel 1: KS4 (age 14) outcomes</i>					
	Grades			N. Exams	
Dep var:	KS4 Eng gr <sup>a</sup>	KS4 Math gr <sup>a</sup>	Ks4 science gr	n exams ks4	n exams ks5 <sup>c</sup>
1=TS	0.001 (0.031)	-0.026 (0.028)	-0.065** (0.027)	0.438** (0.210)	-0.021 (0.022)
N	1332413	1339792	1690325	1690451	860615
ymean	0.022	0.021	0.000	10.303	3.416
<i>Panel 2: KS5 (age 16) outcomes</i>					
Dep var:	1=KS 5	1=KS5 math	1=KS5 Bio	1=KS5 Che	1=KS5 Phy
1=TS	-0.009 (0.010)	0.035*** (0.005)	0.037*** (0.004)	0.025*** (0.003)	0.024*** (0.005)
N	1690451	1690451	1690451	1690451	1690451
ymean	0.509	0.056	0.040	0.026	0.065
<i>Panel 3: University outcomes<sup>b</sup></i>					
Dep var:	1=uni	1=grad	1=Russell	1=uni med	1=grad STEM
1=TS	0.044* (0.025)	0.041 (0.025)	0.022* (0.011)	0.013** (0.007)	0.033*** (0.011)
N	966777	966777	966777	966777	966777
ymean	0.318	0.207	0.046	0.019	0.034

Additional controls: year and school fixed effects; student controls: gender, Free School Meal Eligible, Special Education Needs, primary school grade in science, math and english; schools controls: school size. All dependent variables are set equal to 0 if students do not continue studying or if they do not take the considered subjects. Robust standard errors clustered by school in parentheses. \* denotes significance at 10%, \*\* denotes significance at 5%, \*\*\* denotes significance at 1%.

<sup>a</sup> Grades go from 0 to 7, but are standardized to have mean 0 and standard deviation 1.

<sup>b</sup> The results on university outcomes use only the 2005-2008 sample because otherwise there will be no information on the graduation outcomes.

Table 4: Gender Heterogeneity

Dep var:	<i>age 16</i>		<i>uni- age 19</i>		<i>uni- age 22</i>	
	1=KS5 sci [1]	1=Russell [2]	1=STEM [3]	1=medicine [4]	1=grad [5]	1=grad STEM [6]
Girls						
1=TS	0.047*** (0.008)	0.027 (0.021)	0.003 (0.015)	0.023** (0.009)	0.049 (0.040)	0.015 (0.013)
N	849149	486068	486068	486068	486068	486068
ymean	0.080	0.053	0.020	0.030	0.239	0.019
Boys						
1=TS	0.053*** (0.007)	0.018 (0.013)	0.037** (0.017)	0.005 (0.006)	0.033 (0.029)	0.045*** (0.016)
N	841234	480646	480646	480646	480646	480646
ymean	0.088	0.040	0.054	0.008	0.174	0.049

Additional controls: year and school fixed effects; student controls: gender, Free School Meal Eligible, Special Education Needs, primary school grade in science, math and English; schools controls: school size. All dependent variables are set equal to 0 if students do not continue studying or if they do not take the considered subjects. Robust standard errors clustered by school in parentheses. \* denotes significance at 10%, \*\* denotes significance at 5%, \*\*\* denotes significance at 1%.

Table 5: Summarizing effects on other subjects

	$\Delta$ age 16 (ks5) courses			$\Delta$ uni major		
	All	Girls	Boys	All	Girls	Boys
High achievers	0.197*** (0.019)	0.168*** (0.028)	0.220*** (0.023)	0.022*** (0.007)	0.021** (0.011)	0.028*** (0.008)
Female-dominated	-0.042*** (0.018)	-0.016 (0.027)	-0.058*** (0.020)	-0.007 (0.008)	0.014 (0.011)	-0.023** (0.010)

The coefficients are computed as  $\sum_j \beta_j q_j$  where  $j$  indicates subjects,  $\beta_j$  is the subject specific coefficient estimated in Tables A4 and A5 and  $q_j$  is either 'high achievers' (the average primary school grade of taking the course  $j$  in out of sample academic years (2002-2005), standardized to have mean 0 and standard deviation 1) or 'female dominated' (the share of girls attending course  $j$  in out of sample academic years). Standard errors are computed through the delta method.

Table 6: Characterizing compliers

Sample	Everybody [1]	Only Girls [2]	Only Boys [3]
<i>Panel 1: Entire Sample</i>			
$Z_{st}$	0.175*** (0.004)	0.161*** (0.005)	0.188*** (0.005)
N	1690451	849184	841267
<i>Panel 2: Quintiles science grade in primary school</i>			
	subgroup: 1st quintile av. primary school grade		
$Z_{st}$	0.009*** (0.001)	0.008*** (0.001)	0.009*** (0.001)
N	339951	174093	165858
Ratio wrt tot FS	0.051	0.050	0.048
	subgroup: 2nd quintile av. primary school grade		
$Z_{st}$	0.038*** (0.001)	0.035*** (0.002)	0.041*** (0.002)
N	341063	171845	169218
Ratio wrt tot FS	0.217	0.217	0.218
	subgroup: 3rd quintile av. primary school grade		
$Z_{st}$	0.099*** (0.003)	0.092*** (0.003)	0.105*** (0.004)
N	336767	168450	168317
Ratio wrt tot FS	0.566	0.571	0.559
	subgroup: 4th quintile av. primary school grade		
$Z_{st}$	0.222*** (0.005)	0.208*** (0.006)	0.234*** (0.006)
N	344551	171725	172826
Ratio wrt tot FS	1.269	1.292	1.245
	subgroup: 5th quintile av. primary school grade		
$Z_{st}$	0.449*** (0.009)	0.417*** (0.011)	0.479*** (0.010)
N	328119	163071	165048
Ratio wrt tot FS	2.566	2.590	2.548
<i>Panel 3: Socio-Economic Status</i>			
	subgroup: Low SES students (yes FSM <sup>a</sup> )		
$Z_{st}$	0.084*** (0.002)	0.077*** (0.003)	0.092*** (0.003)
N	223375	114446	108929
Ratio wrt tot FS	0.480	0.478	0.489

The Table reports results from the first stage for different subgroups of the population. Dependent variable: a dummy equal to 1 if the student takes triple science. Additional controls: year and school fixed effects. Robust standard errors clustered by school in parentheses. \* denotes significance at 10%, \*\* denotes significance at 5%, \*\*\* denotes significance at 1%.

<sup>a</sup> Free School Meal Eligible.

Table 7: Selection

	av KS2 gr <sup>a</sup> [1]	sci KS2 gr <sup>b</sup> [2]	1=FSM [3]	1=KS5 sci [4]	1=uni [5]	1=STEM [6]	1=grad STEM [7]
<i>Without school specific trends</i>							
$Z_{st}^{11}$	-0.005 (0.005)	-0.008 (0.006)	0.002 (0.002)	0.005*** (0.001)	-0.002 (0.004)	0.001 (0.002)	0.001 (0.002)
N	2882341	2882341	2882341	2882341	1468169	1468169	1468169
School fe	Yes	Yes	Yes	Yes	Yes	Yes	Yes
School trend	No	No	No	No	No	No	No
<i>With school specific trends</i>							
$Z_{st}^{11}$	0.002 (0.006)	0.002 (0.002)	0.007** (0.003)	0.004** (0.002)	-0.003 (0.005)	0.001 (0.002)	0.001 (0.002)
N	2285735	2285735	2285735	2285735	1309004	1309004	1309004
School fe	Yes	Yes	Yes	Yes	Yes	Yes	Yes
School trend	Yes	Yes	Yes	Yes	Yes	Yes	Yes

Additional controls years dummies, school fixed effects. Robust standard errors clustered by school in parentheses. The dependent variables in column 4, 5 and 7 are set equal to 0 if students do not continue studying or if they do not take that subject. \* denotes significance at 10%, \*\* denotes significance at 5%, \*\*\* denotes significance at 1%.

<sup>a</sup> average grade in English, math and science.

<sup>b</sup> grade in science.

Table 8: Balancing Test

	OLS [1]	OLS-Fe [2]	Altonji [3]	IV [4]	IV-Fe [5]	IV-Fe tr [6]
Dep var:	1=Average Grade age 11 <sup>a</sup>					
1=TS	0.927*** (0.013)	0.788*** (0.015)	0.802*** (0.054)	0.363*** (0.052)	0.042 (0.026)	0.045 (0.034)
mfemale				0.232*** (0.053)		
mfsm				-1.545*** (0.051)		
N	1337202	1337202	1337202	1337202	1337202	1337202
School Fe	No	Yes	No	No	Yes	Yes
School time trends	No	No	No	No	No	Yes

Additional controls: years dummies. Robust standard errors clustered by school in parentheses.

\* denotes significance at 10%, \*\* denotes significance at 5%, \*\*\* denotes significance at 1%.

<sup>a</sup> Average grade in the KS4 exams in English, math and science.

Table 9: Robustness: offer age 16 (KS5) Science

Dep var:	Sch level regr (offer)		Stud in schools wo sixth form	
	1=Offer KS5 Science [1]	1=offer KS5 Math [2]	All schools Dep var: 1=KS5 [3]	only offer KS4 Science [4]
$Z_{st}^{11}$	0.002 (0.004)	-0.000 (0.004)		
1=TS			0.050*** (0.006)	0.053*** (0.009)
N	5294	5294	1690451	751721
ymean	0.477	0.467	0.084	0.060

Column 1 and 2 are run at the school-year level. Columns 3 and 4 are run at the student level. Additional controls: year and school fixed effects; student controls: gender, Free School Meal Eligible, Special Education Needs, primary school grade in science, math and English; schools controls: school size. The dependent variables in columns 3, and 4 are set equal to 0 if students do not continue studying or if they do not take the considered subjects. Robust standard errors clustered by school in parentheses. \* denotes significance at 10%, \*\* denotes significance at 5%, \*\*\* denotes significance at 1%.

Table 10: Identification based on changes in catchment areas

Dep. Var.:	1=age 16 (KS5) Science			1=STEM major		
	[1]	[2]	[3]	[4]	[5]	[6]
1=TS	0.111*** (0.013)	0.120*** (0.042)	0.108** (0.043)	0.028*** (0.010)	0.042 (0.054)	0.035 (0.060)
% reach school off $TS_{t-1}$	0.001 (0.002)			-0.002 (0.002)		
av. qual reach school			0.018** (0.007)			0.010 (0.007)
N	2847133	2850675	2850675	2392486	2395787	2392319
Neigh Fe	No	Yes	Yes	No	Yes	Yes

Additional controls: year fixed effects; student controls: gender, Free School Meal Eligible, Special Education Needs, primary school grade in science, math and English. All dependent variables are set equal to 0 if students do not continue studying or if they do not take the considered subjects. Robust standard errors clustered by neighbourhood in parentheses. \* denotes significance at 10%, \*\* denotes significance at 5%, \*\*\* denotes significance at 1%.

Table 11: Peers

Dep var:	$Q_{ist}^a$	1=KS5 sci	1=Russell	1=STEM	1=medic	1=grad	1=grad STEM
	[1]	[2]	[3]	[4]	[5]	[6]	[7]
Z offer*ks2 sci q1	-0.095*** (0.011)						
Z offer*ks2 sci q2	-0.060*** (0.008)						
Z offer*ks2 sci q3	-0.031*** (0.007)						
Z offer*ks2 sci q4	0.024*** (0.007)						
Z offer*ks2 sci q5	0.055*** (0.007)						
Z offer*ks2 sci q6	0.099*** (0.008)						
1=TS		0.053*** (0.006)	0.022** (0.011)	0.024** (0.012)	0.013* (0.008)	0.042* (0.025)	0.034*** (0.011)
qual peer (std)		0.021*** (0.005)	0.018*** (0.004)	0.003 (0.004)	-0.001 (0.003)	0.014 (0.009)	0.004 (0.004)
N	1648926	1621765	935630	935630	935630	935630	935630

Additional controls: year and school fixed effects; student controls: gender, Free School Meal Eligible, Special Education Needs, primary school grade in science, math and english; schools controls: school size. All dependent variables are set equal to 0 if students do not continue studying or if they do not take the considered subjects. Gr sci refers to sixtiles of the grade distribution in the science exam at the end of primary school (KS2). F statistic: 35.

<sup>a</sup> quality (based on science grade in ks2 (age 11) of peers in the same science class. Robust standard errors clustered by school in parentheses. \* denotes significance at 10%, \*\* denotes significance at 5%, \*\*\* denotes significance at 1%.

# Appendix

Table A1: Heterogeneity

Dep var:	1=KS5 sci [1]	1=Russell [2]	1=STEM [3]	1=medicine [4]	1=grad [5]	1=grad STEM [6]
<i>Panel 1: Quintiles science grade in primary school</i>						
3rd quintile						
1=TS	0.019 (0.015)	-0.002 (0.035)	-0.002 (0.037)	0.015 (0.028)	0.036 (0.089)	0.032 (0.036)
N	336723	203148	203148	203148	203148	203148
ymean	0.045	0.024	0.026	0.017	0.188	0.023
4th quintile						
1=TS	0.032*** (0.010)	0.041* (0.021)	0.076*** (0.021)	0.017 (0.014)	0.084* (0.046)	0.086*** (0.019)
N	344500	197276	197276	197276	197276	197276
ymean	0.104	0.053	0.045	0.024	0.277	0.042
5th quintile						
1=TS	0.053*** (0.007)	0.018 (0.016)	0.010 (0.015)	0.005 (0.008)	0.016 (0.023)	0.012 (0.015)
N	328076	181689	181689	181689	181689	181689
ymean	0.254	0.146	0.097	0.040	0.414	0.090
<i>Panel 2: Socio-Economics Status</i>						
High SES students (no FSM)						
1=TS	0.048*** (0.006)	0.024** (0.011)	0.020 (0.013)	0.015* (0.008)	0.037 (0.026)	0.033*** (0.012)
N	1431595	818880	818880	818880	818880	818880
ymean	0.093	0.052	0.041	0.020	0.226	0.037
Low SES students (yes FSM)						
1=TS	0.063*** (0.018)	-0.008 (0.044)	0.042 (0.039)	-0.003 (0.035)	0.100 (0.090)	0.024 (0.036)
N	258804	147854	147854	147854	147854	147854
ymean	0.034	0.015	0.018	0.010	0.103	0.016

Additional controls: year and school fixed effects; student controls: gender, Free School Meal Eligible, Special Education Needs, primary school grade in science, math and English; schools controls: school size. All dependent variables are set equal to 0 if students do not continue studying or if they do not take the considered subjects. Robust standard errors clustered by school in parentheses. \* denotes significance at 10%, \*\* denotes significance at 5%, \*\*\* denotes significance at 1%.



Table A2: Other balancing tests

	RF	RF	IV	IV
	[1]	[2]	[3]	[4]
Dep var:	1=Grade English prim school			
$Z_{st}$	-0.000 (0.004)	0.005 (0.005)		
1=TS			-0.001 (0.023)	-0.002 (0.023)
N	1690451	1690451	1690451	1690451
ymean	0.015	0.015	0.015	0.015
Dep var:	1=female			
$Z_{st}$	-0.002 (0.001)	-0.001 (0.002)		
1=TS			-0.009 (0.009)	-0.009 (0.009)
N	1690451	1690451	1690451	1690451
ymean	0.502	0.502	0.502	0.502
Dep var:	1=FSM			
$Z_{st}$	-0.000 (0.001)	-0.000 (0.002)		
1=TS			-0.001 (0.008)	-0.001 (0.008)
N	1690451	1690451	1690451	1690451
ymean	0.153	0.153	0.153	0.153
School Fe	Yes	Yes	Yes	Yes
School trend	No	Yes	No	Yes

Additional controls years dummies. All dependent variables are set equal to 0 if students do not continue studying or if they do not take that subject. Robust standard errors clustered by school in parentheses. \* denotes significance at 10%, \*\* denotes significance at 5%, \*\*\* denotes significance at 1%.

Table A3: Effect on other KS4 subjects (age 14)

Dep. var	All		Girls		Boys	
	[1] Coeff.	[2] Se	[3] Coeff.	[4] Se	[5] Coeff.	[6] Se
English lit	0.068**	(0.030)	0.075**	(0.030)	0.061*	(0.032)
Statistics	0.011	(0.034)	0.010	(0.038)	0.011	(0.034)
Food (voc.)	-0.027*	(0.016)	-0.047**	(0.024)	-0.009	(0.013)
Graphics (voc.)	-0.015	(0.014)	-0.002	(0.017)	-0.027	(0.017)
Material (voc.)	-0.014	(0.014)	0.000	(0.011)	-0.024	(0.022)
Art design	-0.008	(0.019)	0.001	(0.025)	-0.015	(0.019)
History	-0.032*	(0.019)	-0.045*	(0.023)	-0.022	(0.021)
Geogr	0.007	(0.020)	0.010	(0.024)	0.005	(0.022)
French	-0.015	(0.028)	-0.010	(0.033)	-0.020	(0.027)
German	-0.065***	(0.018)	-0.072***	(0.022)	-0.060***	(0.018)
Business	-0.012	(0.019)	-0.012	(0.020)	-0.014	(0.021)
Drama	0.007	(0.014)	-0.001	(0.020)	0.013	(0.014)
Inf tech	-0.034	(0.031)	-0.020	(0.032)	-0.048	(0.035)
Music	-0.001	(0.008)	-0.012	(0.011)	0.009	(0.010)
Media	-0.012	(0.022)	-0.016	(0.025)	-0.009	(0.023)
Fine art	0.005	(0.014)	0.007	(0.019)	0.004	(0.013)
Office technology	0.016	(0.028)	0.008	(0.032)	0.022	(0.028)
Applied buss	-0.001	(0.014)	-0.004	(0.015)	0.000	(0.015)
Health care	0.003	(0.011)	0.009	(0.022)	-0.002	(0.004)
Applied IT	-0.009	(0.021)	-0.009	(0.021)	-0.008	(0.024)

Each line represents a different regression. Columns 1, 3 and 5 display the coefficients on the independent variable  $1 = TS$ . All dependent variables are set equal to 0 if students do not take that subject. Usual controls. Robust standard errors clustered at the school level. \* denotes significance at 10%, \*\* denotes significance at 5%, \*\*\* denotes significance at 1%. I exclude math and English because compulsory in KS4.

Table A4: Effect on other KS5 subjects (age 16)

Dep. var	All		Girls		Boys	
	Coeff.	Se	Coeff.	Se	Coeff.	Se
Biology	0.035***	(0.005)	0.037***	(0.008)	0.034***	(0.006)
Chemistry	0.037***	(0.004)	0.032***	(0.006)	0.040***	(0.005)
Physics	0.025***	(0.003)	0.012***	(0.003)	0.036***	(0.005)
Math	0.024***	(0.005)	0.016**	(0.007)	0.031***	(0.007)
Textile (voc.)	-0.003*	(0.002)	-0.005	(0.003)	-0.001*	(0.000)
History	0.005	(0.005)	0.004	(0.008)	0.005	(0.006)
Economics	0.003	(0.003)	0.002	(0.003)	0.004	(0.005)
Law	-0.007**	(0.003)	-0.007	(0.005)	-0.008**	(0.004)
Psychology	-0.010*	(0.006)	-0.015	(0.011)	-0.006	(0.005)
Media film tv	-0.012***	(0.005)	-0.013*	(0.007)	-0.011**	(0.005)
German	-0.003**	(0.001)	-0.002	(0.002)	-0.003**	(0.001)
Music tech	-0.004***	(0.001)	-0.001	(0.001)	-0.008***	(0.002)
Accounting	-0.002*	(0.001)	-0.002	(0.002)	-0.002	(0.002)

Each line represents a different regression. Columns 1, 3 and 5 display the coefficients on the independent variable  $1 = TS$ . All dependent variables are set equal to 0 if students do not continue studying or if they do not take that subject. Usual controls. Robust standard errors clustered at the school level. \* denotes significance at 10%, \*\* denotes significance at 5%, \*\*\* denotes significance at 1%.

Table A5: Effect on other university majors (age 18)

Dep. variables	All		Girls		Boys	
	Coeff.	Se	Coeff.	Se	Coeff.	Se
Physics	0.006***	(0.002)	0.001	(0.003)	0.009***	(0.003)
Math	0.001	(0.002)	-0.002	(0.002)	0.003	(0.004)
Engineering	0.007***	(0.002)	0.003**	(0.001)	0.011***	(0.003)
Biology	-0.001	(0.003)	-0.001	(0.005)	-0.002	(0.004)
Veterinary agric	-0.001	(0.001)	-0.001	(0.002)	0.000	(0.001)
Computer sci	-0.001	(0.001)	-0.001	(0.001)	-0.000	(0.002)
Technology	-0.000	(0.001)	-0.000	(0.001)	-0.000	(0.001)
General science	-0.000	(0.001)	-0.001	(0.002)	0.000	(0.001)
Medicine	0.003*	(0.001)	0.006**	(0.002)	0.001	(0.001)
Allied medicine	0.004*	(0.002)	0.008*	(0.004)	0.000	(0.002)
Architecture	-0.003***	(0.001)	-0.002*	(0.001)	-0.004**	(0.002)
Other languages	0.000	(0.000)	-0.000	(0.001)	0.000	(0.001)
History	0.001	(0.002)	0.003	(0.003)	-0.001	(0.002)
Art design	-0.000	(0.003)	0.001	(0.005)	-0.002	(0.003)
Education	-0.001	(0.002)	-0.001	(0.004)	-0.001	(0.001)
Soc studies	0.003	(0.003)	0.005	(0.005)	0.001	(0.003)
Law	-0.004*	(0.002)	-0.006*	(0.003)	-0.002	(0.002)
Business	0.001	(0.003)	0.001	(0.004)	-0.000	(0.004)
Communication	0.000	(0.002)	0.001	(0.003)	-0.001	(0.002)
Ling classic	0.005**	(0.002)	0.004	(0.004)	0.006***	(0.002)
Eu languages	-0.000	(0.001)	-0.000	(0.002)	-0.000	(0.001)

Each line represents a different regression. Columns 1, 3 and 5 display the coefficients on the independent variable  $1 = TS$ . All dependent variables are set equal to 0 if students do not continue studying or if they do not take that subject. Usual controls. Robust standard errors clustered at the school level. \* denotes significance at 10%, \*\* denotes significance at 5%, \*\*\* denotes significance at 1%.

Table A6: Robustness: exclusion restriction

Dep var:	1=KS5 sci [1]	1=Russell [2]	1=STEM [3]	1=medicine [4]	1=grad [5]	1=grad STEM [6]
1=TS	0.057*** (0.007)	0.024* (0.014)	0.022 (0.013)	0.010 (0.009)	0.039 (0.028)	0.026** (0.012)
N	1613226	948058	948058	948058	948058	948058
ymean						

The sample includes only schools where the triple science class is not likely to be oversubscribed (class size not around a multiple of 30). Additional controls: year and school fixed effects; student controls: gender, Free School Meal Eligible, Special Education Needs, primary school grade in science, math and English; schools controls: school size. The dependent variables in columns 3, 4, 5 and 6 are set equal to 0 if students do not continue studying or if they do not take the considered subjects. Robust standard errors clustered by school in parentheses. \* denotes significance at 10%, \*\* denotes significance at 5%, \*\*\* denotes significance at 1%.

Table A7: Teachers

Dep. variable:	N teachers [1]	N qualified teachers [2]
1=TS	1.604 (1.267)	1.577 (1.249)
N	1022489	1022489
ymean	70.567	66.654

Additional controls: year and school fixed effects; student controls: gender, Free School Meal Eligible, Special Education Needs, primary school grade in science, math and english; schools controls: school size. Robust standard errors clustered by school in parentheses. \* denotes significance at 10%, \*\* denotes significance at 5%, \*\*\* denotes significance at 1%.

RECENTLY PUBLISHED “TEMI” (\*)

- N. 1082 – *Price dispersion and consumer inattention: evidence from the market of bank accounts*, by Nicola Branzoli, (September 2016).
- N. 1083 – *BTP futures and cash relationships: a high frequency data analysis*, by Onofrio Panzarino, Francesco Potente and Alfonso Puorro, (September 2016).
- N. 1084 – *Women at work: the impact of welfare and fiscal policies in a dynamic labor supply model*, by Maria Rosaria Marino, Marzia Romanelli and Martino Tasso, (September 2016).
- N. 1085 – *Foreign ownership and performance: evidence from a panel of Italian firms*, by Chiara Bentivogli and Litterio Mirenda (October 2016).
- N. 1086 – *Should I stay or should I go? Firms’ mobility across banks in the aftermath of financial turmoil*, by Davide Arnaudo, Giacinto Micucci, Massimiliano Rigon and Paola Rossi (October 2016).
- N. 1087 – *Housing and credit markets in Italy in times of crisis*, by Michele Loberto and Francesco Zollino (October 2016).
- N. 1088 – *Search peer monitoring via loss mutualization*, by Francesco Palazzo (October 2016).
- N. 1089 – *Non-standard monetary policy, asset prices and macroprudential policy in a monetary union*, by Lorenzo Burlon, Andrea Gerali, Alessandro Notarpietro and Massimiliano Pisani (October 2016).
- N. 1090 – *Does credit scoring improve the selection of borrowers and credit quality?*, by Giorgio Albareto, Roberto Felici and Enrico Sette (October 2016).
- N. 1091 – *Asymmetric information and the securitization of SME loans*, by Ugo Albertazzi, Margherita Bottero, Leonardo Gambacorta and Steven Ongena (December 2016).
- N. 1092 – *Copula-based random effects models for clustered data*, by Santiago Pereda Fernández (December 2016).
- N. 1093 – *Structural transformation and allocation efficiency in China and India*, by Enrica Di Stefano and Daniela Marconi (December 2016).
- N. 1094 – *The bank lending channel of conventional and unconventional monetary policy*, by Ugo Albertazzi, Andrea Nobili and Federico M. Signoretti (December 2016).
- N. 1095 – *Household debt and income inequality: evidence from Italian survey data*, by David Loschiavo (December 2016).
- N. 1096 – *A goodness-of-fit test for Generalized Error Distribution*, by Daniele Coin (February 2017).
- N. 1097 – *Banks, firms, and jobs*, by Fabio Berton, Sauro Mocetti, Andrea Presbitero and Matteo Richiardi (February 2017).
- N. 1098 – *Using the payment system data to forecast the Italian GDP*, by Valentina Aprigliano, Guerino Ardizzi and Libero Monteforte (February 2017).
- N. 1099 – *Informal loans, liquidity constraints and local credit supply: evidence from Italy*, by Michele Benvenuti, Luca Casolaro and Emanuele Ciani (February 2017).
- N. 1100 – *Why did sponsor banks rescue their SIVs?*, by Anatoli Segura (February 2017).
- N. 1101 – *The effects of tax on bank liability structure*, by Leonardo Gambacorta, Giacomo Ricotti, Suresh Sundaresan and Zhenyu Wang (February 2017).
- N. 1102 – *Monetary policy surprises over time*, by Marcello Pericoli and Giovanni Veronese (February 2017).
- N. 1103 – *An indicator of inflation expectations anchoring*, by Filippo Natoli and Laura Sigalotti (February 2017).
- N. 1104 – *A tale of fragmentation: corporate funding in the euro-area bond market*, by Andrea Zaghini (February 2017).

---

(\*) Requests for copies should be sent to:

Banca d’Italia – Servizio Studi di struttura economica e finanziaria – Divisione Biblioteca e Archivio storico – Via Nazionale, 91 – 00184 Rome – (fax 0039 06 47922059). They are available on the Internet [www.bancaditalia.it](http://www.bancaditalia.it).

2015

- AABERGE R. and A. BRANDOLINI, *Multidimensional poverty and inequality*, in A. B. Atkinson and F. Bourguignon (eds.), *Handbook of Income Distribution*, Volume 2A, Amsterdam, Elsevier, **TD No. 976 (October 2014)**.
- ALBERTAZZI U., G. ERAMO, L. GAMBACORTA and C. SALLESO, *Asymmetric information in securitization: an empirical assessment*, *Journal of Monetary Economics*, v. 71, pp. 33-49, **TD No. 796 (February 2011)**.
- ALESSANDRI P. and B. NELSON, *Simple banking: profitability and the yield curve*, *Journal of Money, Credit and Banking*, v. 47, 1, pp. 143-175, **TD No. 945 (January 2014)**.
- ANTONIETTI R., R. BRONZINI and G. CAINELLI, *Inward greenfield FDI and innovation*, *Economia e Politica Industriale*, v. 42, 1, pp. 93-116, **TD No. 1006 (March 2015)**.
- BARDOZZETTI A. and D. DOTTORI, *Collective Action Clauses: how do they Affect Sovereign Bond Yields?*, *Journal of International Economics*, v. 92, 2, pp. 286-303, **TD No. 897 (January 2013)**.
- BARONE G. and G. NARCISO, *Organized crime and business subsidies: Where does the money go?*, *Journal of Urban Economics*, v. 86, pp. 98-110, **TD No. 916 (June 2013)**.
- BRONZINI R., *The effects of extensive and intensive margins of FDI on domestic employment: microeconomic evidence from Italy*, *B.E. Journal of Economic Analysis & Policy*, v. 15, 4, pp. 2079-2109, **TD No. 769 (July 2010)**.
- BUGAMELLI M., S. FABIANI and E. SETTE, *The age of the dragon: the effect of imports from China on firm-level prices*, *Journal of Money, Credit and Banking*, v. 47, 6, pp. 1091-1118, **TD No. 737 (January 2010)**.
- BULLIGAN G., M. MARCELLINO and F. VENDITTI, *Forecasting economic activity with targeted predictors*, *International Journal of Forecasting*, v. 31, 1, pp. 188-206, **TD No. 847 (February 2012)**.
- CESARONI T., *Procyclicality of credit rating systems: how to manage it*, *Journal of Economics and Business*, v. 82, pp. 62-83, **TD No. 1034 (October 2015)**.
- CUCINIELLO V. and F. M. SIGNORETTI, *Large banks, loan rate markup and monetary policy*, *International Journal of Central Banking*, v. 11, 3, pp. 141-177, **TD No. 987 (November 2014)**.
- DE BLASIO G., D. FANTINO and G. PELLEGRINI, *Evaluating the impact of innovation incentives: evidence from an unexpected shortage of funds*, *Industrial and Corporate Change*, v. 24, 6, pp. 1285-1314, **TD No. 792 (February 2011)**.
- DEPALO D., R. GIORDANO and E. PAPAPETROU, *Public-private wage differentials in euro area countries: evidence from quantile decomposition analysis*, *Empirical Economics*, v. 49, 3, pp. 985-1115, **TD No. 907 (April 2013)**.
- DI CESARE A., A. P. STORK and C. DE VRIES, *Risk measures for autocorrelated hedge fund returns*, *Journal of Financial Econometrics*, v. 13, 4, pp. 868-895, **TD No. 831 (October 2011)**.
- CIARLONE A., *House price cycles in emerging economies*, *Studies in Economics and Finance*, v. 32, 1, **TD No. 863 (May 2012)**.
- FANTINO D., A. MORI and D. SCALISE, *Collaboration between firms and universities in Italy: the role of a firm's proximity to top-rated departments*, *Rivista Italiana degli economisti*, v. 1, 2, pp. 219-251, **TD No. 884 (October 2012)**.
- FRATZSCHER M., D. RIMEC, L. SARNOB and G. ZINNA, *The scapegoat theory of exchange rates: the first tests*, *Journal of Monetary Economics*, v. 70, 1, pp. 1-21, **TD No. 991 (November 2014)**.
- NOTARPIETRO A. and S. SIVIERO, *Optimal monetary policy rules and house prices: the role of financial frictions*, *Journal of Money, Credit and Banking*, v. 47, S1, pp. 383-410, **TD No. 993 (November 2014)**.
- RIGGI M. and F. VENDITTI, *The time varying effect of oil price shocks on euro-area exports*, *Journal of Economic Dynamics and Control*, v. 59, pp. 75-94, **TD No. 1035 (October 2015)**.
- TANELI M. and B. OHL, *Information acquisition and learning from prices over the business cycle*, *Journal of Economic Theory*, 158 B, pp. 585-633, **TD No. 946 (January 2014)**.

- ALBANESE G., G. DE BLASIO and P. SESTITO, *My parents taught me. evidence on the family transmission of values*, Journal of Population Economics, v. 29, 2, pp. 571-592, **TD No. 955 (March 2014)**.
- ANDINI M. and G. DE BLASIO, *Local development that money cannot buy: Italy's Contratti di Programma*, Journal of Economic Geography, v. 16, 2, pp. 365-393, **TD No. 915 (June 2013)**.
- BARONE G. and S. MOCETTI, *Inequality and trust: new evidence from panel data*, Economic Inquiry, v. 54, pp. 794-809, **TD No. 973 (October 2014)**.
- BELTRATTI A., B. BORTOLOTTI and M. CACCAVAIO, *Stock market efficiency in China: evidence from the split-share reform*, Quarterly Review of Economics and Finance, v. 60, pp. 125-137, **TD No. 969 (October 2014)**.
- BOLATTO S. and M. SBRACIA, *Deconstructing the gains from trade: selection of industries vs reallocation of workers*, Review of International Economics, v. 24, 2, pp. 344-363, **TD No. 1037 (November 2015)**.
- BOLTON P., X. FREIXAS, L. GAMBACORTA and P. E. MISTRULLI, *Relationship and transaction lending in a crisis*, Review of Financial Studies, v. 29, 10, pp. 2643-2676, **TD No. 917 (July 2013)**.
- BONACCORSI DI PATTI E. and E. SETTE, *Did the securitization market freeze affect bank lending during the financial crisis? Evidence from a credit register*, Journal of Financial Intermediation, v. 25, 1, pp. 54-76, **TD No. 848 (February 2012)**.
- BORIN A. and M. MANCINI, *Foreign direct investment and firm performance: an empirical analysis of Italian firms*, Review of World Economics, v. 152, 4, pp. 705-732, **TD No. 1011 (June 2015)**.
- BRANDOLINI A. and E. VIVIANO, *Behind and beyond the (headcount) employment rate*, Journal of the Royal Statistical Society: Series A, v. 179, 3, pp. 657-681, **TD No. 965 (July 2015)**.
- BRIPI F., *The role of regulation on entry: evidence from the Italian provinces*, World Bank Economic Review, v. 30, 2, pp. 383-411, **TD No. 932 (September 2013)**.
- BRONZINI R. and P. PISELLI, *The impact of R&D subsidies on firm innovation*, Research Policy, v. 45, 2, pp. 442-457, **TD No. 960 (April 2014)**.
- BURLON L. and M. VILALTA-BUFI, *A new look at technical progress and early retirement*, IZA Journal of Labor Policy, v. 5, **TD No. 963 (June 2014)**.
- BUSETTI F. and M. CAIVANO, *The trend-cycle decomposition of output and the Phillips Curve: bayesian estimates for Italy and the Euro Area*, Empirical Economics, V. 50, 4, pp. 1565-1587, **TD No. 941 (November 2013)**.
- CAIVANO M. and A. HARVEY, *Time-series models with an EGB2 conditional distribution*, Journal of Time Series Analysis, v. 35, 6, pp. 558-571, **TD No. 947 (January 2014)**.
- CALZA A. and A. ZAGHINI, *Shoe-leather costs in the euro area and the foreign demand for euro banknotes*, International Journal of Central Banking, v. 12, 1, pp. 231-246, **TD No. 1039 (December 2015)**.
- CIANI E., *Retirement, Pension eligibility and home production*, Labour Economics, v. 38, pp. 106-120, **TD No. 1056 (March 2016)**.
- CIARLONE A. and V. MICELI, *Escaping financial crises? Macro evidence from sovereign wealth funds' investment behaviour*, Emerging Markets Review, v. 27, 2, pp. 169-196, **TD No. 972 (October 2014)**.
- CORNELI F. and E. TARANTINO, *Sovereign debt and reserves with liquidity and productivity crises*, Journal of International Money and Finance, v. 65, pp. 166-194, **TD No. 1012 (June 2015)**.
- D'AURIZIO L. and D. DEPALO, *An evaluation of the policies on repayment of government's trade debt in Italy*, Italian Economic Journal, v. 2, 2, pp. 167-196, **TD No. 1061 (April 2016)**.
- DOTTORI D. and M. MANNA, *Strategy and tactics in public debt management*, Journal of Policy Modeling, v. 38, 1, pp. 1-25, **TD No. 1005 (March 2015)**.
- ESPOSITO L., A. NOBILI and T. ROPELE, *The management of interest rate risk during the crisis: evidence from Italian banks*, Journal of Banking & Finance, v. 59, pp. 486-504, **TD No. 933 (September 2013)**.
- MARCELLINO M., M. PORQUEDDU and F. VENDITTI, *Short-Term GDP forecasting with a mixed frequency dynamic factor model with stochastic volatility*, Journal of Business & Economic Statistics, v. 34, 1, pp. 118-127, **TD No. 896 (January 2013)**.
- RODANO G., N. SERRANO-VELARDE and E. TARANTINO, *Bankruptcy law and bank financing*, Journal of Financial Economics, v. 120, 2, pp. 363-382, **TD No. 1013 (June 2015)**.



- ALESSANDRI P. and H. MUMTAZ, *Financial indicators and density forecasts for US output and inflation*, Review of Economic Dynamics, v. 24, pp. 66-78, **TD No. 977 (November 2014)**.
- MOCETTI S. and E. VIVIANO, *Looking behind mortgage delinquencies*, Journal of Banking & Finance, v. 75, pp. 53-63, **TD No. 999 (January 2015)**.
- PATACCHINI E., E. RAINONE and Y. ZENOU, *Heterogeneous peer effects in education*, Journal of Economic Behavior & Organization, v. 134, pp. 190-227, **TD No. 1048 (January 2016)**.

#### FORTHCOMING

- ADAMOPOULOU A. and G.M. TANZI, *Academic dropout and the great recession*, Journal of Human Capital, **TD No. 970 (October 2014)**.
- ALBERTAZZI U., M. BOTTERO and G. SENE, *Information externalities in the credit market and the spell of credit rationing*, Journal of Financial Intermediation, **TD No. 980 (November 2014)**.
- BRONZINI R. and A. D'IGNAZIO, *Bank internationalisation and firm exports: evidence from matched firm-bank data*, Review of International Economics, **TD No. 1055 (March 2016)**.
- BRUCHE M. and A. SEGURA, *Debt maturity and the liquidity of secondary debt markets*, Journal of Financial Economics, **TD No. 1049 (January 2016)**.
- BURLON L., *Public expenditure distribution, voting, and growth*, Journal of Public Economic Theory, **TD No. 961 (April 2014)**.
- CONTI P., D. MARELLA and A. NERI, *Statistical matching and uncertainty analysis in combining household income and expenditure data*, Statistical Methods & Applications, **TD No. 1018 (July 2015)**.
- DE BLASIO G. and S. POY, *The impact of local minimum wages on employment: evidence from Italy in the 1950s*, Regional Science and Urban Economics, **TD No. 953 (March 2014)**.
- FEDERICO S. and E. TOSTI, *Exporters and importers of services: firm-level evidence on Italy*, The World Economy, **TD No. 877 (September 2012)**.
- GIACOMELLI S. and C. MENON, *Does weak contract enforcement affect firm size? Evidence from the neighbour's court*, Journal of Economic Geography, **TD No. 898 (January 2013)**.
- MANCINI A.L., C. MONFARDINI and S. PASQUA, *Is a good example the best sermon? Children's imitation of parental reading*, Review of Economics of the Household, **TD No. 958 (April 2014)**.
- MEEKS R., B. NELSON and P. ALESSANDRI, *Shadow banks and macroeconomic instability*, Journal of Money, Credit and Banking, **TD No. 939 (November 2013)**.
- MICUCCI G. and P. ROSSI, *Debt restructuring and the role of banks' organizational structure and lending technologies*, Journal of Financial Services Research, **TD No. 763 (June 2010)**.
- MOCETTI S., M. PAGNINI and E. SETTE, *Information technology and banking organization*, Journal of Financial Services Research, **TD No. 752 (March 2010)**.
- NATOLI F. and L. SIGALOTTI, *Tail co-movement in inflation expectations as an indicator of anchoring*, International Journal of Central Banking, **TD No. 1025 (July 2015)**.
- RIGGI M., *Capital destruction, jobless recoveries, and the discipline device role of unemployment*, Macroeconomic Dynamics, **TD No. 871 July 2012)**.
- SEGURA A. and J. SUAREZ, *How excessive is banks' maturity transformation?*, Review of Financial Studies, **TD No. 1065 (April 2016)**.
- ZINNA G., *Price pressures on UK real rates: an empirical investigation*, Review of Finance, **TD No. 968 (July 2014)**.