Promoting scientific faculties: does it work? Evidence from Italy

Maestri, Virginia

AIAS-University of Amsterdam

April 2009

Promoting scientific faculties: does it work? Evidence from Italy

Virginia Maestri*
AIAS, University of Amsterdam

First version: April 2009
This version: October 2010

Abstract

The object of this article is to assess the causal impact of promotions policies on students' choice of the field of study. We match the records of the students enrolled in two large universities with the records of the participating schools. Within the participating schools, some students took part in the program, while others did not. We adopted an "exposure" approach in which we define as treated all students of a cohort that were eligible for these activities. We find, on average, a positive and significant effect of the policy on targeted and non-targeted scientific bachelor's degrees and positive cross-treatment effects across subjects. However, if the policy has a considerable influence on male students' choices, it does not appear to have any effect on female students' choices. These findings suggest that the policy helped students in correcting their labor market expectations for graduating in science.

Keywords: economic impact, educational economics, school choice
JEL classification: H43; I23; I28

*Current correspondence: AIAS, University of Amsterdam, Plantage Muidergracht 12, 1018 TV Amsterdam, The Netherlands. I would like to thank: the Marie Curie Research Training Network "The Economics of Education and Education Policy in Europe" for founding the research; CREST for hosting me during the fall of 2008 and, in particular, Philippe Zamora and Pierre-Yves Cabannes; Daniele Checchi; Ugo Cosentino, Nice Terzi, Simona Binetti, Giorgio Sironi, and all the other persons with various titles involved in the "Progetto Lauree Scientifiche" for providing valuable information; Polo Qualità that kindly provided the data about participation to the PLS and the statistical offices of the University of Milan and Milan-Bicocca for providing the university enrolments' data; seminar participants to the Doctoral Workshop in Economic and Policy Analysis 2009, Warsaw International Economic Meeting 2009, Dutch Ministry of Education, Culture and Science 2011, Michela Braga and Veruska Oppedisano for useful comments.

1 Introduction

In the last decade, many developed countries have shown concern about scientific education. In some cases the absolute number of students enrolled in scientific and technological tracks increased, masking an overall expansion of secondary and tertiary education. The share of students enrolled in science and technology is, perhaps, a more appropriate indicator of the trend, as it normalizes changes in the student population and demographic trends. The evolution of the share of enrollments in scientific and technological studies at the secondary and tertiary educational level has been decreasing overall in most of the OECD countries over the last twenty years, though the number of students in Engineering and Computer Science has increased. The picture is more serious when considering traditional scientific disciplines such as Mathematics and Physics and when considering PhD programs, in which the decline also occurred in absolute terms (OECD, 2008).

A reversal of this downward trend should be desirable for a variety of reasons. First, R&D may become more difficult over time and a small number of scientists engaged in R&D can hinder economic growth (Segerstrom, 1998). Moreover, investing in studies that "bear" a higher chance of over-education[1] is inefficient. For instance, in the Italian context, investing in quantitative fields (including Science) not only increases the participation to the labor market and the probability of employment, but also early earnings[2] (Buonanno and Pozzoli, 2007). Furthermore, Webbink and Oosterbeek (1997) show that there is an unexploited technical potential. That is, there are some types of students who do no choose technical education, but when they do they perform better than others. For example, female students choose less often technical studies, but when they do, they perform better than male students. In addition, the choice of the degree (considering that girls choose less often scientific studies than boys) can explain 13% and up to 36% of the gender wage gap (Machin and Puhani , 2005). Therefore, more female students in science would translate to a smaller gender gap. After having briefly illustrated the importance of science, it is natural to wonder what drives the choice of the field of study.

The existing literature agrees on the importance of expected earnings in explaining the choice of the field of study. In a simultaneous model of field and length of studies, Beffy *et al.* (2009) find that a 10% increase of expected earnings in a given field results in a significant impact on the allocation of students between fields. Montmarquette *et al.* (2002) define a model in which the utility of choosing major k depends on the characteristics of the individual and his expected income corresponding to major k, that in turn

[1] See, for example, Frenette (2004).

[2] However, this results could be due to the relative scarcity of those who graduated in scientific fields (Ballarino and Bratti, 2006) and to the signal for high ability and flexibility attached to Italian scientific graduates (Convert, 2005).

depends on his perceived probability of success in major k, his expected earnings after graduation for major k and the earnings alternatives. In their model, preferences are an unobserved random component and they also assume that expected earnings are always realized. They find that expected earnings play a crucial role in college major choice, though the importance of earnings is lower for females and for non-whites. In another work, Boudabart and Montmarquette (2007) introduce a weight for expected income as a function of student characteristics, such as family background, and the probability of job-education skill match. They find a correlation between male students choosing science and a vocation for high income and for the acquisition of skills, while for female students choosing science they do not find such a correlation. Berger (1988) stresses the importance of the streams of future earnings *versus* early income (at the time of the choice) as driving the students' choice of college majors. Freeman *et al.* (2008) introduce the role of the knowledge content of a job. That is, the choice to undertake a specific major may be affected by the importance of the competencies provided by that major in the labor market, which may vary over years. They find that women, when choosing majors, are more responsive to this aspect, while men are more responsive to changes in the wage return of the knowledge content in a field. The different behavior of males and females is a constant feature in the literature about the choice of college major[3], but this difference does not seem to be addressed at the level of policy design. On the other hand, Arcidiacono (2004) argues that the large exogenous monetary premiums for attending science (and business) courses cannot explain the ability sorting across majors, due instead to differing preferences for majors across abilities.

However, a non-decreasing payoff of science cannot explain why scientific fields of study are chosen less often today than in past years. We agree with Convert (2005) that this apparently puzzling fact can be explained by a general (among the developed countries) expansion and democratization of the education system, where the students more "adverse" to science or less informed about studying science, like females and students with a low family background, have entered higher education. At this point, a more general question arises: can policy affect the choice of the field of study?

Some governments have attempted to attract more students in scientific tracks. The policy instruments that have been used fall under two categories: reduction of tuition fees and promotion of science studies. The first tool merely reduces the direct cost of investments in scientific education, as it is commonly assumed that the non-monetary costs of studying science are higher than those of studying, for instance, humanities. The second kind of intervention may play a role in the formation of the expectations about the probability of finding a (proper) job and about earnings after majoring in

[3] See for example Zafar (2009).

science and it may increase the non pecuniary returns of studying science by stimulating the interest in scientific subjects. Moreover, promotion activities may help students to familiarize themselves with scientific subjects and, as a possible consequence, they may upgrade their perception of the probability of success in scientific majors.

Italy is not an exception of the general trend of declining majoring in science. In the next years there will be a generational turnover of high school teachers and, in particular, new teachers of science are needed. Moreover, the contribution of the private sector to the founding of research, a crucial aspect for science, is 40% with respect to a European average of 53.4% (Ministero della Pubblica Istruzione, 2007). The performance of 15 year-old Italian students in science, reported by the survey OCSE-PISA 2006, is not encouraging. Aware of this situation, in 2005 the Ministry of Education and Research, the National Employers Organization (*Confindustria*) and the Universities launched a program to boost enrollment in scientific bachelors and to increase the number of graduates in science. With respect to the first aim, the content of this policy was to promote scientific studies to high school students.

This paper analyzes the impact of the sponsoring policy on enrollments in Chemistry, Physics, Mathematics and Materials Science[4] to see whether it has been effective and, more in detail, for whom it worked and in which fields. We match the records of the students enrolled in the two main public universities of Milan with the records of the secondary schools that participated to the program. We use an "exposure" approach to identify the "intention" effect of the treatment. We find that the "Progetto Lauree Scientifiche" increased the probability to choose a scientific bachelor, on average, by about 1.5%. The effect is not limited to the selected disciplines and there are positive cross effects of treatment's subject for Mathematics and Physics. When considering a differential effect of the policy for males and females, we find that for male students the policy was even more successful, with a shift in the probability of enrollment in science equal to 3.5 percentage points, while for female students it seems the policy had no effect.

The remainder of the paper is organized as follows. Section 2 discusses the content and the functioning of the policy. Section 3 presents the available data. Section 4 explains the "exposure" approach proposed to evaluate the effect of the policy. Section 5 reports the descriptive statistics of our sample. Section 6 presents our findings and Sections 7 adds some robustness checks. Finally, Section 8 draws some conclusive remarks.

[4]Materials Science is a course of study in chemistry and physics of materials.

2 The PLS policy

The "Progetto Lauree Scientifiche" (PLS) is a policy launched in 2005 by an agreement between the Ministry of Education and Research, the National Employers Organization and the National Committee of Science and Technological Universities to increase the enrollments and the number of graduates in Chemistry, Physics, Mathematics and Materials Science. The distinguishing characteristic of this project, with respect to previous interventions, is the considerable amount of its founding and its coverage of the national territory: 11 million euro invested and more than 30 universities involved[5]. The program includes different interventions: 4 about the sponsoring of science to students and the training of teachers at the secondary school level, 3 concerning university education, stage and post-graduate studies and 1 for scholarships and other activities at the tertiary level (Ministero della Pubblica Istruzione, 2007). The first block of interventions aimed at increasing enrollments in scientific bachelors programs and was intended for high school students, while the other blocks aimed at increasing the number of graduates in science and were intended for university students.

The present analysis focuses on what we think is the most innovative content of this policy: the activities to promote scientific studies to secondary school students. The project was initially introduced for two years: 2005 and 2006. The content of this sponsoring policy is summarized by an evocative sentence: *"We need to change the idea that Math is boring, Physics difficult and Chemistry dangerous"* (page 10 of Ministero della Pubblica Istruzione (2007)). The activities to promote scientific studies to high school students aimed, on the one hand, at stimulating the interest in subjects that are commonly deemed to be boring and, on the other, they aimed at filling the gap between the perception of the professions one can undertake after these studies and the variety of the applications they can have (both in terms of the width of labor opportunities and the social utility of these studies). In practice, the project organized lab activities stimulating an active participation of students and experiments to show the links between science and everyday life. A second, but not minor, pillar upon which the project is focused is the involvement of secondary school teachers to stimulate the interest of their students in science, realized through training and support for lab activities provided by university professors and researchers.

The target of the policy was students of the last three years of the secondary cycle of education. 50000 students, 20000 teachers and 2000 high schools[6] participated in these activities in the entire national territory (Ministero della Pubblica Istruzione, 2007). The PLS deliberately intended to attract students according to their interest. The rationale of the policy was

[5]Only two regions were not involved in the program: Valle d'Aosta and Molise.

[6]Schools are counted twice if they participated to two activities.

to boost an underlying propensity or ability for scientific studies and not to create it. The explicit aim was to increase the matriculations in the bachelors of Chemistry, Physics, Mathematics and Materials Science.

Originally, all secondary schools received an official communication about the PLS policy, but few of them were actually involved in this way. In most of the cases the universities that organized these activities relied on e-mails, personal contacts with teachers and principals and on teacher-to-teacher advertising. Once the organizers of the activities had a reference person in a school, through one or more teachers or the principal, the participation could occur with one or more classes or with small groups of students, depending on the type of activity. The teacher of reference was not necessarily the teacher of the students involved, as the program was limited to the last three years of high school. Furthermore, many activities were mainly targeted to the last or before last year students of high school and could take place in the first or the second semester of the school year. From direct interviews that we had with the organizers of the activities, it often happened that the decision of a school to participate with the 4^{th} or 5^{th} grade depended on the collocation of the activity in the first or second semester. In the second semester many 5^{th} graders were busy preparing for their final exam. Students could participate in one or more activities, in one or more of the four subjects. Finally, the program organized a survey of the participants to assess their evaluation of the activities.

The PLS policy may have affected three aspects regarding the choice of the field of study. First, practical examples of professions linked with every day life may have "corrected" the labor market expectations related to graduates in science. Second, the proposal of non-standard activities may have increased the non-pecuniary returns of studying science. Finally, the time spent following the PLS activities may have helped students to familiarize themselves with scientific subjects and increased students' expectations of success in studying science. While the first aspect may, according to the literature, affect mostly male students, the second aspect could be more important for female students.

3 Data

We anticipate the section about the data as the (non) availability of (proper) data has constrained the empirical strategy. Thus, we explain which data we have and in the next section, given the data, which strategy we use. We had access to the data of the students' enrollments in the two public universities of Milan: Università degli Studi di Milano and Università degli Studi Milano-Bicocca, containing the records of the high school attended by the student and its address, the year of enrollment, the year, age and mark of the final high school diploma, the type of school, gender, family income bracket and the chosen bachelor. The data cover four school years: from 2004 to 2007[7].

A second dataset contains the information of the high schools participating in the program, with the address of the school, the subject of the program in which the school was involved, the year of involvement and the grades involved or the grade that individual students attended when participating. This dataset corresponds to the questionnaires submitted to the participants of the PLS after they completed the activities. About the questionnaires the program defined three cases: for some activities the questionnaires had to be filled in by individual students, for others they had to be filled in at the class level, whereas for some activities (especially for the more general ones, such as the showing of movies about science, museum visits and so on) the program did not request to fill out any questionnaires. Some information about school participation was missing. We collected the remaining information through direct contact with the professors who organized these activities. We then matched these two datasets by the address of the secondary school.

It was not possible to link the information about program's participation at the individual level, even for the questionnaires filled out by the students, as the dataset about the program's participation does not contain an individual identifier. Since the corrections we made on the two data-sets and the recovery of the missing information was considerably time consuming, we restricted the analysis to the students enrolled in the two universities of Milan, having attended a high school in the province of Milan.

It is worth noting that among the students having attended the secondary school in the province of Milan and among those that decided to continue studying at the tertiary level, about 90% enrolled at a university in the same region and among them, 93% enrolled at a university in the province of Milan[8]. However, Milan has several universities: the Catholic University (15.6%[9]), the *IULM* (2.8%), the Bocconi University (7.2%), the *Politecnico* (21.4%) and few medical and art institutions (2.2%). The Catholic University of Milan does not offer bachelors degrees in science: they are instead offered at a campus located in a city 85 km away from Milan[10]. Similarly, the *IULM*, the Bocconi University and the *Politecnico* do not offer bachelors in science, though the *Politecnico* offers "competing" scientific bachelors degrees such as Engineering. The Statale and Bicocca Universities enroll about 51% of all the students enrolled in one of the universities of Milan, including those who do not provide courses in science and including students from

[7] These years correspond to the school year 2004/2005 to 2007/2008.

[8] See www.anagrafe.miur.it.

[9] The total is the overall number of students enrolled in one of the universities of Milan, including students from other provinces, regions and countries. See www.sistemauni.it.

[10] Brescia.

other provinces, regions and countries[11]. Furthermore, the activities of the PLS program were organized by the two universities for which we have the data. It might be that the policy pushed some students to undertake a scientific bachelor's degree, but at a university for which we do not have the data: for instance at the University of Pavia. For the facts listed above we think that the two public universities of Milan constitute a natural outcome for most of the students for which the policy could have had an effect. In any case, this aspect should not invalidate our analysis, as long as the policy does not have an effect on the choice of the place of study.

The final dataset contains the records of students coming from 320 different secondary schools, out of which 56 were involved in the program. The participating schools "sent" to the two universities 14172 students from 2004 to 2007. From this matched sample we exclude the (few) schools that do not have a stable number of students enrolled in the two universities over the four year period 2004-2007. We also drop the students with a gap year between the diploma and the first year of matriculation. The final sample includes 6333 observations for post-policy years (2006 and 2007). The schools could participate in the PLS in 2005 and/or 2006, with the 5^{th} or the 4^{th} grade. In these schools 46% of the students were exposed to the activities organized by the PLS program since its inception, corresponding to 3371 individuals for 2006 and 2007.

From the data about the project's participation for which the number of participants is available[12], we know that an average of approximately 16 students participated for each high school's cohort. We also know that in our final data about university enrollments, each high school (of the province of Milan) "sent", on average, 64 students (living in the province of Milan) per year. If all the participants to the PLS enrolled in the two public universities of Milan, then 25% of the "exposed" cohorts were actually treated[13].

4 Empirical strategy

The aim of the analysis is to estimate the effect of the PLS sponsoring policy on the enrollment of students in the targeted scientific bachelors programs. As this policy was not designed to be a randomized experiment, the effect of the participation to the PLS may be confounded by several factors. First, the schools involved may be the schools that better prepare students for scientific faculties. Second, the students who participated to the program could be the students most interested in science. Third, the classes that participated can have, for instance, better teachers of scientific subjects.

[11] See www.sistemauni.it.

[12] The number of participants is not available for all the PLS activities and, in any case, for the individual questionnaires it just corresponds to the number of participants who actually answered the questionnaire.

[13] However, we do not know who they are.

Last, a school may have participated with some classes of the 4^{th} grade cohort because there were not similar students in the 5^{th} grade cohort.

In order to avoid the confounding effect of schools' self-selection, we only consider the secondary schools involved in the sponsoring policy, where we can find both the students who participated to the policy and the students who did not. The results are limited to the effect of the policy for the schools that chose to participate. To avoid the confounding effect of students' and classes' self-selection and, also, for the constraints imposed by the data[14], we adopt an "exposure" approach[15]. That is, we assign the treatment to the whole cohort of each school where one or more classes or some students of the same cohort were involved in the program. In this way the treatment effect is comparable to an intention to treat effect. Usually, the intention to treat effect is used in randomized trials when non compliance may be non random[16]. In this case we use the intention to treat because the actual treatment is likely endogenous and because we do not have information on actual participants. If we had the information on actual participants we could have used it to estimate an upper bound of the treatment or we could have used the school/cohort "exposure" as an instrumental variable for actual participation. In order to be eligible for the treatment the student should have been in school S in grade g in year t. The actual treatment effect corresponds to the treatment coefficient divided by the proportion of students for each cohort in each school that actually participated in the program. The intention to treat coefficient is a lower bound of the actual treatment effect.

Just to give an example, let's say that school B was involved in the project in 2005 with one class in the 4^{th} grade cohort. A student in the 4^{th} grade in 2005 can enroll at the university in 2007. We define as treated all the students from that cohort of school B. If school C participated in the program in 2006 with some students of the 4^{th} grade, school C is included in the analysis and its students enrolled at the university in year 2006 and 2007 are included in the control group. Indeed, "exposed" students of school C can enroll at the university at the earliest in 2008, but our data only include enrollments in 2006 and 2007. The Table below illustrates the mechanism of the treatment's assignment:

[14] We cannot link the (in any case, incomplete) individual data about PLS' participation to the university administrative data.

[15] The "exposure" approach is generally used when a given policy has an effect spread over a certain subpopulation, identified, for instance, by the date and region of birth. For an application see Bratti *et al.* (2008) and Oppedisano (2009).

[16] See for example Angrist and Lavy (2002).

year and class PLS	$\rightarrow$	*year enroll.*	school A	school B	school C	school D
2005 5th grade	$\rightarrow$	**2006**	**T**	**C**	**T**	**C**
2005 4th (2006 5th)	$\rightarrow$	**2007**	**C**	**T**	**T**	**C**
2006 4th grade	$\rightarrow$	2008	C	C	C	T

The identification of the treatment effect together with school-specific effects and a time trend is allowed by the within and between school year and/or grade variation of the exposure to the treatment. The interpretation of the treatment effect relies on the assumption that school-specific cohorts are comparable within school and between (participating) schools, conditioning on the pre-policy share of students of the school enrolled in scientific faculties. In other words, in this framework school participation reveals a pre-existing interest toward science, selection into the program occurs at school, class and individual level, but the choice of the timing and the grade for the participation to the program is not correlated with an underlying motivation of the school-specific cohort toward science and we assume that in each grade of each school we can always find a portion of students and/or teachers more interested in science than others.

The basic idea of this approach is that the cohort's composition of students and teachers in each school should be more stable and less selective than class variation or students' variation, in terms of motivation toward science. For example, science teachers can teach in more cohorts in the same school and the potential manipulation of student allocation may occur when forming classes and not for the cohorts, that are determined by the age of the students. However, it could still be possible that a cohort in a school is systematically different from another cohort in the same school or from a cohort in another school and that this difference drives the participation (of a portion of these cohorts) into the program. In Section 7 we will introduce some variables to control for observable school/cohort characteristics. The resulting model is:

$$y *_{ics} |(U=1) = \alpha + \beta C_c + \gamma S_s + \xi X_{ics} + \delta treat_{cs} + \varepsilon_{ics} \tag{1}$$

$$y_{ics} = \begin{cases} 1 & \text{if } y*_{ics} > 0 \\ 0 & \text{if } y*_{ics} \leq 0 \end{cases}$$

$$\varepsilon \sim (0, \pi^2/3)$$

where $y*_{ics}$ is the latent propensity to attend a scientific faculty of the i^{th} student in the c^{th} cohort in the s^{th} school, C_c is a dummy for the cohort to capture the trend, S_s is the pre-policy rate of enrollment in scientific bachelor programs for each school and X is a vector of students' characteristics (gender, final grade of diploma, having repeated one or two grades during high school, family income bracket). The variable $treat_{cs}$ is defined as belonging to a high school's cohort in which some classes or some students participated to one or more activities organized by the PLS. These students are not necessarily actual participants but, among them, some actually participated. Thus, the coefficient δ is an intention to treat effect and the average treatment effect for the treated in the participating schools.

As we consider only the participating schools and we do not assign the treatment at the individual or class level, we do not have to deal with possible unobservable and confounding factors such as η_s, υ_i and ν_g, respectively, at the school, individual and class level, unless the intensity and the timing of the treatment is systematically related to the underlying motivation toward science. The outcome y_{ics} is a binary variable for whether or not the student attends a bachelor in Chemistry, Physics, Mathematics or Materials Sciences. We have to condition on U (equal to one if the student is enrolled at the university) as we do not have data about the students of the s^{th} high school that do not enroll at the university. Hence, with this model we cannot disentangle the effect of the treatment on the probability to go to the university to follow a scientific track from the effect on the probability to choose a scientific track, once chosen to go to the university (independently from the policy). The coefficient δ includes both mechanisms. Finally, since the error terms may be correlated within school and cohort and the coefficient of interest is defined at this level, we cluster by school and cohort to capture common unobservable shocks to students in the same school and cohort. We allow for an unrestricted correlation structure.

5 Descriptive statistics

Figure 1 shows the trend of the share of the enrollments at the two universities of Milan in Chemistry, Physics, Mathematics and Materials Sciences from 2004 to 2007, separately for the schools participating in the PLS and those that did not. The vertical line corresponds to the year in which the policy may have started to have an effect on these enrollments. The positive change in the trend at the point the policy was introduced for the group of the schools involved is clearcut, especially in the first year of its implementation. Figure 2 shows the same picture, but with the absolute number of the enrollments in the four sciences rather than the share on the total number of enrollments. The difference in the trends for the two groups of schools is slightly more pronounced than in Figure 1. The participating schools already had a larger share and absolute number of students enrolled in science before the policy was introduced, but the steep increase after 2005 is considerable. In the first year after the introduction of the policy the effect seems to be driven by Physics, while in the second year by Chemistry. Indeed, in 2006 13% of students were exposed to a treatment in Physics, while only 6% the following year. The opposite is true for Chemistry: 5% were treated in 2006 and 42% in 2007.

[insert Figure 1 about here] [insert Figure 2 about here]

Figure 3 shows the trend in science's enrollments for males and females, in non participating and participating schools. If for females the increase in enrollments for the participating schools seems to happen only in the first year after the introduction of the policy, for male the increase is notable and constant over the following two years. Figure 4 shows the difference between the share of students enrolled in science between the cohorts exposed and not exposed to the treatment, only for the participating schools. Figure 4 seems to confirm that the gain from being exposed to the policy is higher for males than for females.

[insert Figure 3 about here]

[insert Figure 4 about here]

Tables 1 and 2 report the descriptive statistics of the sample of participating schools, by treatment status. Table 1 reports mean values using students as the unit of measurement, while Table 2 is calculated using cohorts as the unit of measurement. When we assign equal weight to each cohort, the only significant difference between treated and controls is the year of enrollment. On the other hand, if we assign a weight to the cohorts proportional to their number of students we find that the treated and control students differ in the pre-policy share of enrollments in science and in the share of males of the high school of provenience. However, the share of males is not a pre-treatment variable. Thus, the descriptive characteristics reported in Tables 1 and 2 do not seem to seriously invalidate our cohort approach, where we would basically like to have the treated and non treated cohorts as similar as possible.

[insert Table 1 about here]

[insert Table 2 about here]

6 Empirical findings

Table 3 reports the estimation of equation 1. In order to interpret the results in terms of probabilities we also report the corresponding marginal effects. The attended high school's pre-policy share of students enrolled in a scientific track is the major determinant of the choice of a scientific bachelor. Male students are 8% more likely to choose a scientific track and the higher the mark of the diploma, the higher this chance. On the other side, students who failed a grade during secondary school are less likely to choose science at the university. There is no significant trend in the enrollment of science between 2006 and 2007, that we could interpret as a stable (perceived) labor market reward for choosing a scientific track. Introducing a control for family income does not change the results. The possibility to participate in the PLS project significantly increases by 1.4% the probability to enroll in scientific faculties. By using a liner probability model the effect raises to 1.9%. Thus, for the cohorts who had the chance to participate, the treatment shifts upward the probability to enroll in science. We gather all four scientific bachelor degree in one dummy for science studies or "other studies". We have to remember that only a small fraction of the cohorts defined as treated have actually been treated. For this reason we expect the actual (individual) effect of the PLS policy to be higher than that found here.

[insert Table 3 about here]

As the figures reported in this paper suggest, it is worthwhile to investigate whether the effect of the treatment is the same for males and for females. By calculating a proper interaction effect between treatment and gender[17], we indeed find that the positive effect of the policy is driven by male students. The effect of treatment for males is 3.5% higher than for females with a z-value of 2.1 and this effect dominates for most of the values of the other control variables. Figure 5 shows this result.

[insert Figure 5 about here]

The treatment includes activities in Chemistry, Physics, Mathematics or Materials Sciences. We expect an effect of the subject-specific treatment on the choice of the same subject as the activity attended, but we do not exclude that it can have an effect on the other three scientific subjects. Table 4 shows the cross treatment effect for each of the four scientific subjects. We restrict the analysis to the students who participated in the activities in only one subject, to avoid the correlation between participations in different subject-specific activities. As expected, the effects are found on the diagonal. However, we cannot give a causal interpretation of these coefficients, as cohorts more inclined toward one subject may choose to participate (even the small fraction) in the activities in that specific subject. Interestingly, there are cross effects for Mathematics and Physics. Participating in activities about Physics has a positive effect on choosing Mathematics and *vice versa*.

[insert Table 4 about here]

In addition, the policy was intended to promote the four mentioned scientific subjects, but there may be supplementary effects. First, if the

[17]In a non linear model framework, this interaction effect corresponds to the cross derivative of the bachelor choice equation with respect to gender and treatment and not to the single derivative with respect to the interaction gender and treatment. For the calculation we used the Stata command inteff (Ai and Norton, 2003).

policy diverted potential students from other scientific fields not included in the policy's target, it would not be considered a full success. As already mentioned, the policy can act in two ways: by redistributing students across faculties and by increasing the number of students that choose to purse their studies at the university to follow a scientific track. The first mechanism can, in turn, work by redistributing students from overcrowded tracks or from other scientific and reasonably non rival tracks, according to the philosophy of the policy. Last, the policy can have a positive effect on other bachelors. Table 5 shows that the policy does not appear to have diverted students from other scientific and quantitative bachelors. Moreover, there is a strong positive and significant effect of the treatment on Pharmacy, reasonably due to the fact that this faculty also includes a course of study in Chemistry and Pharmaceutical Technologies.

[insert Table 5 about here]

This result, together with the increase in the absolute number of male enrollments from the participating schools, might point in favor of the mechanism of the policy as boosting students to go to the university in order to follow a bachelors degree in science.

7 Robustness checks

The treated and untreated groups differ in some (omitted) characteristics that could bias the result about the project's participation. In fact, the cohort's composition in each school may reflect the learning environment of the students of our sample, through the peer effect or as a result of common shocks as in a case where teachers are assigned to cohorts according to the characteristics of the students. In other words, the difference in some observable characteristics between treated and untreated cohorts can have an effect itself on the choice of scientific bachelors and not merely through the treatment status. In order to check this possibility we perform an estimation including controls for cohorts' average characteristics in each school[18]. Results are reported in the first column of Table 6. Column 2 reports the marginal effects. Cohorts' average characteristics do play a role in the choice of college major, especially the share of grade repeaters and the proportion of males. Conditional on the average characteristics, the treatment coefficient is slightly reduced to 1.3% as well as the corresponding standard error. Unfortunately, we do not have pre-treatment characteristics. Therefore, the selection of cohorts into treatment may have not been based on these characteristics and they could not represent the actual learning environment of the student. They could be, instead, an effect of the treatment

[18] Average characteristics are calculated excluding the i^{th} observation for each observation.

on the university's enrollment. In particular, we cannot exclude that the project's participation had an effect on male university's enrollment. In column 3 and 4 of Table 6 we report the results by replacing the share of males for the school/cohort combination that the student belongs to with the average share of males in the same school for the two cohorts preceding the introduction of the policy. With this latter specification the treatment coefficient is confirmed to increase by around 1.4-1.5% the probability to choose a scientific bachelor, with a confidence level within 5%.

[insert Table 6 about here]

As explained in Section 2 the PLS policy also provided training courses for secondary school teachers. We know the year of participation, the school the teachers were involved with and that they were teachers of the last three grades of the high school. We cannot identify the effect of teacher training as all cohorts of students in our data, coming from the schools where these teachers were working, would be defined as treated. In other words, teacher training is introduced as a school level control. As a robustness check we insert in our specification a control for teacher training for the students that could have had these teachers. If we cannot interpret the effect of teacher training on student enrollment in science, the omission of this variable could distort our results. For instance, if the students that are classified in the control group are the students of the teachers that received the training, the coefficient of the treatment for the students' activities could be underestimated. This more precise specification strengthens our results: conditional on teachers' training, the treatment increases by 1.5% the probability to choose a scientific bachelor, with a confidence level of 4.4% (regression not reported).

8 Final remarks

The intuition of the PLS policy seems right: policy interventions should be made before the choice of the field is made by the students. This is contrary to interventions such as the reduction of tuition fees in science that in Italy, for example, act *ex-post*. The policy succeeded in increasing enrollments in Chemistry, Physics, Mathematics, Materials Sciences and, unintentionally, in Pharmacy. Participating in activities in one subject is correlated with the probability of enrolling in a bachelor's program in that subject but not exclusively, as in the case of the students participating in the activities of Physics and Mathematics. For these subjects, the PLS project seems to boost a general attitude toward a scientific approach, rather than a specific interest in the selected disciplines.

Nonetheless, a more accurate insight into the effect of the policy leads to a somehow different conclusion. The policy was very effective for male students, but there seems to be no effect for female students. Having participated in the treatment raises the probability of enrolling in a scientific track by 3.5 percentage points for males with respect to females and this result holds for most of the values of the other characteristics of the student. If we follow the existing literature on the choice of college major, as depending on expected income, being a function of the probability to find a (proper) job and of expected earnings, and on a non-pecuniary utility of studying science, we believe that the PLS policy was more effective in tackling the first issue. Indeed, if males react more than females to changes in expected income and treated males benefit the most from the policy, we can imagine that these activities helped students in correcting their labor market expectations for graduating in science.

Overall, our results are robust to different specifications and poorly sensitive to controls for additional factors such as cohorts' average characteristics and teacher training activities. The effect of the treatment on the probability to choose a scientific bachelor ranges from 1.3% to 1.8%. It is worthwhile to note that the effect that we identify with the cohort approach is an intention to treat effect. Consequently, we expect the actual (individual) effect of participating in the PLS activities to be somehow higher and the higher, the lower the proportion of students actually treated on the size of the same cohort enrolled at the university. For instance, if this proportion were 20% of the cohorts classified as treated, the actual policy effect could go up to 17.5% for males. From the analysis we can also deduce that the policy worked through pushing students that would not have chosen to pursue their studies at the university to follow a scientific track, rather than by redistributing students across bachelors programs. Indeed, we do not find a negative effect of the PLS policy on the enrollments in other bachelors and the data show an increase in the absolute number of the students of science.

The overall cost of the 7 projects organized by the universities of Milan (2 for Chemistry, 2 for Physics, 2 for Mathematics and 1 for Materials Science) was 717838 euro, including the activities to train high school teachers. A 1.5% effect of the policy[19] for our sample of treated schools/cohorts including 3371 students means that 51 students chose a scientific bachelor thanks to the policy. Thus, the (maximum) cost to have one student more studying science is about 14000 euro[20]. An extension of this research could be to estimate the benefit of having one student more studying science.

The limitation of our study is the short sight of our evaluation, due to the fact that the "Progetto Lauree Scientifiche" is a policy introduced very

[19] We are just considering the effect of the sponsoring activities for high school students and not the effect of teacher training (as we do not know the magnitude of this effect).

[20] This per capita estimation is valid only if the effect of teacher training on the students' choice of scientific bachelor is zero, otherwise this amount is an overestimation of the costs.

recently. We do not know if what we identify as a success will become a later drop-out from science bachelors and/or if it translates to a lower academic performance of the students boosted to study science.

It would be interesting to compare these findings with other studies on the effect of policies sponsoring scientific majors, especially for policies designed to be a randomized experiment. Our results, supported by the existing literature, would suggest to paying more attention to gender differences in the choice of college major when designing policies to sponsor scientific studies.

References

Ai C., Norton E. C., 2003. Interaction terms in logit and probit models. Economics Letters 80, 123-129.

Angrist J. D., Lavy V., 2002. The effect of high school Matriculation awards: evidence from randomized trials. NBER Working Paper 9389.

Arcidiacono P. (2004). Ability sorting and the returns to college major. Journal of Econometrics 121, 343-375.

Ballarino G., Bratti M., 2006. Fields of study and graduates' occupational outcomes in Italy during the 90s. Who won and who lost? DEAS Working Paper 2006-17.

Beffy M., Fougère D., Maurel A., 2007. The field and the length of studies in the French post-secondary education: the effect of expected earnings. IZA DP n. 4127.

Berger M. C., 1988. Predicted future earnings and choice of college major. Industrial and labor Relations Review 41 (3), 418-429.

Boudabart B., Montmarquette C., 2007. Choice of fields of study of Canadian university graduates: the role of gender and their parents' education. IZA Discussion paper 2552.

Buonanno P., Pozzoli D., 2007. Early labor market returns to college subject. Quaderni di Ricerca del Dipartimento di Scienze Economiche "Hyman P. Minsky" 5.

Bratti M., Checchi D., De Blasio G., 2008. Does the expansion of higher education increase equality of opportunities?. Labor 22, 53-88.

Convert B., 2005. Europe and the crisis in scientific vocations. European Journal of Education 40 (4), 361-366.

Frenette M., 2004. The overqualified Canadian graduate: the role of the academic program in the incidence, persistence, and economic returns to over-qualification. Economics of Education Review 23, 29-45.

Freeman J. A., Hirsch B. T., 2008. College majors and the knowledge content of jobs. Economics of Education Review 27, 517-535.

Machin S., Puhani P., 2005. The contribution of degree subject to the gender wage gap for graduates. A comparison of Britain, France and Germany. Report to the Anglo-German Foundation.

Ministero della Pubblica Istruzione, 2007. Il progetto "Lauree Scientifiche". Annali della Pubblica Istruzione 2-3.

Montamarquette C., Cannings K., Mahseredjian S., 2002. How do young people choose college majors? Economics of Education Review 21, 543-556.

OECD, 2008. Encouraging student interest in science and technology studies. OECD Publications: Paris.

Oppedisano V., 2008. The (adverse) effects of expanding higher education from Italy. Geary Institute Working Paper 45/2009.

Segerstrom P. S., 1998. Endogenous growth without scale effects. American Economic Review 88 (5), 1290-1310.

Webbink D., Ooosterbeek H., 1997. Is there a hidden technical potential? De Economist 145 (2), 159-177.

Zafar B., 2009. College major choice and the gender gap. Federal Reserve Bank of New York Staff Report 364.

Figure 1: Trend of the share of enrollments in science, by school participation in PLS

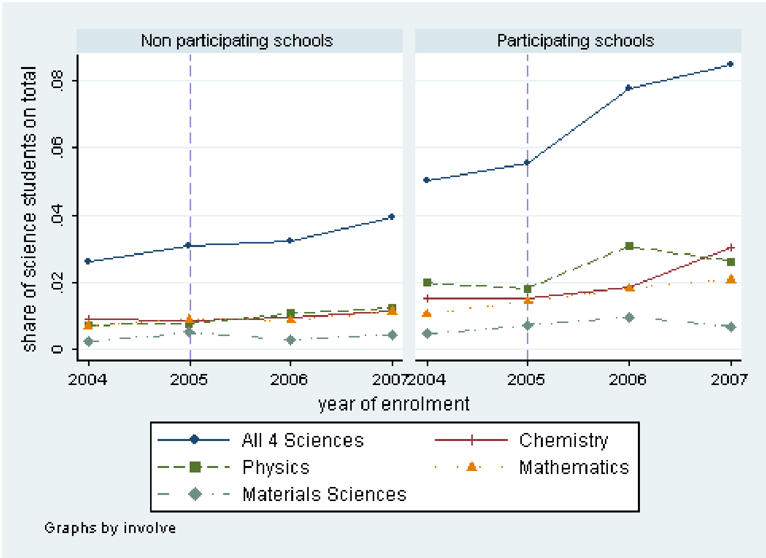

Figure 2: Trend of the number of enrollments in science, by school participation in PLS

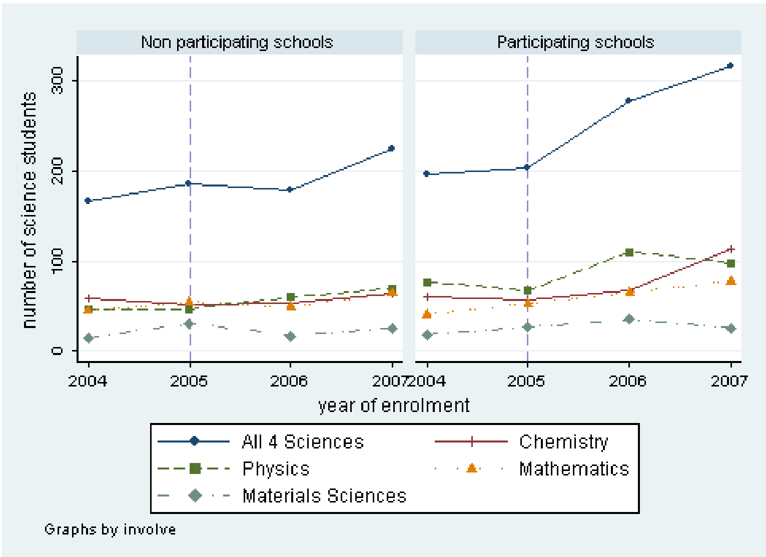

Figure 3: Trend of the number of enrollments in science, by school participation and gender

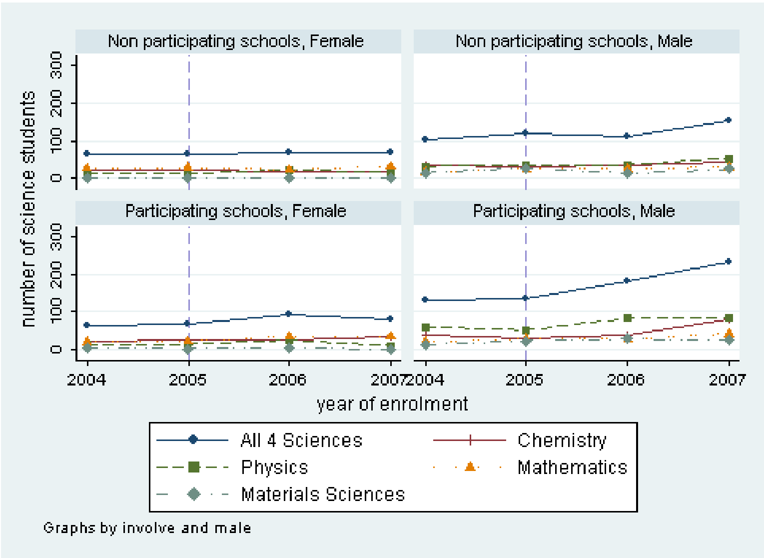

Figure 4: Gap in the share of enrollments in science between treated and control group, by gender

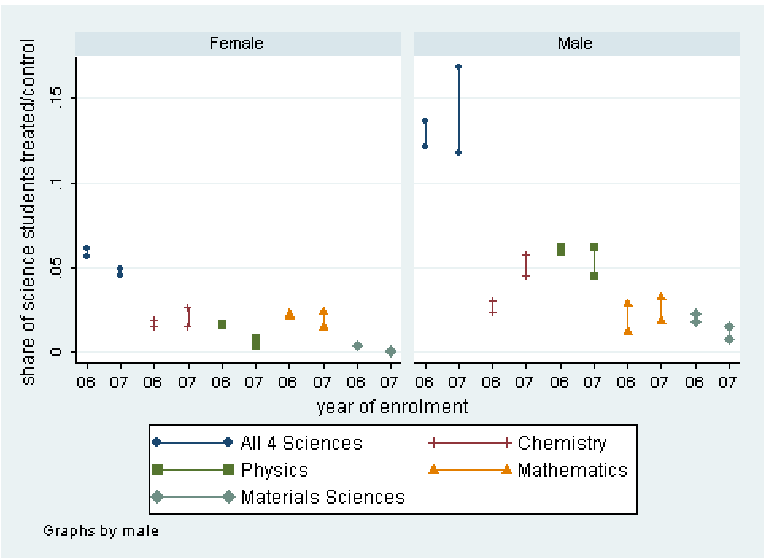

Table 1: Descriptive statistics (with students as unit of observation)

	control group	treated group	difference
Science	0.080	0.105	-0.025**
	(0.005)	(0.005)	(0.007)
Pre-policy share science (per school)	0.052	0.057	-0.005**
	(0.001)	(0.001)	(0.001)
enrollment 2007	0.346	0.671	-0.325**
	(0.001)	(0.001)	(0.012)
Male	0.434	0.487	-0.053**
	(0.009)	(0.009)	(0.013)
Mark diploma	0.031	-0.012	0.042
	(0.018)	(0.018)	(0.025)
Fail	0.136	0.147	-0.011
	(0.006)	(0.006)	(0.009)
N	3102	3231	6333

Table 2: Descriptive statistics (with cohorts as unit of observation)

	control cohorts	treated cohorts	difference
Science	0.080	0.106	-0.026
	(0.011)	(0.010)	(0.015)
Pre-policy share science (per school)	0.058	0.057	-0.002
	(0.005)	(0.004)	(0.007)
enrollment 2007	0.348	0.691	-0.343**
	(0.058)	(0.063)	(0.086)
Male	0.524	0.550	-0.027
	(0.033)	(0.031)	(0.046)
Mark diploma	-0.044	-0.014	0.030
	(0.040)	(0.049)	(0.063)
Fail	0.172	0.171	-0.000
	(0.019)	(0.015)	(0.025)
N	69	55	124

Table 3: Main effect of the participation in the PLS program

	Logit	Marginal effects	OLS
Pre-policy share science	9.4269**	0.5536**	0.7770**
	(1.5574)	(0.0920)	(0.1334)
enrollment 2007	-0.0144	-0.0008	0.0016
	(0.1244)	(0.0073)	(0.0092)
Male	1.2758**	0.0816**	0.0981**
	(0.0939)	(0.0076)	(0.0094)
Mark diploma	0.7687**	0.0451**	0.0641**
	(0.0548)	(0.0031)	(0.0059)
Failure	-0.4173*	-0.0216*	-0.0230*
	(0.2071)	(0.0096)	(0.0101)
Treatment	0.2450*	0.0144*	0.0186*
	(0.1235)	(0.0073)	(0.0092)
N	6333	6333	6333

Legend: † p<0.10 $*$ p<0.05 $**$ p<0.01. Standard errors are clustered by school and cohort. Marginal effects calculated at mean values.

Figure 5: PLS treatment effect heterogenous by gender

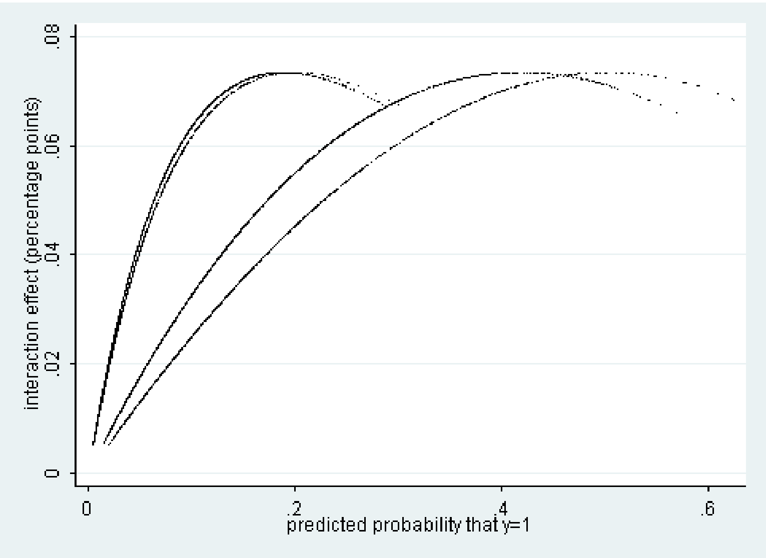

Table 4: Cross subject-specific treatment effects on each subject

	Chemistry	*Physics*	*Math*	*Mat Sc*
Chemistry treatment	0.4874†	-0.3253	0.0166	0.2392
	(0.2767)	(0.4579)	(0.2848)	(0.3994)
Physics treatment	-0.0632	0.5620*	1.1245**	-1.2700
	(0.2617)	(0.2585)	(0.2728)	(0.9217)
Mathematics treatment	-0.1178	0.5449*	0.6292*	-0.5599
	(0.2105)	(0.2592)	(0.2522)	(0.7159)
Materials Sciences treatment	-0.1281	0.0044	-0.0837	0.8906*
	(0.3458)	(0.2440)	(0.3626)	(0.3635)
N	5697	5697	5697	5697

Legend: † p<0.10 $*$ p<0.05 $**$ p<0.01. Control variables for each estimation not reported in the table are as in Table 3. Standard errors are clustered by school and cohort.

Table 5: PLS treatment effect on other bachelor programs

	treatment
Biotechnology	0.0047
	(0.1621)
Biology	0.1471
	(0.1348)
Pharmacy	0.3393*
	(0.1341)
Environmental Sciences	-0.2718
	(0.4076)
Geology	0.5024
	(0.3405)
Economics	-0.2000
	(0.1706)
Medicine	0.2028
	(0.2554)
Statistics	-0.2987
	(0.3143)

Legend: † p<0.10 $*$ p<0.05 $**$ p<0.01. Standard errors are clustered by school and cohort.

Table 6: Estimation with cohort's average characteristics

	Logit	Mfx	Logit	Mfx
Pre-policy share science	7.4287**	0.4242**	7.9608**	0.4579**
	(1.6249)	(0.0926)	(1.6281)	(0.0928)
enrollment 2007	-0.0137	-0.0008	-0.0135	-0.0008
	(0.1090)	(0.0062)	(0.1107)	(0.0064)
Male	1.2142**	0.0751**	1.2234**	0.0763**
	(0.1032)	(0.0074)	(0.1047)	(0.0076)
Mark diploma	0.7706**	0.0440**	0.7705**	0.0443**
	(0.0549)	(0.0029)	(0.0547)	(0.0029)
Fail	-0.3845	-0.0195*	-0.3838	-0.0196*
	(0.2078)	(0.0095)	(0.2075)	(0.0096)
Cohort's share male	1.3889**	0.0793**		
	(0.3310)	(0.0185)		
Cohort's share male (pre)			1.1639**	0.0669**
			(0.3129)	(0.0181)
Cohort's av mark diploma	0.1023	0.0058	0.0978	0.0056
	(0.2570)	(0.0147)	(0.2656)	(0.0153)
Cohort's share fail	-3.0149**	-0.1722**	-2.7224**	-0.1566**
	(0.7445)	(0.0414)	(0.6889)	(0.0397)
Treatment	0.2244*	0.0128*	0.2539*	0.0146*
	(0.1102)	(0.0063)	(0.1123)	(0.0065)
N	6333	6333	6333	6333

Legend: † $p<0.10$ $*$ $p<0.05$ $**$ $p<0.01$. Standard errors are clustered by school and cohort. Marginal effects in columns 2 and 4 are calculated at mean values.