The Indirect Effects of Educational Expansions: Evidence from a Large Enrollment Increase in STEM Majors

Nicola Bianchi*

January 5, 2016

Abstract

Increasing access to education may have consequences that go beyond the effects on marginal students induced to enroll. It may change school quality, peer effects, and returns to skill. This paper studies the effects of an educational expansion on student learning, exploiting an Italian reform that changed the admission requirements for university STEM majors. Newly collected administrative data on 27,236 students indicate that the reform decreased learning in STEM fields due to overcrowded universities and negative peer effects. The analysis of long-run incomes suggests that the reform might have had a long-lasting negative effect on the returns to STEM degrees.

JEL Classification: I21, I26, I28, J24, N34.
Keywords: educational expansion, STEM majors, higher education, quality of education, class heterogeneity.

*Kellogg School of Management, nicola.bianchi@kellogg.northwestern.edu. I thank Ran Abramitzky, Caroline Hoxby, Petra Moser, and Luigi Pistaferri for extensive feedback on this project. I also thank Dan Barron, Michael Dinerstein, Pascaline Dupas, Michela Giorcelli, Guzman Gonzalez-Torres, Daniel Grodzicki, Ben Jones, Davide Malacrino, Niko Matouschek, Giovanni Peri, Pietro Tebaldi, Stephen Terry, Constantine Yannelis, and seminar participants at Bank of Italy, Chicago, Columbia, Indiana, Kellogg, LSE, NBER, RPI, Stanford, Yale, and the All-UC Graduate Student Workshop in Economic History for their helpful comments and suggestions. This research was supported by the George P. Shultz grant and the Haley-Shaw fellowship through the Stanford Institute for Economic Policy Research, as well as the Graduate Research Opportunity Award through the School of Humanities and Sciences.

1 Introduction

This paper studies the effects of increasing access to postsecondary education on the learning of university students, exploiting a policy that changed the admission requirements for Italian universities. Historically, only graduates of academic-track high schools could enroll in university studies. In 1961, an education reform allowed students with a technical high school diploma to enroll in university STEM majors for the first time. In only a few years, the Italian policy led to an abrupt increase in enrollment and to a more diverse student population in the affected programs. In this context, increased enrollment might have congested students' access to teaching resources, such as professors and teaching assistants. In addition, learning could have been more difficult in classes where students had more heterogeneous precollegiate qualifications. The Italian setting, then, makes it possible to jointly study how changes in these two important inputs—access to university resources and quality of peers—affect student learning. This paper tests whether more congestion of teaching resources and higher class heterogeneity lowered human-capital accumulation in STEM majors among the incumbent academic-track students, who could enroll in STEM before and after 1961.

To analyze the reform, this paper uses newly-collected administrative data for the population of 27,236 students who completed high school in Milan—the Italian city with the thickest market for university graduates— between 1958 and 1968. Data from high school registries include the grades students receive on the exit exam. University transcripts contain detailed information on each course students enroll in, each exam they take, and the degrees they attain. Thanks to the high level of detail of university transcripts, the measure of learning used in this paper is the grade received by academic-track students in each university course. The sample is restricted to courses that were required for graduation, because their syllabi were less likely to vary over time. It is important to notice that grades were not assigned on a curve, especially in this time period when most tests were oral and students were graded sequentially.

The identification strategy exploits the fact the the reform differentially increased crowding and class heterogeneity between otherwise similar courses within the same major. To study crowding, this paper focuses on the student–faculty ratio, a measure of congested access to professors and teaching assistants. Before the reform, courses with similar enrollment could have different student–faculty ratios, mainly because some courses—those linked to a chair endowed by the state—had more power to hire teaching assistants. After the reform, the student–faculty ratio grew more in courses where resources were scarcer before the reform, even though enrollment increased similarly in all compulsory courses within a major.

This paper compares how the grades of academic-track students (those who could enroll in STEM majors before 1961) changed after the reform in courses where the student–faculty ratio increased more, relative to other courses in the same STEM major. To study class heterogeneity, this paper exploits the fact that newly-enrolled technical students were more likely to generate negative peer effects and disrupt teaching in university courses where they were less prepared by their high school studies. This paper compares how the grades of the incumbent academic-track students changed after the reform in university courses that were not taught in technical high schools, relative to other courses in the same major.

This paper finds that the grades of academic-track students decreased disproportionately in courses with more congestion and in courses with higher class heterogeneity after the reform. From 1961 to 1964, temporary enrollment caps limited the entry of technical students in STEM majors, leading to only moderate congestion of teaching resources. In the average STEM course, the student–faculty ratio increased by only 3.36 students per teaching fellow and the grades of academic-track students decreased by 0.004σ. After enrollment caps for technical students were removed in 1965, the student–faculty ratio increased by 38.95 students per teaching fellow and the grades of academic-track students decreased by 0.091σ in the average STEM course. In courses where technical students were not prepared, the grades of academic-track students decreased by 0.076σ during the first phase and by 0.165σ during the second phase. Overall, these estimates suggest that the educational expansion had a negative impact on learning of incumbent students.

To check the robustness of these findings, this paper addresses several concerns about changes over time in the sample of academic-track students in STEM majors. The results are robust to the inclusion of different time trends, which intend to capture the possible effect on grades of the post-World War II increase in school participation. Another concern is that the reform might have changed the way in which students sorted into different high schools. If the sample is restricted to cohorts who chose a high school before the reform and could not respond along this dimension, the main findings do not change. There is, however, suggestive evidence that some academic-track students might have avoided STEM majors after the reform, in favor of other programs still not accessible to technical students after 1961. This behavior could be explained by more difficult learning in STEM majors, by increased competition for STEM jobs, or by a weaker signal of high aptitude provided by STEM degrees to the labor market. If the sample is restricted to students who were not likely to change their major choice after the reform—graduates of scientific academic-track schools with low high school grades—, the results still indicate that grades decreased disproportionately in courses with more congestion and higher class heterogeneity. In addition, the results are robust to the use of sampling weights that match the average observable characteristics of pre-reform and post-reform cohorts. The main results are also robust to the inclusion of professor fixed effects, which indicates that the findings are not driven by professors' characteristics.

If students in STEM majors acquired less human capital, the reform might have lowered the returns to a STEM degree. To test this hypothesis, the education data are merged with the (former) students' 2005 income information from their tax returns. This is the only year the Italian government made available, following a short-lived attempt to foster the dissemination of income data. Since the individuals were between 56 and 67 years old in 2005, the observed income is a long-run measure of earnings. At this stage of the life cycle, the returns to education should be realized in full and the signaling content of education should be limited. The paper first compares the long-run income of academic-track students with a STEM degree to the long-run income of students with a different university degree, before and after the reform. The analysis shows that the income premium associated with a STEM degree decreased among the cohorts who completed high school from 1965 to 1968. In this period, the effects of congestion of teaching resources and negative peer effects were especially high. The effects are larger among academic-track students with low high school grades, who were less likely to change major choice after the reform. Lower human capital could account for three quarters of the total income loss, based on a simple model that describes how an educational expansion affects individual earnings. In the absence of other unaccounted factors, the remaining share could be attributed to higher competition for STEM jobs in the labor market.

This paper then uses a second identification strategy to corroborate the results on long-run income. Technical students in the industrial-track and commercial-track attended different technical schools, but faced the same constraints in the admissions to the university before the reform. In spite of their similarities, only the industrial-track students were allowed to enroll in STEM majors in 1961. If learning decreased in STEM majors, the reform might have reduced the returns to STEM degrees for the newly-admitted industrial students. The negative effects should be larger among the cohorts who enrolled during the second phase of the reform (1965 to 1968), when congestion and peer effects led to larger learning losses. This paper compares the long-run income of industrial-track technical students (allowed to enroll in STEM majors in 1961) to the long-run income of commercial-track technical students (not allowed to enroll in STEM majors in 1961), before and after the reform. The analysis shows that the average income of industrial students increased among the cohorts who completed high school from 1961 to 1964. These cohorts were allowed to enroll in the high-paying STEM majors when the negative effects of educational expansion were still limited. In spite of a dramatic surge in the university graduation rates of industrial students after 1965, the average income among the cohorts who completed high school from 1965 to 1968 did not increase, relative to the previous cohorts. Overall, the evidence provided by long-run incomes appears consistent with the hypothesis that the reform lowered the returns to a STEM degree, in particular for the cohorts who experienced more congestion and higher class heterogeneity.

By studying the effects of an educational expansion on student learning, this paper is related to three main strands of the literature. First, a few papers explore how crowding of university inputs affects academic outcomes (Stapleton and Young, 1988; Bound and Turner, 2007). This literature relies primarily on cross-cohort variation in the congestion of university resources. By combining a shock to the number of university students with preexisting differences in school inputs between university courses, this paper identifies the effect of crowding based solely on within-major, within-cohort, between-course variation.

Second, a large literature examines how the composition of a class affects students' outcomes. Most of these papers exploit natural variations in class composition across the units of observation (usually classes, schools, or neighborhoods) or time (Figlio and Page, 2002; Lavy, Paserman and Schlosser, 2012; Anelli and Peri, 2013), while a few rely on policy changes (Hoxby and Weingarth, 2006; Cooley, 2010) or randomized control trials (Duflo, Dupas and Kremer, 2011). The literature on class composition focuses primarily on precollegiate education, in part because self-selection into different universities or majors makes identification more problematic. This paper, however, is able to estimate the effect of class heterogeneity at the university by exploiting within-major, within-cohort, between-course variation in student preparedness.[1]

Third, a few papers study how education policies might have general equilibrium effects on skill prices (Lee, 2005; Lee and Wolpin, 2006; Abbott et al., 2013). Heckman, Lochner and Taber (1998*a*; 1998*b*) build a life-cycle model, in which educational expansion affects skill prices through changes in the relative supply of different types of human capital. In a different context, Duflo (2004) shows how a large school-construction program in Indonesia decreased the wages of older and untreated individuals by increasing the supply of educated workers in the economy. This paper adds to these findings, highlighting how educational expansions could also decrease the returns to education through lower human-capital accumulation.

The rest of the paper is organized as follows. Section 2 outlines the policy change. Section

[1] The literature on peer effects at the university is mainly based on random housing assignments. For this reason, it identifies the effect of roommates—and not necessarily classmates—on achievement. De Giorgi, Pellizzari and Woolston (2012), however, estimate the effects of class composition at the university in a context with random assignment of students across sections within a course. The identification strategy in this paper is different because it relies on a variation in the student precollegiate preparedness across courses within a major.

3 describes the data. Section 4 presents a model of educational expansion. Section 5 shows evidence that the reform lowered learning and addresses several concerns about the empirical analysis. Section 6 presents results on long-run income. Section 7 concludes.

2 Institutional Details

2.1 The Italian High School System

Italian high schools offer different diplomas. General education schools (*licei*; type A schools) focus on either the humanities (*licei classici*) or the sciences (*licei scientifici*) and traditionally prepare students for the university. Their curricula range from philosophy, Latin, and Ancient Greek to mathematics and physics (Appendix Table B1). Technical high schools, instead, train professionals for specific economic sectors.[2] The two main types of technical schools are industrial schools (*istituti industriali*; type B schools), which prepare students for jobs such as chemists and surveyors, and commercial schools (*istituti commerciali*; type C schools), which prepare students for jobs in accounting and the service sector. The technical curricula focus on applied disciplines and have a narrower scope than academic-track programs.

At age 14, students choose a high school. They can self-select in different tracks, because admission into public high schools does not depend on past performance and is typically granted to all applicants. On the one hand, type A schools provide a better preparation for most university majors but are characterized by a heavy workload and are academically more challenging. On the other hand, technical schools grant access to well-paid professions that do not require a university degree, but they do so at the expenses of a more general education. In Italy, the family background is a strong predictor of high school choice. For instance, data from the Bank of Italy's Survey of Household Income and Wealth (SHIW) show that 74.9 percent of technical graduates have a father with 8 or fewer years of completed education, compared with only 53 percent of academic-track students (Appendix Table B2). Similarly, 80.1 percent of technical graduates have a mother with 8 or fewer years of completed education, compared with 66 percent of academic-track students.

2.2 Expanding Access to University STEM Programs

In Italy, high school participation rose steadily after the end of World War II. The share of 14 to 18 years old enrolled in high school increased from 8.5 percent in 1950 to 20.3 percent

[2] Technical schools are different from vocational and trade schools, which last only 3 years (instead of 5) and have a narrower scope.

in 1960 (Istat). Relative to other institutions, enrollment in type B schools grew at a faster pace. In 1950, type B students constituted only 1.5 percent of the high school–age population (14 to 18 years old), while type A students made up 4.1 percent (Appendix Figure B1). By 1960, type B students made up 4.7 percent and type A 5.7 percent of the high school–age population.

In this setting of steady increases in enrollment, this paper studies an abrupt surge in the number of students enrolled in university STEM majors. Until 1961, type A students could enroll in any university major, while technical students could enroll only in business economics and statistics.[3] In July 1961, a policy change known as law 685 allowed type B students to enroll in university STEM programs for the first time. Following the reform, type B students could major in engineering, mathematics, physics, chemistry, biology, geology, natural science, and agricultural science.

This reform intended to increase the number of university-educated STEM workers and to make the university less elitist. These two issues became more pressing as the Italian economy boomed and the enrollment in technical high schools increased. The exact timing of the policy, however, was plausibly exogenous and determined primarily by politics. In July 1960, a faction of the ruling party, which was more receptive towards the requests of the center-left, formed a new government and prepared the ground for a university reform.

The reform was carried out in two phases. Until 1964, universities set caps on the number of type B students who could enroll. If demand exceeded this cap, type B applicants were chosen with an admission test. As mandated by the original law, the enrollment caps for type B students were eliminated in 1965. During the first phase, the number of STEM freshmen rose by 35.3 percent from 12,222 students in 1960 to 16,643 in 1964 (Figure 1, Panel A). In terms of share of 19-year-olds, STEM freshmen increased from 1.5 percent in 1960 to 2.4 percent in 1964 (Appendix Figure B2). During the second phase, freshmen increased by an additional 132.1 percent to 38,627 students (4.6 percent of the 19-year-old population) in 1968. In addition to increasing the number of enrolled students, the reform led to a major compositional change in the student population. By 1967, type B freshmen had become more numerous than type A (Figure 1, Panel B).

Even after the reform, many majors remained accessible to only type A students: law, medicine, the humanities, the social sciences (with the exception of business economics, accessible to type B students before 1961), and architecture. The number of freshmen in majors with restricted access increased by only 24 percent during the first phase, from 20,382

[3] The major in business economics (*economia e commercio)* focused primarily on accounting and was preparatory for becoming an accountant or tax preparer. All technical students (type B and type C) had access to this major, even though only type C graduates had received a precollegiate preparation in accounting.

students in 1960 to 25,280 in 1964, and by 64.3 percent during the second phase, to 41,533 students in 1968.

3 Education Data

To analyze the effects of the policy, I collected the school transcripts of 27,236 students who completed high school between 1958 and 1968 in Milan, Italy. Traditionally an industrial powerhouse, Milan is also a major financial and commercial center. It has the thickest market for university graduates and university-type jobs, and it is believed to be the place where a university graduate can earn the highest returns. In 2012, its income per capita of €31,761 was 87.5 percent higher than the national average of €16,937.[4] Due to the undisputed prominence of Milan in the Italian economy, the dataset assembled for this paper could be described as the Italian counterpart of an individual-level dataset of New York City students in the US.

3.1 High School Registries

I collected and digitized official registries containing the grades of all students who completed high school between 1958 and 1968 in Milan. The final sample includes 27,236 students from 17 high schools (7 type A, 6 type B, and 4 type C) out of 19 public schools that were operating throughout this time period (Appendix Figure A1). Of type B students, 98.6 percent were male, compared with 66.8 percent of type A students and 56.8 percent of type C students (Table 1). On average, type B and C students were one year older than type A.

At the end of their fifth year, students take a national examination (*maturità*) in order to graduate. The registries report the outcome of the exit exam (a numerical score from 0 to 10 with 6 as passing grade) for each discipline in the curriculum. On average, type A students graduated with a 6.48 GPA, type B with a 6.36 GPA, and type C with a 6.38 GPA. The registries identify home-schooled students (7.2 percent of type A, 8.4 percent of type B, and 16.9 percent of type C) and repeaters (9.7 percent of type A, 8.2 percent of type B and type C).[5] The high school score standardized by year and school and the binary indicators for home-schooled students and repeaters are used as measures of precollegiate ability.

[4] Data on average per capita personal income at the municipality level can be accessed from this link: http://www1.finanze.gov.it/analisi_stat.

[5] "Home-schooled" students (*privatisti*) take the national exam in the school, although they are not enrolled during the regular school year. Some "home-schooled" students come from private schools that are not allowed to administer the national exam, while others are educated at home. Repeaters are 20 years old or older, when they attempt the national exam.

In each high school, cohorts are divided randomly into classes of 20–30 students at the beginning of freshman year. Usually, the initial assignment remains unchanged in the following years, except for students who move to new classes after failing a grade. For each student, her other classmates' average score is used as a measure of precollegiate peer quality.

3.2 University Transcripts

For the same sample of students who completed high school in Milan between 1958 and 1968, I collected and digitized full transcripts from the two public universities of Milan—*Università Statale* (State University) and *Politecnico* (Polytechnic)—and from the private *Università Cattolica* (Catholic University).[6]

The transcripts are an incredibly rich source of information. For each course attended, they contain the course title, the exam date, and the grade received (from 0 to 30 cum laude with 18 as passing grade). The transcripts include the final mark (*voto di laurea*), a function of the GPA and a final thesis, that is salient to potential employers during a job search. The transcripts also contain the start and end dates of each student's university career, together with a description of the final outcome (graduation, dropout, or transfer). Of type A students, 86.7 percent enrolled in university and 63.7 percent graduated (Table 1). In comparison, only 40.8 percent of type B students and 40.2 percent of type C students enrolled in university, while only 16.1 percent of type B students and 10.5 percent of type C students graduated.

The fact that the transcripts are only available from the three local institutions could raise some concerns about the representativity of the sample, if a large share of students who attended high school in Milan enrolled in universities located elsewhere. More than 90 percent of high school students from Milan, however, chose a local university (Istat). Moreover, the portion of students who stayed in Milan for their university studies remained high over the period under consideration (94.1 percent in 1956 and 93.5 percent in 1967).

4 A Model of Educational Expansion

This section proposes a model in which educational expansion can lower learning and returns through three channels: crowding of university resources, higher class heterogeneity, and lower skill prices.

[6] The sample does not contain transcripts for Università Bocconi, a private university located in Milan. This does not affect the analysis for two reasons. First, Bocconi specializes in business and economics and admission into these majors was not restricted before 1961. Second, Bocconi was the only highly selective university in the Italian system, due to restricted admission and high tuition fees.

In the model, students choose between a university major or work (HS). The majors are divided into three groups: STEM ($STEM$), majors not accessible to type B students after 1961 (R for restricted majors), and majors accessible to type B students before and after 1961 (NR for nonrestricted majors). Aggregate production in the economy is determined by a CES production function that uses the four types of human capital:

$$Y = (S_{HS}H_{HS}^{\rho} + S_{STEM}H_{STEM}^{\rho} + S_R H_R^{\rho} + S_{NR}H_{NR}^{\rho})^{\frac{1}{\rho}}, \tag{1}$$

where H_m is the aggregate supply of human capital with education m, S_m are share parameters, $\rho = \frac{\phi-1}{\phi}$, and ϕ is the elasticity of substitution between the four types of human capital. Assuming perfect competition in the labor market, the skill price of human capital $m \in \{HS,\ STEM,\ R,\ NR\}$ is $w_m = Y^{1-\rho}S_m H_m^{\rho-1}$. In turn, the wage of individual i with education m is the product of skill prices, which are the same for all workers with the same education, and human capital, which varies across individuals ($W_i^m = w_m \cdot h_i^m$). Individual human capital is the weighted sum of the knowledge acquired in each university course (k_{ic}^m): $h_i^m = \sum_{c \in N_m} \mu_c \cdot k_{ic}^m$, where N_m is the set of courses in major m. The individual education production function in the university course c is:

$$k_{ic}^m = \gamma_0^m + \gamma_1^m X_i + \gamma_2^m C_c + \gamma_3^m Q_c + \gamma_4^m \mathrm{CH}_{ic} + u_{ic}^m, \tag{2}$$

where X_i are individual characteristics, C_c are course characteristics, Q_c is quality of education in course c, and CH_{ic} measures class heterogeneity. The utility of choosing major m is the sum of net nonmonetary preferences, the log of monetary returns, and a random idiosyncratic shock.

In this setting, the entry of type B students has three main effects. First, educational expansion can congest university resources. Knowledge acquired in course c is a function of the quality of education, which depends positively on the amount of public resources assigned to the course (r_c) and negatively on the number of enrolled students (E_c). If resources do not vary with enrollment,[7] a marginal increase in the enrollment of type B students affects the quality of education according to

$$\frac{\mathrm{d}Q_c}{\mathrm{d}E_{B,STEM}} = \frac{\partial Q_c}{\partial E_{STEM}} \cdot (1 + \frac{\partial E_{A,STEM}}{\partial E_{B,STEM}}). \tag{3}$$

[7] This case fits the Italian scenario, where tertiary education was heavily subsidized by lump-sum transfers and tuition fees were low. For example, tuition fees were equal to 20 percent of total revenues at *Università Statale* and 11 percent at *Politecnico*. This assumption can be relaxed to allow for cases in which only a part of university resources respond to enrollment.

As total enrollment increases, resources are shared among a larger amount of students, access to university inputs becomes more crowded, and quality of education decreases ($\frac{\partial Q_c}{\partial E_{STEM}}$). The sign of equation (3), however, hinges on how total enrollment changes. Type A students, in fact, might decide to move out of STEM fields towards majors with restricted access after 1961, where quality of education is unaffected. The overall effect is negative if less than one type A student leaves STEM for each incoming type B student ($\frac{\partial E_{A,STEM}}{\partial E_{B,STEM}} > -1$).

Second, educational expansion can modify the composition of skills in a classroom by admitting students who differ from the incumbents in terms of their precollegiate qualification. In the resulting classroom, teaching might be more challenging and negative peer effects more likely (Lazear, 2001). For student i in course c, class heterogeneity is a function of the number of students of the same type ($E_{i,STEM}$)—in this case, with the same high school diploma—and the number of students of a different type ($E_{-i,STEM}$). A marginal increase in type B enrollment modifies class heterogeneity according to

$$\frac{\mathrm{dCH}_{ic}}{\mathrm{d}E_{B,STEM}} = \frac{\partial CH_{i,c}}{\partial E_{B,STEM}} + \frac{\partial CH_{i,c}}{\partial E_{A,STEM}} \cdot \frac{\partial E_{A,STEM}}{\partial E_{B,STEM}}. \tag{4}$$

For type A students in STEM majors, class heterogeneity increases after the reform: type B students enter and make classes more diverse ($\frac{\partial CH_{i,c}}{\partial E_{B,STEM}}$), while some type A students who would have chosen STEM might decide to enroll elsewhere to avoid the negative consequences of educational expansion ($\frac{\partial CH_{i,c}}{\partial E_{A,STEM}} \cdot \frac{\partial E_{A,STEM}}{\partial E_{B,STEM}}$).

Third, an increase in STEM enrollment might drive up the aggregate supply of STEM skills, which can decrease prices of STEM skills. The change in prices of STEM skills following a marginal increase in type B enrollment is

$$\frac{\mathrm{d}log(w_{STEM})}{\mathrm{d}E_{B,STEM}} = -\phi^{-1} \cdot \frac{1}{H_{STEM}} \cdot (\frac{\partial H_{STEM}}{\partial E_{B,STEM}} + \frac{\partial H_{STEM}}{\partial E_{A,STEM}} \cdot \frac{\partial E_{A,STEM}}{\partial E_{B,STEM}}) \tag{5}$$

with aggregate production Y fixed. As seen above, the sign depends on two offsetting changes. On the one hand, more type B students enroll in STEM fields and drive up the aggregate supply of STEM skills ($\frac{\partial H_{STEM}}{\partial E_{B,STEM}}$). On the other hand, some type A students switch to other fields and decrease the supply of STEM human capital ($\frac{\partial H_{STEM}}{\partial E_{A,STEM}} \cdot \frac{\partial E_{A,STEM}}{\partial E_{B,STEM}}$).

5 The Indirect Effects of Educational Expansion on Human Capital

This section describes the effects of educational expansion in STEM majors on the human-capital accumulation of type A students and addresses several concerns about the empirical strategy. The unit of analysis is a type A student i in a compulsory course c in period p. The outcome of interest is the standardized grade received by each student in different courses. The sample is restricted to type A students in order to estimate the indirect effects of educational expansion on the students who were supposed to be inframarginal to the reform. The sample is further restricted to compulsory courses (*insegnamenti fondamentali*) for two reasons. First, students could not opt out of compulsory courses, if they wanted to complete their degree. Second, because compulsory courses focused on the fundamental topics of each discipline, their suppression or any major change to their syllabi would have required a complex bureaucratic process.

The empirical strategy exploits variation along two dimensions. The first source of variation is time, because a large number of type B students enrolled in STEM majors after 1961. This wave of new students, however, affected courses in different academic years. First-year courses, for example, went through phase I of the reform (regulated entry of type B students) from 1961 to 1964 and through phase II (free entry) from 1965, while second-year courses went through phase I from 1962 to 1965 and through phase II from 1966. In the analysis, period 0 always identifies when a course received type B students for the first time. The second source of variation is based on the fact that the entry of type B students differentially affected crowding of university resources and peer effects between courses within a major.

5.1 Congestion of Teaching Resources: Setup and Results

Out of all possible university resources that enter the education production function, this paper focuses on the student–faculty ratio for several reasons. First, the student–faculty ratio is a measure of access to professors and teaching assistants, arguably one of the most important university inputs. Teaching assistants were especially important for pedagogy. While professors had mainly a lecturing role, teaching assistants held office hours, led review sections, supervised undergraduate theses (mandatory to graduate), carried out administrative tasks, helped with oral examinations, and sometimes delivered the main lectures (Marbach, Rizzi and Salvemini, 1969). In summary, teaching assistants were the best chance for students to receive assistance with coursework.

Second, the number of teaching fellows (professors and assistants) did not adjust to accommodate higher enrollment. Italian universities, in fact, were mostly funded through government grants, which were not modified in 1961. More students did not bring more resources to hire additional faculty. In Milan, the student–faculty ratio in compulsory STEM courses increased only slightly during the first phase of the reform (Appendix Figure C1).

On average, it increased by 3.36 student per teaching fellow or 7.5 percent in the first four periods after the policy implementation. This moderate increase, however, was the result of a larger positive change in engineering, physics, and the natural sciences (+5.15) and a small decrease in math and chemistry (-3.19). During the second phase of the reform, the student-faculty ratio increased in most STEM majors (+38.95).

Third, the student–faculty ratio could vary substantially between two otherwise similar compulsory courses (within the same major, within the same field of study, within the same curriculum year, and with similar enrollment). Some compulsory courses, in fact, were linked to an endowed chair (funded by the state) and to a permanent "tenured" professor. The decision to create an endowed chair (*cattedra*) was not in the direct control of the universities and required several governmental approvals. The other compulsory courses were taught by professors on temporary assignment. In this setting, "tenured" professors had more power and incentives to hire teaching assistants (Clark, 1977).[8] Data from universities in Milan confirm that a compulsory STEM course taught by a "tenured" professor had on average 3.2 additional teaching fellows, relative to other compulsory STEM courses (Appendix Table C1). All other course characteristics have limited explanatory power: for example, increasing enrollment by 10 units correlates to only 0.28 additional teaching fellows.

In spite of a similar enrollment increase, the courses with fewer assistants (and higher preexisting student–faculty ratio) experienced a larger surge in student–faculty ratio and more congestion of teaching resources after the reform (Appendix Figure C2). Assuming a negative correlation between the student–faculty ratio and the quality of education,[9] the average learning of type A students should have decreased more in courses with higher preexisting student–faculty ratio. To test this hypothesis, the baseline specification is:

$$g_{icp} = \alpha + \beta \frac{E_c}{fac_c} + \gamma_p + \sum_p \delta_p \left(\frac{E_c}{fac_c} \times P_p \right) + \eta X_{ip} + \zeta Z_{cp} + \kappa V_c + u_{icp}, \tag{6}$$

where g_{icp} is the standardized grade of student i in the compulsory course c in period p. $\frac{E_c}{fac_c}$ is the average preexisting student–faculty ratio in course c.[10] fac_c measures the number of teaching fellows assigned to course c, computed from the yearly records of the universities.

[8] "Tenured" professor sat in the hiring committees and were more likely to have their requests approved. More importantly, because they were permanently assigned to the same course, "tenured" professors had access to a much larger pool of former students to hire as assistants.

[9] This assumption is not directly testable, because quality of education is hard to measure. A piece of evidence that is consistent with—although only suggestive of— a negative relationship between quality of education and student–faculty ratio is the negative empirical correlation between average grades and student–faculty ratio (Appendix Figure C3).

[10] It is computed as the average student-faculty ratio between period -3 and period -1. The sample does not include the courses with a preexisting student–faculty ratio in the bottom or top decile to downplay the influence of outliers.

The actual attendance in each course, however, is not readily available. E_c, instead, measures the number of students enrolled in a given major and year (maximum attendance possible) multiplied by course-specific weights. Using the transcript data, the weights are computed as the share of students from Milan, who attempted the final exam, per course and period. Because many enrolled students were not actively attending classes (a well-known problem of the Italian university), controlling for participation is important to estimate the congestion of teaching resources. γ_p are period fixed effects. P_p is 1 for post-reform periods. X_{ip} are student characteristics: gender, a quadratic polynomial of age, high school fixed effects, measures of precollegiate ability, and major and university fixed effects. Z_{cp} are time-varying course characteristics: the "tenure" and gender of the professor, and a dummy variable for institute directors.[11] V_c are time-invariant course characteristics: institute and curriculum year fixed effects.

The estimates of δ_p, the parameter of interest, show that the grades of type A students decreased in courses with a higher pre-existing student–faculty ratio and more post-reform congestion of teaching resources, relative to other courses in the same major (Table 2, column 1). During the first phase of the reform, the effect of increased congestion is precisely estimated, but small: a marginal increase in the student–faculty ratio decreased grades by 0.0013σ. Since the student–faculty ratio increased by 3.36 units during the first phase, grades decreased by 0.004σ (s.e.=0.001) in the average STEM course. During the second phase of the policy, the magnitude of the effect becomes larger. After the free entry of type B students, grades of type A students decreased by 0.091σ in the average STEM course. Because these estimate measure lower learning in one course only, the congestion of teaching resources might have led to much larger losses over the course of a STEM degree, where compulsory courses were usually more than 10. These results are not driven by the fact that some students opted out of courses with more congestion after the reform (Table 2, column 2). Including student fixed effects, the grades of type A students decreased by 0.006σ (s.e.=0.001) during phase I and by 0.098σ (s.e.=0.015) during phase II.

Yearly estimates confirm that the crowding effect was negative and small during phase I of the reform. During phase II, however, the coefficients become more negative, as an increasing number of new entrants led to more congestion of teaching resources (Appendix Figure C4). These effects describe how learning changed in the majority of compulsory courses, where the student–faculty ratio increased after the reform. In the cases where the students–faculty ratio did not grow (for example, because resources increased), grades of type A students increased by 0.031σ (s.e.=0.020) during phase I and by 0.002σ (s.e.=0.012) during phase II (Appendix Table C3, row 2). Neither coefficient is statistically significant.

[11] An institute is a bureaucratic entity that groups courses within a major in the same field of study.

Similarly, the effects are larger in magnitude in engineering, physics, and the natural sciences, where the student–faculty ratio increased the most. In these majors, grades decreased by -0.009σ (s.e.=0.002) during phase I and by -0.135σ (s.e.=0.021) during phase II (Appendix Table C3, row 3).

5.2 Class Heterogeneity: Setup and Results

Relative to the pre-reform years when all students in STEM majors had a type A diploma, the educational expansion increased significantly the heterogeneity of STEM classes. At the time of university enrollment, in fact, type A and B students differed greatly in their scientific knowledge (Appendix Table B1). While type A students had studied formal, physical, and life sciences (for example, math, physics, chemistry, and biology) during high school, type B students had studied only applied sciences such as mechanics, basic engineering, and topography.

The empirical analysis exploits the fact that the composition of enrolled students changed in all STEM majors after 1961, but the negative effects on learning were larger in courses that were not included in the high school curriculum of type B students. Since type B students were less well prepared in these fields, their entry was more likely to disrupt teaching practices and student-to-student interactions.[12] Engineering students, for example, had to take "Technical Drawing" and "Mathematics I" in their freshman year. Type B students were trained in technical drawing, but not prepared for university-level math. Unlike technical drawing, in fact, math was not part of the curriculum in type B schools. To isolate the effect of increased class heterogeneity, then, this paper measures how the grades of type A students changed after the reform in STEM compulsory courses where type B students were not prepared (like "Mathematics I"), relative to other courses in the same major (like "Technical Drawing"):

$$g_{icp} = \alpha + \beta Not\ in\ B\ cv_c + \gamma_p + \sum_p \delta_p \left(Not\ in\ B\ cv_c \times P_p\right) + \eta X_{ip} + \zeta Z_{cp} + \kappa V_c + u_{icp}, \quad (7)$$

where $Not\ in\ B\ cv_c$ is equal to 1 if course c was not included in the high school curriculum of type B students. Differently from equation (6), Z_{cp} now includes also the student–faculty ratio per course and period. δ_p, then, measures the grade change driven by higher class heterogeneity after the reform, keeping access to teaching resources fixed.

[12] Once type B students were admitted into STEM majors, their grades in the STEM courses not included in their precollegiate curriculum were 0.243σ lower, compared with other STEM courses (Appendix Table D2). In these same courses, grades of type B students were 0.279σ lower than grades of type A students.

In courses where type B students were less well prepared, grades of type A students decreased by 0.076σ (s.e.=0.020) in phase I and by 0.165σ (s.e.=0.022) in phase II, relative to other compulsory courses (Table 3, column 1). These findings cannot be explained by type A students opting out of courses where type B students were less prepared, because the sample is restricted to compulsory courses (Table 3, column 2). Including student fixed effects in equation (7), grades decreased by 0.104σ (s.e.=0.019) in phase I and by 0.176σ (s.e.=0.022) in phase II.

Differently from what observed for congestion, yearly estimates show that the effect of class heterogeneity is negative and large already in phase I (Appendix Figure D1). The magnitude of the coefficients increases substantially during phase II, when more type B students enrolled in STEM majors. These findings might indicate that even a moderate entry of students with different precollegiate qualification (phase I) can lead to negative peer effects on the inframarginal students. These results, however, cannot identify the channel through which the negative peer effects operated. Lower human-capital accumulation, for example, could be the result of less effective teaching methodologies or student-to-student negative externalities.

5.3 Additional Findings and Robustness Checks

Accounting for Historical Trends in Enrollment

Starting from the end of World War II, an ever-growing number of students enrolled in high school and university. Although this paper takes advantage of a sharp increase in university STEM enrollment, the underlying trend in education might bias the previous results. This section attempts to control for these historical changes in school participation by including different trends in equations (6) and (7).

The baseline findings would overstate the effects of the policy, if the additional students who enrolled in school over time received lower grades in courses with more congestion and higher class heterogeneity. To address this concern, the baseline regressions are modified to include pre-reform trends for courses with more congestion (Table 2, column 3) and higher class heterogeneity (Table 3, column 3). The pre-reform trends are not statistically significant, while the effects of crowding and peer effects are close to the baseline. These results indicate that the average grades in courses differentially affected by the policy were not on different trajectories in the pre-reform period, corroborating the hypothesis the the increasing school participation does not bias the baseline estimates.

The results are robust to the inclusion of other trends. Linear major-specific trends can control for differential entry of type A students over time and across STEM fields. In courses with more congestion of teaching resources, grades decreased by 0.006σ (s.e.=0.001) in phase I and by 0.100σ (s.e.=0.015) in phase II (Appendix Table C3, row 6). In courses where type B students were less well prepared, grades decreased by 0.105σ (s.e.=0.019) in phase I and by 0.179σ (s.e.=0.022) in phase II (Appendix Table D4, row 2). All estimates are close to the baseline.

The inclusion of a linear trend for each quartile of the precollegiate ability distribution (computed using the high school grades) can control for differential entry of type A students over time and across ability. These trends are important, if the increase in school participation was driven primarily by low-achieving students. Due to crowding, grades decreased by 0.006σ (s.e.=0.001) in phase I and by 0.098σ (s.e.=0.015) in phase II (Appendix Table C3, row 7). Due to higher class heterogeneity, grades decreased by 0.104σ (s.e.=0.019) in phase I and by 0.176σ (s.e.=0.022) in phase II (Appendix Table D4, row 3). Also in this case, the estimates are close to the baseline.

The results are robust also to the inclusion of quadratic major trends, high school trends, or high school–ability trends.

Did the Reform Change the High School Choice?

The reform increased the value of a type B diploma and gave students an alternative path to pursue a STEM degree. The policy, therefore, might have affected how students sorted into type A and B schools. For example, some students who would have enrolled in a type A school without the policy might have chosen a type B school. To gauge whether changing selection into high school drives the previous results, the sample can be restricted to cohorts who could have not responded along this dimension. The first cohort who enrolled in high school after the reform graduated in 1966. The cohorts graduating between 1961 and 1965 could have not modified their high school choice after the reform, because transfers across types of schools were not possible.[13]

Using the sample of type A students who graduated before 1966, grades decreased by 0.006σ (s.e.=0.001) in phase I and by 0.102σ (s.e.=0.016) in phase II in course with more congestion of teaching inputs (Table 2, column 5). In courses where type B students were less well prepared, grades decreased by 0.106σ (s.e.=0.019) in phase I and by 0.151σ (s.e.=0.027) in phase II (Table 3, column 5). The coefficients for phase I are significantly lower (larger in magnitude) than the baseline. Although these estimates do not reveal whether the sample of students who graduated after 1965 sorted differently into high school, they suggest that the

[13]Strong anticipation effects were unlikely. A search on the historical archive of an Italian newspaper (*La Stampa*) reveals that, less than a month before the final ratification, there was uncertainty about the content of the reform and the odds of approval (link).

negative effects of congestion and class heterogeneity on grades cannot be explained solely by changing selection into high school.

Did the Reform Change the Major Choice?

The reform might have induced some type A students, who would have otherwise enrolled in STEM majors, to choose a major with restricted access. This behavior could be explained by multiple, non-mutually exclusive explanations. The entry of type B students, in fact, decreased learning in STEM courses through crowding of university resources and negative peer effects. Moreover, the reform increased the competition for STEM jobs. Finally, the entry of type B students might have weakened the signal of high aptitude provided by STEM degrees to the labor market.

There are several pieces of evidence suggesting that the reform might have affected the major choice of some type A students. In Milan, the share of type A students enrolling in STEM and restricted majors followed a diverging trend after the reform (Figure 2, Panel A). In 1958, 39.6 percent of type A enrolled in STEM, while 37.7 percent enrolled in restricted majors. The two shares stayed constant until 1961. After 1961, however, the share of type A students enrolling in STEM started decreasing, while the share of type A students choosing a restricted major increased. To control for student characteristics, the major choice of type A students can be modeled as a multinomial logit.[14] The estimated cohort effects follow the same diverging trends shown by the raw probabilities (Figure 2, Panel B). Although these cross-cohort comparisons cannot isolate the causal effect of the reform, these findings are robust to additional tests (Appendix E) that control for changes in the characteristics of type A students (for example, the increasing female participation in the university) or for exogenous changes in the returns to different majors.

If the reform affected the selection of type A students into STEM majors, a variation over time in the estimating sample might bias the effects of congestion and class heterogeneity on grades. To address this concern, the sample can be restricted to type A students who were less likely to select out of STEM majors after the reform. Using the same multinomial logit model described above, the major choice of type A students can be estimated on pre-reform observations to predict the behavior of students who enrolled after 1961. The predicted and actual major choices, then, can be compared to identify the students who might have abandoned the STEM majors (Appendix Table E2). Not surprisingly, graduates of scientific type A schools and low-achieving students (scoring in the bottom quartile of the high school

[14]The choice is between STEM, restricted, and nonrestricted majors, and no university as the baseline. The model controls for year of high school graduation (cohort) fixed effects, as well as the same student characteristics included in equations (6) and (7).

exam) were less likely to select out of STEM majors after the reform. These students, in fact, either had strong preference for STEM disciplines (scientific graduates) or lacked the skills to succeed in a wide array of programs (low-achieving students).

Based on these considerations, the sample is further restricted to low-achieving scientific students. Due to congestion of teaching resources in phase I, the grades of low-achieving scientific students decreased by 0.005σ (s.e.=0.003) in the average STEM course (Table 2, column 6). This effect is significant at the 10 percent level. During phase II, the grades of low-achieving scientific students decreased by 0.120σ (s.e.=0.036) in the average STEM course. Due to higher class heterogeneity, the grades of low-achieving scientific students decreased by 0.131σ (s.e.=0.048) during phase I and by 0.174σ (s.e.=0.051) during phase II (Table 3, column 6). The coefficients are statistically different from zero, but not from the baseline estimates.

Alternatively, the post-reform observations can be reweighted to match the average characteristics of the pre-reform sample. This technique, which essentially generate sampling weights from propensity scores, can account for sample selection based on observables (DiNardo, Fortin and Lemieux, 1996). Using the sampling weights, grades decreased by 0.006σ (s.e.=0.001) in phase I and by 0.102σ (s.e.=0.015) in phase II in course with more congestion of teaching inputs (Table 2, column 7). In courses where type B students were less well prepared, grades decreased by 0.113σ (s.e.=0.021) in phase I and by 0.179σ (s.e.=0.024) in phase II (Table 3, column 7). The coefficients are significantly different from zero and close to the baseline.

In summary, there is suggestive evidence that some type A students might have selected out of STEM majors after the reform, especially humanities graduates and high-achieving students. The baseline results appear robust to tests that restrict the sample to students who were less likely to change their major choice, and that reweight the sample to keep average observables constant.

Accounting for Professor Fixed Effects

What is the role of professors' characteristics? "Untenured" professors, for example, were more likely to teach courses with high preexisting student–faculty ratio, where crowding became more severe after the reform. The effect attributed to congestion, then, might depend on the inability of "untenured" professors to deal with increased enrollment after the reform. In this case, the channel through which higher enrollment decreased learning could have been larger class size, not congested access to teaching resources. Because some professors were teaching one course "with tenure" (linked to an endowed chair) and other courses "without tenure" (not linked to an endowed chair) during the same academic year, the inclusion of professor-fixed effects addresses this concern. Due to increased congestion of teaching resources, grades decreased by 0.006σ (s.e.=0.001) during phase I and by 0.124σ (s.e.=0.024) during phase II in the average STEM course (Table 2, Columns 4). Similarly for class heterogeneity, results are robust to the inclusion of professor fixed effects. In courses were type B students were less prepared, the grades of type A students decreased by 0.135σ (s.e.=0.022) during phase I and by 0.261σ (s.e.=0.030) during phase II (Table 3, column 4). All coefficients are statistically different from zero and more negative than the baseline estimates.

Did the Reform Affect Grading?

The entry of type B students in STEM courses might have affected how grades were assigned. In this case, the results might not reflect actual variations in human capital, but just changes in the composition of university classes. In the Italian university, however, exams were not intended to be graded on a curve. In the period under consideration, assigning grades on a curve would have been especially difficult, because most courses had oral tests and students were graded sequentially. Moreover, the empirical analysis focuses on compulsory courses, in which substantial revisions of the syllabi were less likely.

Even in this scenario, however, professors might have started grading more leniently after the reform to avoid failing a large share of students. If the change in grading was the same across all courses within a major, the difference-in-differences setup would cancel this effect. If, instead, grading became more lenient in courses that experienced more congestion or higher class heterogeneity (because there were more students struggling), the baseline estimates should be interpreted as upper bounds of the negative effect of crowding and class heterogeneity on human capital. The opposite scenario, harsher grading after the reform, is less credible, because students incurred a low penalty from failing an exam. Failed students could just retake the exam in one of the many sessions spread throughout the year.

In summary, it is not possible to rule out the hypothesis that professors changed their grading policies after the reform. Under the most plausible scenarios, the baseline estimates should either be unaffected or treated as upper bounds of the true negative effects.

6 Long-run Income

If the educational expansion decreased student learning in STEM majors, the policy might have lowered the returns to STEM degrees. This section uses the long-run income of the (former) students to address this question.

6.1 Income Tax Returns in 2005

The income from personal income tax returns in 2005 is the sum of all individual earnings that are taxable under the Italian personal income tax after allowed deductions. It includes labor earnings for the employees, profits for the self-employed, pensions, rents, and interests. The main excluded categories are dividends and capital gains, both taxed separately. Because the release of the 2005 income tax data was an extraordinary event, similar data from other years are not available.[15]

Using name and birthdate, I uniquely match 83 percent of the high school graduates to income earners in year 2005, when the estimated survival rate for these cohorts was equal to 91 percent.[16] Matched individuals are more likely to be men and have higher high school grades (Appendix Table A1).[17] These findings suggest that the attrition in matching students to income earners might be primarily driven by individuals (women and low-achieving students) with income below the minimum taxable threshold (€7,500), and not by high-skilled expatriates. Importantly, selective attrition based on gender and school achievements does not vary across cohorts (Appendix Table A2).

In 2005, 95 percent of the students in the dataset were between 56 and 67 years old, when the returns to education were realized in full and income was less likely to be affected by temporary shocks. Having only one cross section, however, limits the analysis in two dimensions. First, the effect of the policy on income dynamics cannot be examined. Second, in a single cross section, any comparison between students from different cohorts combines age and cohort effects. That is, after controlling for observable characteristics, the incomes of two otherwise similar students in different cohorts could differ both because they completed high school in different years (cohort effects) and because their ages were different in 2005 (age effects).

To take age effects into account, this paper uses repeated cross sections of the SHIW data to estimate age effects as a function of observable characteristics, allowing for cohort and year effects (Appendix F). Owing to the fact that cohort, age, and year effects cannot be freely estimated, this paper assumes that year effects are smooth, because the period under consideration does not contain sharp year events (like a war). The results, however, are robust to different assumptions. The resulting out-of-sample estimates of age effects are used to predict income at age 65 (the age of retirement for men) for all the in-sample

[15] In March 2008, the Italian Treasury published the 2005 income tax data (with identifiers like name, birthdate, and city of residence) on its website to facilitate the diffusion of income data. The data files were removed within 24 hours due to strong opposition from the public. If downloaded in March 2008, the data can now be used for research purposes.

[16] The estimated survival rate can be found at http://www.mortality.org.

[17] The gender difference is statistically significant, but small (81.8 percent for women and 83.3 for men).

observations. This paper, then, relies on this age-adjusted income to measure the effects of the policy.

Adjusting income to account for the age effects, however, is not crucial for the analysis. Age effects for older Italian workers, in fact, are small, both because their preretirement earnings are hardly rising with age and because their postretirement earnings tend to be a strict percentage of their pay in their last several years before retirement. In this case, the rigidity of the Italian pay and pension systems (Holden and Wulfsberg, 2008) is useful.

6.2 Income of Type A Students

The main results of this paper indicate that the congestion of teaching resources and the negative peer effects reduced human capital in STEM majors. This loss of skills might have reduced the returns to a STEM degree in the long run. To test this hypothesis, the baseline specification examines how the long-run income of type A students with a STEM degree changed after reform, relative to type A students with other university degrees:

$$log(\text{income}_{imt}) = \alpha + \beta \text{STEM}_m + \gamma_t + \sum_t \delta_t [\text{STEM}_m \times Post_t] + \zeta X_{imt} + u_{imt}, \qquad (8)$$

where the unit of observation is a student i, who completed high school in year t and chose major m after high school. The dependent variable income_{imt} is the taxable personal income in 2005, $Post_t$ is 1 if a student completed high school after the implementation of the reform (1961), STEM_m is 1 if a student received a STEM degree, γ_t are cohort fixed effects, and X_{imt} is the usual set of student characteristics with the addition of the final degree mark and the year of university graduation.

The estimation of δ_t suggests that the income of type A students with a university STEM education increased by 1.7 percent (s.e.=0.081) from 1961 to 1964 (Table 4, column 1).[18] The coefficient is not statistically significant. During the first phase of the reform, the limited increase in congestion and class heterogeneity could explain the lack of negative effects on the long-run income of type A students. Between 1965 and 1968, however, the income of type A students with a STEM education decreased by 12.9 percent (s.e.=0.075). This estimate is significant only at the 10 percent level. Overall, there is weak evidence that the income of type A students with a STEM degree decreased after 1965.

The previous estimates could be biased, if sorting into different programs changed after the reform. The post-World War II increase in school participation, for example, might have

[18] In this context, the phrase "from 1961 to 1964" refers to the cohorts who completed high school from 1961 to 1964.

induced students with lower qualifications to enroll in high school and university. If these new students were more likely to choose a STEM major, the income premium of STEM degrees would have decreased over time. A pre-reform trend for type A students with a STEM degree is not statistically significant, suggesting that the graduates of STEM majors and other programs were on the same income trajectory before the reform (Table 4, column 2). The inclusion of this trend, however, increases the magnitude of the estimated effects. The income of students with a STEM degree decreased by 21.7 percent (s.e.=0.181) from 1961 to 1064 and by 32.8 percent (s.e.=0.178) from 1965 to 1968. Only the second coefficient is statistically different from zero. Similarly, the inclusion of a trend for each quartile of the precollegiate ability distribution (measured using the high school grades) can control for linear income variations over time and across ability groups (Table 4, column 3). After adding a linear trend for each ability level, the income of type A students with a STEM education increased by 1.5 percent (s.e.=0.081) from 1961 to 1964 and decreased by 13.4 percent (s.e.=0.074) from 1965 to 1968, relative to other type A students. The baseline findings are robust also to the inclusion of ability-specific quadratic trends and linear and quadratic trends for each combination of high school and quartile of the precollegiate ability distribution (Appendix Table G1).

An important concern is that some type A students might have responded to increased enrollment in STEM majors by choosing a different program. To assess the influence of this form of selection on the baseline findings, the sample can be restricted to low-achieving (scoring in the bottom quartile of the high school exam) type A students, who were less likely to change major after the policy (Table 4, column 4). The income of low-achieving type A students with a STEM degree decreased by 17.8 percent (s.e.=0.155) from 1961 to 1964 and by 34.4 percent (s.e.=0.141) from 1965 to 1968. The second coefficient is statistically significant and lower than the baseline. These results suggest that the changing selection into university majors might bias toward zero the baseline estimates. As an additional piece of evidence, an intent-to-treat analysis compares how the income of low-achieving type A students from scientific schools changed after the reform, relative to low-achieving type A students from humanistic schools (Table 4, column 6). On the one hand, in fact, low-achieving scientific students had a high probability of enrolling in STEM majors (42.3 percent) and a relatively low predicted propensity to leave after the reform (Appendix Table E2). On the other hand, very few low-achieving students from humanistic schools enrolled in STEM majors even before the reform (16.9 percent). The income of low-achieving scientific students decreased by 25.1 percent (s.e.=0.164) from 1961 to 1964 and by 25.4 percent (s.e.=0.146) from 1965 to 1968. The coefficients are statistically significant at the 10 and 5 percent level, respectively.

Alternatively, the post-reform samples of type A students in STEM majors and other programs can be reweighted to match the average observable characteristics of the two pre-reform samples (Table 4, column 5). The estimation of equation (8) with these sample weights indicates that the income of type A students with a STEM education increased by 1.4 percent (s.e.=0.082) from 1961 to 1964 and decreased by 14.5 percent (s.e.=0.076) from 1965 to 1968.

According to the proposed model of educational expansion, the marginal log income change generated by the reform is the sum of the marginal change in the log skill prices and the marginal change in the log human capital. The estimated effects of congestion of university resources (section 5.1) and higher class heterogeneity (section 5.2) are used to compute the change in learning generated by the educational expansion per student and course (leave-one-out estimator, Abadie, Chingos and West, 2014). For each student, the learning changes in each university course are summed to compute the total variation in human capital. This procedure indicates that the congestion of university resources in STEM majors might have decreased income of low-achieving students by 12.9 percent after 1965. Similarly, higher class heterogeneity might have decreased income of low-achieving students by 12.6 percent after 1965. Since the income of low-achieving type A students with a STEM education declined by 34.4 percent from 1965 to 1968, lower human capital could account for three quarters of the lost returns to a STEM degree. In the absence of other effects, lower skill prices (due to a higher supply of STEM-educated workers) would account for the remaining share.

6.3 Income of Type B Students

The reform was very successful in expanding access to the university among type B students. University graduation rates of type B students increased by 5.9 percentage points from 1961 to 1964 and by 15 percentage points from 1965 to 1968, relative to an average graduation rate of 8.2 percent before 1961 (Appendix Table G3). In the absence of negative indirect effects of educational expansion, the average income of type B students should have increased with the level of completed education. If, however, the educational expansion led to a loss of human capital in STEM majors (and, potentially, more competition for STEM jobs), the reform might have reduced the returns to STEM degrees for type B students.

To test this hypothesis, the empirical analysis compares how the long-run income of type B students changed after the reform, relative to type C students. Type C students attended technical schools that prepared for jobs in accounting and the service sector. Before the reform, type B and C students had access to the same restricted group of university majors.

Despite their similarities with type B students, type C students were not allowed to enroll in university STEM majors in 1961. The baseline regression is:

$$log(\text{income}_{ist}) = \alpha + \beta \text{Type B}_s + \gamma_t + \sum_t \delta_t [\text{Type B}_s \times Post_t] + \zeta X_{ist} + u_{ist}, \tag{9}$$

where the unit of observation is a student i, who completed high school s in year t. Type B_s is 1 if a student attended a type B school.

The long-run income of type B students increased by 14.6 percent (s.e.=0.066) from 1961 to 1964, relative to type C students. During phase II, however, the income of type B students increased by only 6.1 percent (s.e.=0.066), relative to type C students (Table 5, column 1). This estimate is not statistically different from zero or from the phase-I coefficient. This finding suggests that, in spite of higher rates of university graduation, the average income of type B students did not increase after 1965 (relative to the income of pre-reform cohorts or phase-I cohorts). The results are robust to alternative difference-in-differences specifications (Appendix G). Although they are based on different identification strategies, the analyses of the long-run incomes of type A and B students suggest that the reform might have decreased the returns to STEM degrees after 1965.

The inclusion of different trends to account for the historical increase in school participation does not affect the main findings. A pre-reform trend for type B students is not statistically significant, suggesting that type B and C students were not on different income trajectories before the reform (Table 5, column 2). Adding a trend for each quartile of the precollegiate ability distribution, the income of type B students increased by 14.1 percent (s.e.=0.066) from 1961 to 1964 and by 5.8 percent (s.e.=0.066) from 1965 to 1968, relative to type C students (Table 5, column 3). Both coefficients are close to the baseline estimates. Including a trend for each combination of high school and ability level, income of type B students increased by 9.6 percent (s.e.=0.109) from 1961 to 1964 and decreased by 4.3 percent (s.e.=0.192) from 1965 to 1968 (Table 5, column 4). The coefficients are not statistically significant. Overall, the main findings appear robust to the inclusion of different trends.

In addition to historical changes in the student population, some students, who would have chosen a type C school in the absence of the reform, might have decided to enroll in type B schools after the reform. To control for this type of selection, the sample is restricted to the pre-1966 cohorts who chose a high school without knowing about the policy (Table 5, column 5). Within this restricted sample, the income of type B students increased by 14.6 percent (s.e.=0.067) from 1961 to 1964 and decreased by 10.3 percent (s.e.=0.095) in 1965. The estimate for the 1965 cohort is not statistically different from zero and is lower than the phase-I coefficient.

To control for other forms of sample selection based on observables, the post-reform observations can be reweighted to match the average characteristics of the pre-reform sample (Table 5, column 6). Using sample weights, the income of type B students increased by 14.5 percent (s.e.=0.066) from 1961 to 1964 and by 9.3 percent (s.e.=0.066) from 1965 to 1968. The second estimate is not statistically different from zero or from the phase-I coefficient.

Overall, there is not strong evidence that the average long-run income of type B students increased after 1965, in spite of a remarkable surge in the rates of university graduation. The indirect effects of educational expansion (congestion of teaching resources, negative peer effects, lower skill prices), in fact, might have reduced the returns to STEM degrees. Two alternative explanations for these findings, stigma in the labor market and the timing of graduation, appear less plausible in this context. First, employers might have discriminated against all type B students, regardless of their education. Indeed, a type A diploma is associated with 36.1 percent (s.e.=0.022) higher income, compared with a type B diploma. However, the type of high school loses most of its predictive power after controlling for university graduation (coefficient=0.028, s.e.=0.038). Second, the post-1965 cohorts might have entered the labor force during the 1973 oil crisis. The literature that examines the costs of graduating in a recession, however, finds that the negative effects fade out a few years after graduation and do not apply to graduates from high-return majors like STEM (Oreopoulos, von Wachter and Heisz, 2012; Altonji, Kahn and Speer, 2014).

7 Conclusions

This paper uses an education reform to study how increasing access to education affects human-capital accumulation in expanded programs. Using data from high school registries, university transcripts, and long-run income for 27,236 Italian students who completed high school in Milan between 1958 and 1968, this paper finds that the educational expansion lowered learning in STEM majors through congestion of teaching resources and higher class heterogeneity. Moreover, there is evidence that these effects may have induced some of the best students who would have otherwise enrolled in STEM majors to choose other programs. These results provide new insights on the education production function of university students. Finally, the analysis of long-run incomes indicates that the educational expansion might have had long-lasting negative effects on the returns to STEM degrees.

This paper focuses on the cohorts who enrolled in the university immediately after the reform. Did the negative indirect effects of educational expansion persist among the following cohorts? The government, for example, could have increased the resources of public universities to reduce congestion. Some adverse effects of the reform, however, might have lingered. For example, hiring new faculty would not have necessarily restored the pre-policy level of education quality because the new professors would have been trained in the "expanded" university system and, therefore, would have had lower human capital.[19] In addition, if the entry of type B students in STEM fields caused the signal provided by STEM degrees to deteriorate, especially talented type A students could have deserted STEM fields even if the expansion had been more gradual.

This paper exploited few features of the Italian university system to achieve identification, but the estimated effects apply more broadly. If the effects of increased congestion and higher class heterogeneity affect student learning similarly in different countries and over time, the results of this paper could inform about the unintended consequences of many reforms that attempted to increase the number and diversity of university students. In 1999, for example, China increased the admission slots in its universities and enrollment more than doubled in a few years. Other Asian countries like India, Malaysia, and Sri Lanka implemented large-scale affirmative action plans that dramatically changed the composition of university students (Sowell, 2005). All these reforms could have led to overcrowded universities or negative peer effects. Nor is the US necessarily immune by large educational expansions. The recent "America's College Promise Act", for example, promises to grant all US students access to free community college.

What are the broader implications of this paper for the design of education policies? Public intervention is potentially needed to overcome market failures that may lead to suboptimal human-capital investment. The solution to human-capital investment failures, however, is not greatly expanded education provision that attempt to force students into certain fields in state-controlled universities. Indeed, "in-kind" provision of university education by state schools can have distortionary, even perverse, effects on human-capital accumulation relative to the same resources being directed toward student-specific financial aid.[20] Peltzman (1973) and others (Long, 2004; Cellini, 2009; Cohodes and Goodman, 2014) writing about the disconnect between what economic logic suggests and what states often do generally focus on the US case, but in fact the issue is probably minimized in America where state universities compete with a robust private sector (Aghion et al., 2010). In most countries, state universities lack competition so that any distortion they introduce via in-kind

[19] A basic cost–benefit analysis indicates that hiring new faculty in 1961 to prevent any congestion would have been beneficial (Appendix H). This strategy, however, would have required a large stock of available professors in Italy.

[20] If private universities offer higher-quality education and students are willing to forego quality in exchange for cheaper (but lower-quality) education, in-kind subsidies to state universities might lead to an overall decrease in quality-adjusted human capital.

provision is unlikely to be offset by private universities. In short, countries that currently aspire to increase university-level human capital may be reminded by Italy's example that the design of public interventions can magnify the negative effects of educational expansions.

References

Abadie, Alberto, Matthew M. Chingos, and Martin R. West. 2014. "Endogenous Stratification in Randomized Experiments." NBER Working Paper 19742.

Abbott, Brant, Giovanni Gallipoli, Costas Meghir, and Giovanni L. Violante. 2013. "Education policy and intergenerational transfers in equilibrium." NBER working paper 18782.

Aghion, Philippe, Mathias Dewatripont, Caroline M. Hoxby, Andreu Mas-Colell, and André Sapir. 2010. "The governance and performance of universities: evidence from Europe and the US." *Economic Policy Journal*, 25(01): 7–59.

Altonji, Joseph G., Lisa B. Kahn, and Jamin D. Speer. 2014. "Cashier or Consultant? Entry Labor Market Conditions, Field of Study, and Career Success." NBER Working Paper 20531.

Anelli, Massimo, and Giovanni Peri. 2013. "The Long Run Effects of High-School Class Gender Composition." NBER working paper 18744.

Bound, John, and Sarah Turner. 2007. "Cohort crowding: How resources affect collegiate attainment." *Journal of Public Economics*, 91(5-6): 877–899.

Cellini, Stephanie Riegg. 2009. "Crowded Colleges and College Crowd-Out: The Impact of Public Subsidies on the Two-Year College Market." *American Economic Journal: Economic Policy*, 1(2): 1–30.

Clark, Burton R. 1977. *Academic Power in Italy. Bureaucracy and Oligarchy in a National University System.* Chicago:The University of Chicago Press.

Cohodes, Sarah R., and Joshua S. Goodman. 2014. "Merit Aid, College Quality, and College Completion: Massachusetts' Adams Scholarship as an In-Kind Subsidy." *American Economic Journal: Applied Economics*, 6(4): 251–285.

Cooley, Jane. 2010. "Desegregation and the Achievement Gap: Do Diverse Peers Help?" Working Paper.

De Giorgi, Giacomo, Michele Pellizzari, and William Gui Woolston. 2012. "Class Size and Class Heterogeneity." *Journal of the European Economic Association*, 10(4): 795–830.

DiNardo, John, Nicole M. Fortin, and Thomas Lemieux. 1996. "Labor Market Institutions and The Distribution of Wages, 1973 - 1992: A Semiparametric Approach." *Econometrica*, 64(5): 1001 – 1044.

Duflo, Esther. 2004. "The medium run effects of educational expansion: Evidence from a large school construction program in Indonesia." *Journal of Development Economics*, 74(1): 163–197.

Duflo, Esther, Pascaline Dupas, and Michael Kremer. 2011. "Peer Effects, Teacher Incentives, and the Impact of Tracking : Evidence from a Randomized Evaluation in Kenya." *American Economic Review*, 101(5): 1739–1774.

Figlio, David N., and Marianne E. Page. 2002. "School Choice and the Distributional Effects of Ability Tracking: Does Separation Increase Inequality?" *Journal of Urban Economics*, 51(3): 497–514.

Heckman, James J., Lance J. Lochner, and Christopher Taber. 1998*a*. "Explaining Rising Wage Inequality: Explorations with a Dynamic General Equilibrium Model of Labor Earnings with Heterogeneous Agents." *Review of Economic Dynamics*, 1(1): 1–58.

Heckman, James J., Lance J. Lochner, and Christopher Taber. 1998*b*. "General-Equilibrium Treatment Effects: A Study of Tuition Policy." *American Economic Review*, 88(2): 381–386.

Holden, Steinar, and Fredrik Wulfsberg. 2008. "Downward Nominal Wage Rigidity in the OECD." *The B.E. Journal of Macroeconomics*, 8(1): 1935–1690.

Hoxby, Caroline M., and Gretchen Weingarth. 2006. "Taking Race Out of the Equation: School Reassignment and the Structure of Peer Effects." NBER Conference Paper.

Istituto Nazionale di Statistica. n.d.. *Annuario Statistico dell'Istruzione Italiana. Anni 1955-1975.* Roma.

Lavy, Victor, M. Daniele Paserman, and Analia Schlosser. 2012. "Inside the Black Box of Ability Peer Effects: Evidence from Variation in the Proportion of Low Achievers in the Classroom." *Economic Journal*, 122(559): 208–237.

Lazear, Edward P. 2001. "Educational Production." *Quarterly Journal of Economics*, 66(3): 777–803.

Lee, Donghoon. 2005. "An estimable dynamic general equilibrium model of work, schooling, and occupational choice." *International Economic Review*, 46(1): 1–34.

Lee, Donghoon, and Kenneth I. Wolpin. 2006. "Intersectoral labor mobility and the growth of the service sector." *Econometrica*, 74(1): 1–46.

Long, Bridget Terry. 2004. "Does the format of a financial aid program matter? The effect of state in-kind tuition subsidies." *The Review of Economics and Statistics*, 86(3): 767–782.

Marbach, Giorgio, Alfredo Rizzi, and Tommaso Salvemini. 1969. *Gli assistenti universitari e i liberi docenti in Italia.* Roma:Istituto di statistica, Facoltà di scienze statistiche, demografiche ed attuariali, Università degli studi di Roma.

Oreopoulos, Philip, Till von Wachter, and Andrew Heisz. 2012. "The Short- and Long-Term Career Effects of Graduating in a Recession." *American Economic Journal: Applied Economics*, 4(1): 1–29.

Peltzman, Sam. 1973. "The Effect of Government Subsidies-in-Kind on Private Expenditures: The Case of Higher Education." *Journal of Political Economy*, 81(1): 1–27.

Sowell, Thomas. 2005. *Affirmative Action Around the World. An Empirical Study.* Yale University Press.

Stapleton, David C., and Douglas J. Young. 1988. "Educational attainment and cohort size." *Journal of Labor Economics*, 6(3): 330–61.

TABLES AND FIGURES

Table 1: Summary Statistics

	Type A	Type B	Type C
Male	0.668	0.986	0.568
Birth year	1944	1943	1943
High school			
HS exit score (6-10)	6.48	6.36	6.38
Home schooled	0.072	0.084	0.169
Non-repeater	0.903	0.918	0.918
College			
Enrolled	0.867	0.408	0.402
Enrolled—STEM major	0.359	0.204	0.002
Enrolled—restricted major	0.465	0.045	0.037
Enrolled—nonrestricted major	0.043	0.159	0.363
University degree	0.637	0.161	0.105
University degree—STEM major	0.269	0.101	0.001
University degree—restricted major	0.353	0.029	0.016
University degree—nonrestricted major	0.015	0.031	0.088
Grades (18–31)—STEM major	23.93	24.56	21.79
Grades (18–31)—restricted majors	25.64	24.82	25.12
Grades (18–31)—nonrestricted majors	22.80	23.19	22.86
Income in 2005 (€)			
Income	58,657	47,628	41,892
Adjusted Income	65,749	53,812	48,095
Log Income	10.55	10.48	10.04

Notes: Summary statistics of students that completed high school in Milan, Italy; 1958–1968. The sample is composed of 11,433 type A, 8,813 type B, and 6,690 type C. The number of course–student combinations from university transcripts is 144,572 for type A, 15,493 for type B, and 12,456 for type C students. STEM majors are engineering, physics, mathematics, biology, geology, natural science, chemistry, and agricultural science. The restricted majors are medicine, the humanities, political science, law, and architecture. Nonrestricted majors are business, economics, and statistics. Income is winsorized at the 2nd and 98th percentiles. *Adjusted income* is taxable income in 2005 adjusted for age effects. Details on this procedure can be found in appendix F.
Sources: High school archives, university transcripts, and income tax returns in 2005.

Table 2: Congestion of Teaching Resources, Effect on Standardized Grades

	All type A				Pre-1966 Cohorts	Low-skill scientific	Reweighted
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
$\frac{E_c}{fac_c}$ x Phase I_p	-0.004***	-0.006***	-0.006***	-0.006***	-0.006***	-0.005*	-0.006***
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.003)	(0.001)
$\frac{E_c}{fac_c}$ x Phase II_p	-0.091***	-0.098***	-0.105***	-0.124***	-0.102***	-0.120***	-0.102***
	(0.016)	(0.015)	(0.017)	(0.024)	(0.018)	(0.036)	(0.015)
Pre-trend x $\mathbb{1}_{\left(\frac{E_c}{fac_c}>\text{median}\right)}$			-0.004				
			(0.004)				
Student f.e.	No	Yes	Yes	Yes	Yes	Yes	Yes
Professor f.e.	No	No	No	Yes	No	No	No
Observations	41,796	41,796	41,796	41,796	35,284	6,961	41,033

Notes: Effect of the average increase in the student–faculty ratio (+3.36 during phase I and +38.95 during phase II) on grades of type A students. The coefficients are computed from $g_{icp} = \alpha + \beta\frac{E_c}{fac_c} + \gamma_p + \sum_p \delta_p \left(\frac{E_c}{fac_c} \times P_p\right) + \eta X_{ip} + \zeta Z_{cp} + \kappa V_c + u_{icp}$, using data from STEM compulsory courses. g_{icp} is the standardized grade of student i in the compulsory course c in period p. $\frac{E_c}{fac_c}$ is the average preexisting student–faculty ratio in course c. γ_p are period fixed effects. P_p is 1 for post-implementation periods. Period 0 indicates the first academic year when type B students entered each STEM course. It ranges from 1961 for first-year courses to 1965 for fifth-year courses. During phase I (from period 0 to 3), enrollment of type B students in STEM majors was capped. During phase II (from period 4 to period 7), the enrollment caps for type B students were lifted. X_{ip} are student characteristics, Z_{cp} are time-varying course characteristics, andV_c are time-invariant course characteristics. "Pre-1966 Cohorts" limits the sample to students who completed high school before 1966. "Low-skill scientific" refers to type A students who scored in the bottom quartile in the high school exam and received a university degree. "Reweighted" uses sampling weights to match the average characteristics of pre-reform and post-reform observations. Robust standard errors in parentheses, *** p<0.01, ** p<0.05, * p<0.1.
Sources: University transcripts of type A students who completed high school in Milan, Italy; 1958–1968.

Table 3: Class Heterogeneity, Effect on Standardized Grades

	All type A				Pre-1966 Cohorts	Low-skill scientific	Reweighted
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Not in B cv_c x Phase I_p	-0.076***	-0.104***	-0.172	-0.135***	-0.106***	-0.131***	-0.113***
	(0.020)	(0.019)	(0.109)	(0.022)	(0.019)	(0.048)	(0.021)
Not in B cv_c x Phase II_p	-0.165***	-0.176***	-0.244**	-0.261***	-0.151***	-0.174***	-0.179***
	(0.022)	(0.022)	(0.110)	(0.030)	(0.027)	(0.051)	(0.024)
Pre-trend x *Not in B* cv_c			-0.011				
			(0.018)				
Student f.e.	No	Yes	Yes	Yes	Yes	Yes	Yes
Professor f.e.	No	No	No	Yes	No	No	No
Observations	53,502	53,502	53,502	53,502	44,930	8,798	52,646

Notes: Effect of the increase in class heterogeneity on the grades of type A students. The coefficients are computed from $g_{icp} = \alpha + \beta Not\ in\ B\ cv_c + \gamma_p + \sum_p \delta_p \left(Not\ in\ B\ cv_c \times P_p\right) + \eta X_{ip} + \zeta Z_{cp} + \kappa V_c + u_{icp}$, using data from STEM compulsory courses. g_{icp} is the standardized grade of student i in the compulsory course c in period p. *Not in B* cv_c is 1 if course c was not included in the curriculum of type B schools. γ_p are period fixed effects. P_p is 1 for post-implementation periods. Period 0 indicates the first academic year when type B students entered each STEM course. It ranges from 1961 for first-year courses to 1965 for fifth-year courses. During phase I (from period 0 to 3), enrollment of type B students in STEM majors was capped. During phase II (from period 4 to period 7), the enrollment caps for type B students were lifted. X_{ip} are student characteristics, Z_{cp} are time-varying course characteristics, andV_c are time-invariant course characteristics. "Pre-1966 Cohorts" limits the sample to students who completed high school before 1966. "Low-skill scientific" refers to type A students who scored in the bottom quartile in the high school exam and received a university degree. "Reweighted" uses sampling weights to match the average characteristics of pre-reform and post-reform observations. Robust standard errors in parentheses, *** p<0.01, ** p<0.05, * p<0.1.
Sources: University transcripts of type A students who completed high school in Milan, Italy; 1958–1968.

Table 4: Income of Type A Students

	Baseline	Trends		Low-skill	Reweighted	ITT
	(1)	(2)	(3)	(4)	(5)	(6)
STEM x Post 61	0.017	-0.244	0.015	-0.196	0.014	
	(0.081)	(0.181)	(0.081)	(0.155)	(0.082)	
STEM x Post 65	-0.138*	-0.398**	-0.144*	-0.421***	-0.157**	
	(0.075)	(0.178)	(0.074)	(0.141)	(0.076)	
Pre-reform trend x STEM		-0.130				
		(0.081)				
Scientific x Post 61						-0.289*
						(0.164)
Scientific x Post 65						-0.293**
						(0.146)
Trends for quartiles of precollegiate ability	No	No	Yes	No	No	No
Observations	6,398	6,398	6,398	1,731	6,398	1,731

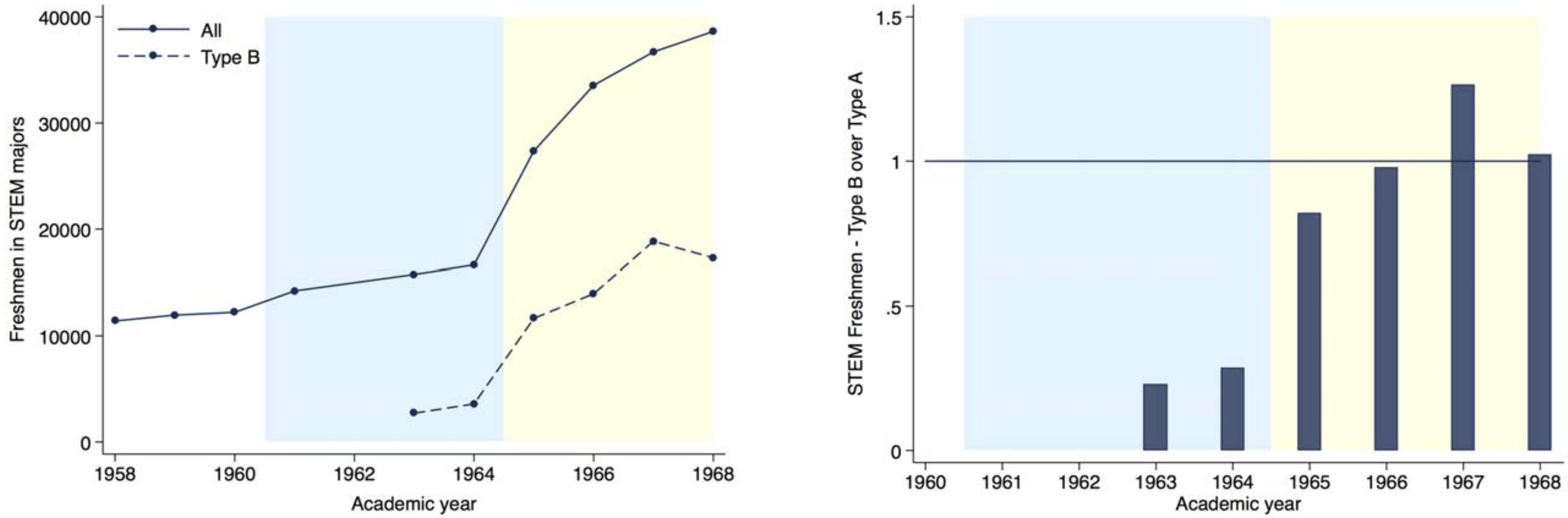
Notes: The coefficients are estimated from $log(\text{income}_{imt}) = \alpha + \beta\text{STEM}_m + \gamma_t + \sum_t \delta_t[\text{STEM}_m \times Post_t] + \zeta X_{imt} + u_{imt}$. The unit of observation is a student i, who completed high school in year t and chose major m after high school. $log(\text{income}_{imt})$ is the log taxable personal income in 2005, $Post_t$ is 1 if a student completed high school after the implementation of the reform (1961), STEM_m is 1 if a student received a STEM degree, γ_t are cohort fixed effects, and X_{imt} are student characteristics. Post 61 is 1 for the cohorts who graduated between 1961 and 1964, and Post 65 is 1 for the cohorts who graduated between 1965 and 1968. "Low-skill" refers to type A students who scored in the bottom quartile in the high school exam and received a university degree. "Reweighted" uses sampling weights to match the average characteristics of pre-reform and post-reform observations. "ITT" compares low-achieving students from scientific academic-track schools to low-achieving students from humanistic academic-track schools. Robust standard errors in parentheses, *** p<0.01, ** p<0.05, * p<0.1.
Sources: University transcripts and income in 2005 of type A students who completed high school in Milan, Italy; 1958–1968.

Table 5: Income of Type B Students

	Baseline	Trends			Pre-1966 Cohorts	Reweighted
	(1)	(2)	(3)	(4)	(5)	(6)
Type B x Post 61	0.136**	0.242**	0.132**	0.092	0.136**	0.135**
	(0.066)	(0.095)	(0.066)	(0.109)	(0.067)	(0.066)
Type B x Post 65	0.059	0.165*	0.056	-0.044	-0.109	0.089
	(0.066)	(0.095)	(0.066)	(0.192)	(0.095)	(0.066)
Pre-reform trend x Type B		0.098				
		(0.064)				
Trends for quartiles of precollegiate ability	No	No	Yes	No	No	No
Trends for ability–high school	No	No	No	Yes	No	No
Observations	13,063	13,063	13,063	13,063	6,621	11,878

Notes: The coefficients are estimated from $log(\text{income}_{ist}) = \alpha + \beta\text{Type B}_s + \gamma_t + \sum_t \delta_t[\text{Type B}_s \times Post_t] + \zeta X_{ist} + u_{ist}$. The unit of observation is a student i, who completed high school s in year t. $log(\text{income}_{ist})$ is the log taxable personal income in 2005, $Post_t$ is 1 if a student completed high school after the implementation of the reform (1961), Type B_s is 1 if a student received a STEM degree, γ_t are cohort fixed effects, and X_{ist} are student characteristics. Post 61 is 1 for the cohorts who graduated between 1961 and 1964, and Post 65 is 1 for the cohorts who graduated between 1965 and 1968. "Pre-1966 Cohorts" limits the sample to students who completed high school before 1966. "Reweighted" uses sampling weights to match the average characteristics of pre-reform and post-reform observations. Robust standard errors in parentheses, *** p<0.01, ** p<0.05, * p<0.1. Sources: University transcripts and income in 2005 of type B and C students who completed high school in Milan, Italy; 1958–1968.

Figure 1: Freshmen in STEM Majors

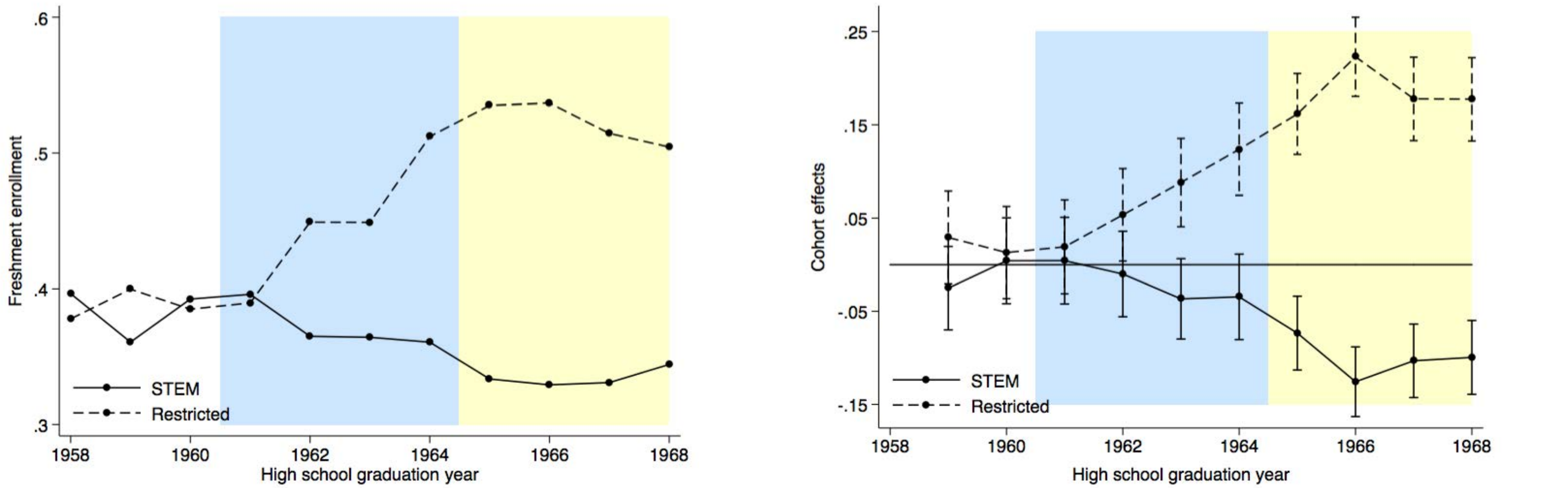


A. Number of Freshmen in STEM Majors

B. Ratio of Type B to Type A

Notes: These graphs show the change in the number and composition of freshmen students in all STEM majors in Italy. Panel A: "All" counts the total number of freshmen students enrolled in a STEM major for each academic year; "Type B" is the number of freshmen students with a type B diploma (1961 and 1962 are missing). STEM majors are engineering, mathematics, physics, chemistry, biology, geology, natural science, and agricultural sciences. The blue shaded area denotes the first phase of the reform: between 1961 and 1964, enrollment of type B students was capped. The yellow shaded area denotes the second phase of the reform: in 1965, the cap to type B enrollment in STEM fields was lifted.
Sources: Annuario Statistico dell'Istruzione Italiana, 1958–1968, Istituto Nazionale di Statistica.

Figure 2: Type A, Enrollment in STEM and Restricted Majors



A. Raw Shares

B. Marginal Cohort Effects

Notes: Panel A shows the raw shares of type A students enrolling in STEM and restricted majors by year of high school graduation. Panel B shows marginal cohort effects with 95 percent confidence intervals from the multinomial logit $ln(\frac{Pr(\text{major}_{it}=m)}{Pr(\text{no university})}) = \alpha_m + \beta_m X_{it} + \gamma_{mt}$, where m is either a STEM, restricted, or nonrestricted major, or no university (baseline). γ_{mt} are year of high school graduation fixed effects with 1958 as omitted category. X_{ic} includes gender, the high school exit score, the average exit score of the high school classmates, high school fixed effects, a dummy for home-schooled students, and a dummy for non-repeaters. STEM majors are engineering, physics, mathematics, biology, geology, natural science, chemistry, and agricultural science. The restricted majors are medicine, the humanities, political science, law, and architecture. The two shaded areas denote phase I and phase II of the reform.

Sources: School data of type A students who completed high school in Milan, Italy; 1958–1968.

Online Appendix - Not For Publication

A Data Collection

The data collection targeted the population of high school students who graduated from a public high school in the city of Milan between 1958 and 1968. The whole process constituted of three main phases.

Between September 2012 and January 2013, I contacted all 19 public high schools in Milan that were granting either a type A (*licei classici e scientifici*), type B (*istituti tecnici industriali*), or type C (*istituti tecnici commerciali*) diploma between 1958 and 1968. 18 schools approved my request to make copies of the student registries (Appendix Figure A2, Panel A), but in one case the archive did not contain the registries for the period under consideration. In some isolated instances, the registries of single school years could not be located in the archives of participating schools. For these reasons, the data cover 74 percent of the high school population in Milan.

Between January 2013 and April 2013, I copied university transcripts from the archives of the three local universities (Appendix Figure A2, Panel B). Two of these universities are public, *Università Statale di Milano* and *Politecnico di Milano*, while the third is private, *Università Cattolica del Sacro Cuore*. The two public universities offered non-overlapping sets of majors: *Politecnico (*Polytechnic*)* focused on engineering and architecture, while *Università Statale* (State University) offered all other majors with the exception of business and economics. *Università Cattolica del Sacro Cuore* (Catholic University of the Sacred Hearth) focused on the humanities majors and the social sciences. The fourth university in Milan, Università Bocconi, was not included in the data collection. Differently from the other Italian universities, Bocconi was charging high tuition fees and admission was highly selective. In addition, Bocconi offered exclusively business and economics majors, which were accessible to type B and type C graduates even before the 1961 reform of university admissions.

Photographic copies of the data were digitized during the months between January 2013 and December 2013 with the help of freelancers hired on a popular online marketplace. The fact that significant portions of the data were hand-written made necessary to hire Italian-speaking typists in order to minimize mistakes in the data entry. The high school registries were transcribed directly into excel spreadsheets. The same procedure, however, was not an option for university transcripts, due to their complex structure and high number of variables. For this reason, I provided each contractor with a data-entry software that I specifically designed to visually reproduce the fields of university transcripts. In addition, I pre-loaded drop-down lists for many string variables, such as course titles. This software sped up the digitization process, lowered the incidence of mistakes, and made data-checking easier.

The resulting dataset of high school graduates was matched with a complete list of personal income tax returns in 2005. Income observations in Italy are extremely rare. The complete list of income tax returns that I used in this paper was published online by the Italian Treasury on March 5, 2008. The goal was to fight tax evasion, allowing every citizen to check the income reported by acquaintances, coworkers, and neighbors. The Italian public strongly opposed this way of disseminating income observations and the data files remained available online for less than 24 hours. The academics that downloaded the data on March 5, 2008 can now use the income observations for research purposes. The data are organized in separate text files for each Italian city. Each file contains the complete list of local income earners with full name, birthdate, the total taxable income after deductions, the due income tax, a coarse indication of the main source of income, and the city of residence in 2005.

Acknowledgments

Many people helped before and during the data collection in Milan. I thank Giovanni Peri for sharing with me his contacts at the Province of Milan and his experience acquired in a similar data collection. I thank Prof. Bruna Sinnone and Prof. Sandra Favi at the Province of Milan. I thank the principals and the staff of the participating schools for their help in the collection of the high school data. For their help in the collection of the university transcripts, I thank Prof. Mauro Santomauro and the staff of the Archivio Storico at Politecnico di Milano; Prof. Daniele Checchi, Emanuela Dellavalle, Idilio Baitieri and the staff of Segreteria Studenti at Università Statale di Milano; Aldo Piacentini, Maurizio Zambon, Claudio Maderna and the staff of Segreteria Studenti at Università Cattolica del Sacro Cuore di Milano. Prof. Aldo Carera shared with me his data on the assignment of faculty to university courses at Università Cattolica.

Table A1: Test of Means, Students With and Without an Income Observation

	Matched	Not Matched	Difference
	(1)	(2)	(3)
Male	0.749	0.730	0.019***
Age in 2005	61.663	62.474	-0.811***
Type A	0.421	0.415	0.006
Type B	0.327	0.304	0.023***
Type C	0.252	0.281	-0.029***
HS exit score	6.420	6.384	0.036***
Observations	22,579	4,657	

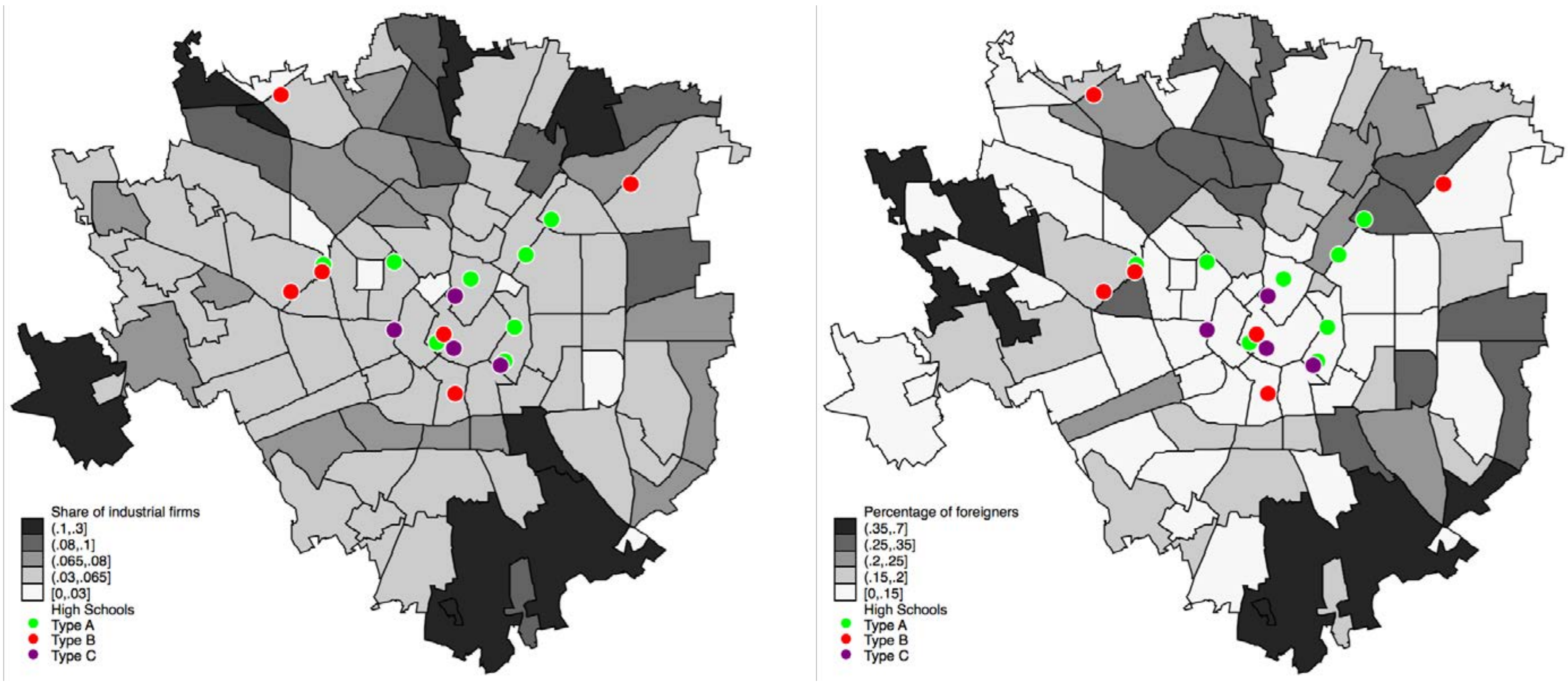
Notes: *Matched* are the 22,579 students that are matched with an income earner from the complete list of personal income tax returns in 2005. Similarly, *Not Matched* are the remaining 4,657 students that did not have a correspondence in the list of income tax returns. *** p<0.01, ** p<0.05, * p<0.1.

Table A2: OLS, Test for Selective Attrition in Matching Students to Income Earners

	Matched	Matched
	(1)	(2)
F-test on student controls	19.43	14.80
	(<0.001)	(<0.001)
F-test on cohort FEs	5.87	0.074
	(<0.001)	(0.689)
F-test on high school FEs	2.38	2.32
	(0.001)	(0.002)
F-test on cohort FEs x Male	Not included	0.91
		(0.527)
F-test on cohort FEs x HS score	Not included	0.16
		(0.999)
Observations	27,206	27,206

Notes: Each column shows the F-statistics and the corresponding p-values (in parenthesis) from separate OLS regressions. The dependent variable is 1 if a student was matched with an income earner in 2005. *** p<0.01, ** p<0.05, * p<0.1.

Figure A1: Location of High Schools in Milan



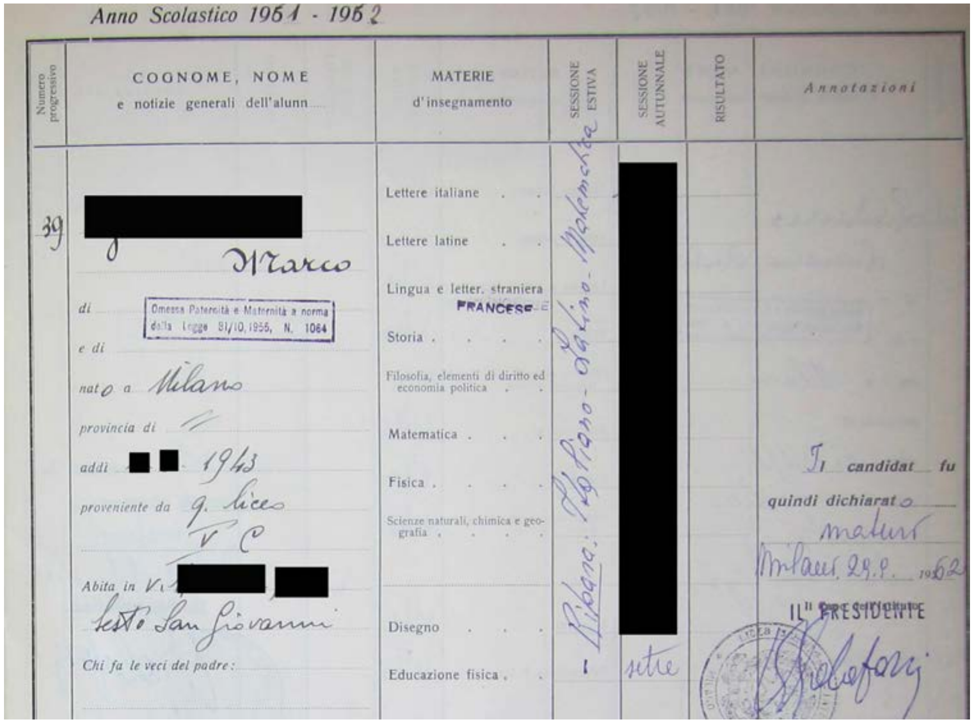
A. Share of Industrial Firms (2011)

B. Share of Foreigners (2011)

Notes: Map of Milan, Italy. The boundaries identify the 88 neighborhoods (*Nuclei di Identità Locale*). The share of industrial firms is the percentage of industrial firms out of all economic activities in the neighborhood (Panel A). The percentage of foreigners is the share of non-nationals over total population living in each neighborhood in year 2011 (Panel B). The colored dots report the location of the public high schools of type A, B, and C that were active throughout the years between 1958 and 1968.

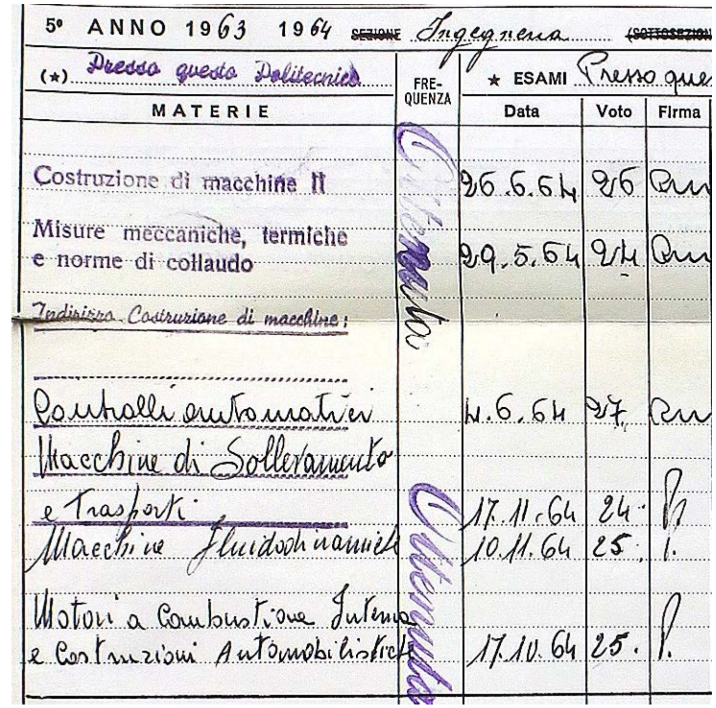
Sources: Data available at http://dati.comune.milano.it/dato/item/61 and http://dati.comune.milano.it/dato/item/201-201-imprese-numero-di-unita-locali-per-settore-di-attivita-e-quartiere-2010.html.

Figure A2: Examples of High School and University Data


Anno Scolastico 1961 - 1962

Numero progressivo	COGNOME, NOME e notizie generali dell'alunn...	MATERIE d'insegnamento	SESSIONE ESTIVA	SESSIONE AUTUNNALE	RISULTATO	Annotazioni
39	[redacted] Marco	Lettere italiane		[redacted]		
	di Omessa Paternità e Maternità a norma della Legge 31/10/1955, N. 1064	Lettere latine				
	e di	Lingua e letter. straniera FRANCESE				
	nato a Milano	Storia				
	provincia di	Filosofia, elementi di diritto ed economia politica				
	addì [redacted] 1943	Matematica				Il candidat fu quindi dichiarato maturo
	proveniente da	Fisica				Milano, 28.9.1962
	Abita in V[redacted] Sesto San Giovanni	Scienze naturali, chimica e geografia				IL PRESIDENTE
	Chi fa le veci del padre:	Disegno				
		Educazione fisica				

A. High School Data


5° ANNO 1963 1964 SEZIONE Ingegneria

(*) Presso questo Politecnico

MATERIE	FREQUENZA	Data	Voto	Firma
Costruzione di macchine II	Ottenuta	26.6.64	26	
Misure meccaniche, termiche e norme di collaudo		29.5.64	24	
Indirizzo Costruzione di macchine:				
Controlli automatici		4.6.64	27	
Macchine di Sollevamento e Trasporti	Ottenuta	17.11.64	24	
Macchine Fluidodinamiche		10.11.64	25	
Motori a Combustione Interna e Costruzioni Automobilistiche		17.10.64	25	

B. University Transcripts

Notes: Panel A shows the information available for one high school students. I blacked out several parts to guarantee anonymity. Panel B shows an excerpt of a university transcript. Specifically, it shows courses attended during the fifth year of engineering with exam dates and outcomes for one university student.

B The Italian Education System

Table B1: High school curricula

	Type A (1)	Type B (2)	Type C (3)
Humanities	Italian, philosophy, history, Latin, Ancient Greek, art history	Italian	Italian
Sciences	Mathematics, physics, chemistry, geography, biology	Chemistry	Financial mathematics, geography
Applied disciplines (type B: not exhaustive)	Technical drawing	Technical drawing, topography, land appraisal, mechanics, optics, electro-chemistry, thermal eng., aeronautical eng., electrical eng., nuclear eng., training in workshops and labs	Accounting
Law and economics	No	Law (basics)	Law, economics
Foreign languages	Yes (1)	No	Yes (2)
Non-academic	P.E.	P.E.	P.E.

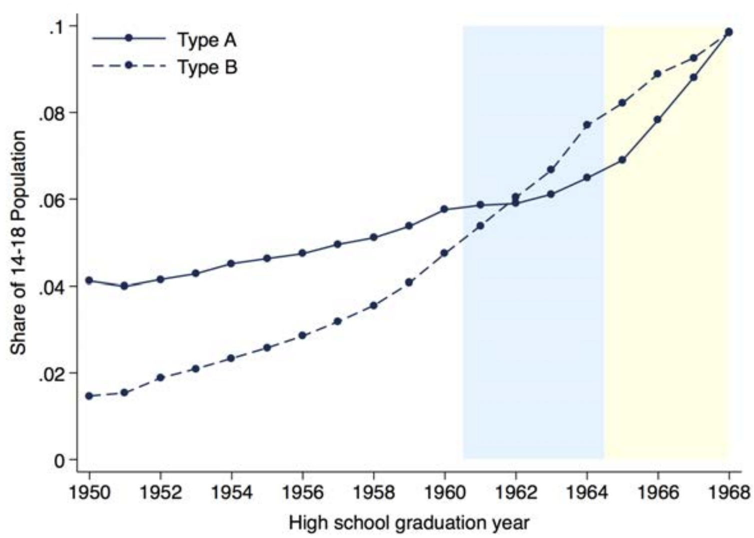
Notes: Type A schools are university-prep high schools, which focus on either the humanities or science. Type B and C are technical high schools, which train professionals for specific economic sectors: type B are industrial schools, which prepare students for jobs in the industry and construction, and type C are commercial schools, which prepare students for jobs in the service sector.

Table B2: Parents' Characteristics and High School Choice

	Father			Mother		
	Type A	Technical	Difference	Type A	Technical	Difference
	(1)	(2)	(3)	(4)	(5)	(6)
Education						
No education	0.057	0.075	-0.018	0.126	0.094	0.032
Low education	0.530	0.749	-0.219***	0.660	0.801	-0.141***
High education	0.413	0.176	0.237***	0.246	0.072	0.174***
Occupation						
Low income	0.275	0.408	-0.133***	0.823	0.856	-0.033
Middle income	0.438	0.461	-0.023	0.148	0.139	0.009
High income	0.287	0.131	0.156***	0.029	0.007	0.022*
Sector						
Public servant	0.277	0.232	0.045	0.483	0.132	0.351***

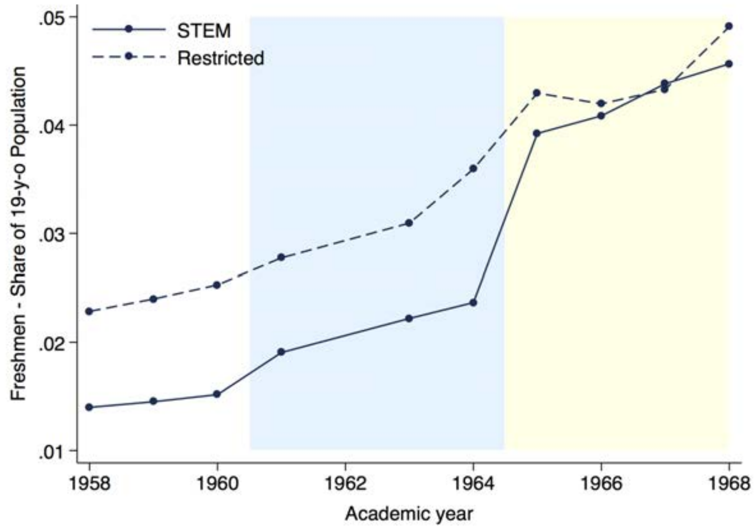
Notes: Data from 1,802 individuals born between 1931 and 1950 and with at least a high school diploma; 710 individuals have a type A diploma, while 1,092 have a technical diploma. The SHIW does not distinguish between types of technical diplomas. Respondents were asked the education, employment status, and sector of activity of their parents at their current age (or earlier, if deceased or retired at that age). Low education: a primary school or lower secondary school certificate. High education: high school diploma or higher. Low income: production workers and not employed. Middle income: clerical workers, teachers, self-employed. High income: managers, professionals, and entrepreneurs. *** p<0.01, ** p<0.05, * p<0.1.
Sources: Bank of Italy's SHIW; 2006, 2008, 2010, 2012 waves.

Figure B1: Type A and Type B Students



Notes: Total number of type A and type B students in Italy as a share of 14–18-year-olds.
Sources: Annuario Statistico dell'Istruzione Italiana, 1958–1968, Istituto Nazionale di Statistica.

Figure B2: Freshmen in STEM and Restricted Majors



Notes: Freshmen students enrolled in STEM and restricted majors in Italy as a share of 19-year-olds.
Sources: Annuario Statistico dell'Istruzione Italiana, 1958–1968, Istituto Nazionale di Statistica.

C Congestion: Additional Results

Table C1: Determinants of the Number of Teaching Fellows in STEM courses

	Compulsory Courses		All Courses	
	(1)	(2)	(3)	(4)
Tenured professor	3.209***	2.720***	3.892***	3.956***
	(0.550)	(0.606)	(0.488)	(0.509)
Institute director	2.339***	1.272	2.066***	0.671
	(0.619)	(1.248)	(0.429)	(0.742)
Female professor	-0.613**		-0.588**	
	(0.299)		(0.237)	
Compulsory course			0.004	0.364
			(0.336)	(0.467)
Number of students	0.028***	0.025***	0.029***	0.024***
	(0.006)	(0.005)	(0.005)	(0.005)
Lagged average grade	0.012	0.294**	0.010	0.146*
	(0.163)	(0.114)	(0.101)	(0.081)
Mean, 1958–68	4.38	4.38	3.39	3.39
Professor FE	No	Yes	No	Yes
Observations	1,388	1,388	2,440	2,440

Notes: The dependent variable is the number of teaching fellows assigned to each STEM course. Each regression controls for major and academic year fixed effects. Standard errors clustered by course in parentheses, *** p<0.01, ** p<0.05, * p<0.1.
Sources: Annals of Università Statale di Milano, Politecnico di Milano, and Università Cattolica del Sacro Cuore. University transcripts of students that completed high school in Milan, Italy; 1958–1968.

Table C2: Pre-Reform Differences, Courses with Low and High Student–Faculty Ratio

	Tenured professor	Institute director	Female professor	Grades (18-31)	Not in B cv
	(1)	(2)	(3)	(4)	(5)
$\mathbb{1}_{\left(\frac{E_c}{fac_c}>\text{median}\right)}$	-0.327***	-0.223**	0.071	-0.566	-0.006
	(0.095)	(0.108)	(0.058)	(0.346)	(0.969)
Observations	289	289	289	281	281

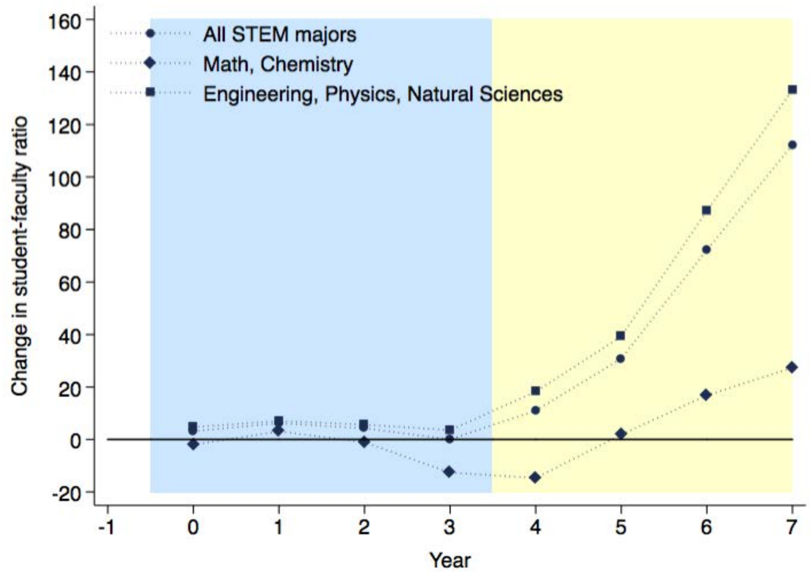
Notes: Data from compulsory courses in STEM majors before the reform. The indicator function $\mathbb{1}_{\left(\frac{E_c}{fac_c}>\text{median}\right)}$ identifies the courses with pre-reform student-faculty ratio above median. The regressions control also for major, period, and curriculum year fixed effects. "Grades" are non-standardized grades. "Tenured", "Institute director", "Female professor" are characteristics of the professor assigned to the course. "Not in B cv" is a binary variable for courses not included in the curricula of type B high schools. Standard errors clustered by course in parentheses, *** p<0.01, ** p<0.05, * p<0.1.
Sources: Annals of Università Statale di Milano, Politecnico di Milano, and Università Cattolica del Sacro Cuore. University transcripts of students that completed high school in Milan, Italy; 1958–1968.

Table C3: Congestion, Additional Results and Robustness Checks

	$\frac{E_c}{fac_c}$ x Phase I_p	$\frac{E_c}{fac_c}$ x Phase II_p	Obs.
	(1)	(2)	(3)
Courses with increasing student–faculty ratio	-0.015***	-0.154***	38,276
	(0.003)	(0.023)	
Courses without increasing student–faculty ratio	0.031	0.002	3,520
	(0.020)	(0.012)	
Engineering, Physics, Natural Sciences	-0.009***	-0.135***	37,378
	(0.002)	(0.021)	
Math and chemistry	0.004	-0.000	4,324
	(0.003)	(0.000)	
Course f.e.	-0.006***	-0.101***	41,796
	(0.001)	(0.015)	
Major-specific linear trends	-0.006***	-0.100***	41,796
	(0.001)	(0.015)	
Trends for quartile of precollegiate ability	-0.006***	-0.098***	41,796
	(0.001)	(0.015)	
Major-specific linear and squared trends	-0.006***	-0.100***	41,796
	(0.001)	(0.015)	
High school-specific linear trends	-0.006***	-0.097***	41,796
	(0.001)	(0.015)	
Major–ability linear trends	-0.006**	-0.100***	41,796
	(0.001)	(0.019)	
S.e. clustered at student level	-0.006***	-0.098***	41,796
	(0.001)	(0.016)	
S.e. clustered at exam level	-0.006***	-0.098***	41,796
	(0.002)	(0.035)	
S.e. clustered at high school–graduation year level	-0.006***	-0.098***	41,796
	(0.002)	(0.019)	

Notes: Each row shows the results of a different specification of equation (6). The coefficients reported measure the effect of the average increase in the student–faculty ratio (+3.36 during phase I and +38.95 during phase II in all STEM courses; +9.18 during phase I and +52.52 during phase II in courses with an increase in congestion; -15.60 during phase I and -8.21 during phase II in courses without an increase in congestion; +5.15 during phase I and +49.84 during phase II in engineering, physics, and the natural sciences; -3.19 during phase I and +0.12 during phase II in math and chemistry) on grades of type A students. Robust (or clustered, where specified) standard errors in parentheses, *** p<0.01, ** p<0.05, * p<0.1.
Sources: University transcripts of students who completed high school in Milan, Italy; 1958–1968.

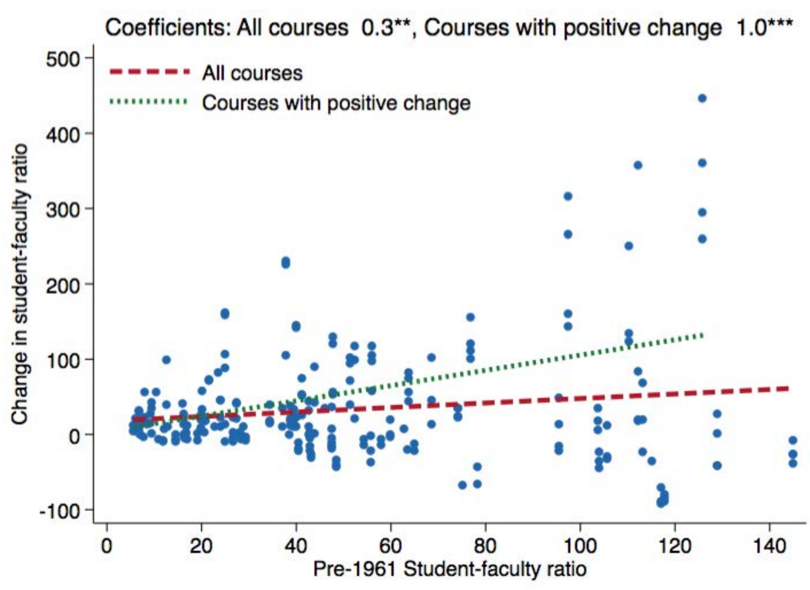
Figure C1: Students–Faculty Ratio in STEM Majors



Notes: The graph shows the average student–faculty ratio for each period in different STEM majors in Milan.

Sources: Annals of Università Statale di Milano, Politecnico di Milano, and Università Cattolica del Sacro Cuore; 1958–1968.

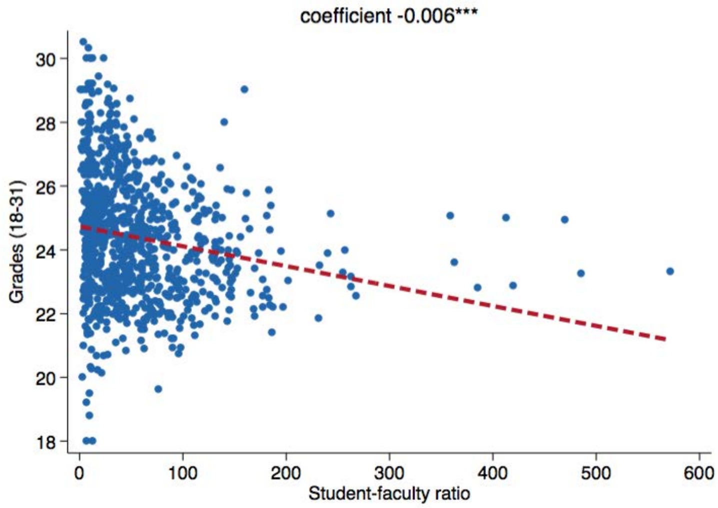
Figure C2: Preexisting Student–Faculty Ratio and Post-Reform Congestion



Notes: The graph shows the existence of a positive correlation between the pre-reform student-faculty ratio and the post-reform change in student–faculty ratio.

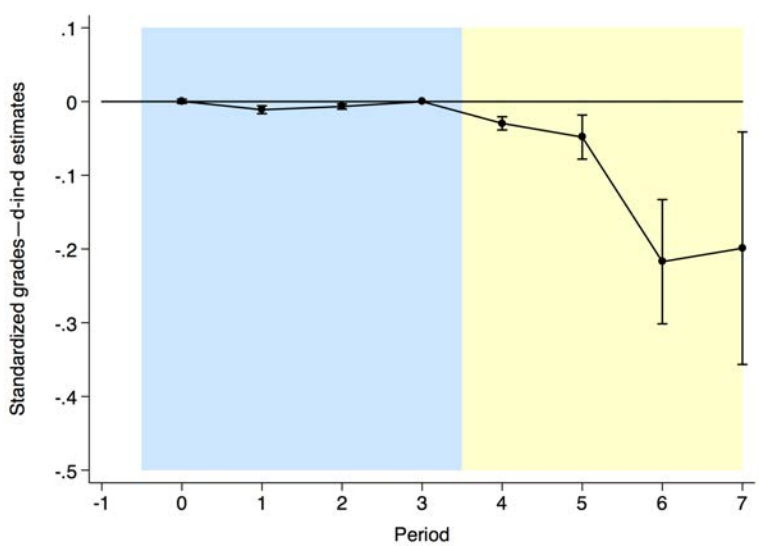
Sources: Annals of Università Statale di Milano, Politecnico di Milano, and Università Cattolica del Sacro Cuore; 1958–1968.

Figure C3: Average Grades and Student–Faculty Ratio



Notes: Each point is a course-academic year combination.
Sources: Annals of Università Statale di Milano, Politecnico di Milano, and Università Cattolica del Sacro Cuore; 1958-1968. University transcripts of students that completed high school in Milan, Italy; 1958–1968.

Figure C4: Congestion, Per-Period Effect on Standardized Grades



Notes: Per-period effect of the average increase in the student–faculty ratio (+3.18 in period 0, +6.17 in period 1, +4.26 in period 2, +0.042 in period 3, +11.03 in period 4, +30.87 in period 5, +72.22 in period 6, +112.21 in period 7) on grades of type A students. The coefficients are computed from equation (6). The bars represent 95 percent confidence intervals. The omitted academic periods are -3 to -1.
Sources: University transcripts of students who completed high school in Milan, Italy; 1958–1968.

D Class Heterogeneity: Additional Results

To create the variable *Not in B* cv_{cm}, I use the institute of affiliation to determine whether a course was taught in type B schools. Unlike course titles, in fact, institutes have unambiguous denominations from which it is easy to infer the field of study. As an example, consider the engineering course "Analytical Mechanics" that studies a branch of mathematical physics and belongs to the institute of mathematical sciences. Based on its title, "Analytical Mechanics" could be misinterpreted as a course in applied mechanics, an area of expertise of type B students. Using the institute of affiliation, however, I correctly categorize this course as not being included in the precollegiate curriculum of type B students.

Table D1: List of Institutes in Type B Cv

Fields	Institutes in type B Cv	Institutes not in type B Cv
Agricultural science	"Chimica agraria", "Chimica organica", "Economia e politica agraria", "Idraulica agraria", "Meccanica agraria"	"Agronomia", "Anatomia e fisiologia degli animali domestici", "Coltivazioni arboree", "Entomologia agraria", "Fisiologia della nutrizione animale", "Industrie agrarie", "Ispezione degli alimenti di origine animale", "Istologia ed embriologia", "Microbiologia agraria", "Morfologia e fisiologia vegetale", "Patologia vegetale", "Scienze botaniche", "Scienze fisiche", "Scienze matematiche", "Tecnologie alimentari", "Zootecnica generale"
Engineering	"Chimica", "Chimica fisica, elettrochimica e metallurgia", "Chimica industriale", "Disegno generale", "Edilizia", "Elettrotecnica ed elettronica", "Elettrotecnica industriale", "Geodesia, topografia, e fotogrammetria", "Idraulica", "Ingegneria aerospaziale", "Ingegneria nucleare", "Macchine", "Meccanica", "Scienza e tecnica delle costruzioni", "Vie e trasporti"	"Costruzioni di ponti", "Fisica", "Fisica tecnica", "Matematica"
Sciences	"Chimica fisica", "Chimica generale ed inorganica", "Chimica industriale", "Chimica organica", "Topografia e cartografia"	"Fisiologia generale", "Geologia", "Igiene", "Istologia ed embriologia", "Mineralogia, petrografia e geochimica", "Paleontologia", "Pedagogia", "Scienze botaniche", "Scienze fisiche", "Scienze matematiche", "Zoologia"

Table D2: Grades of Type B students in STEM Courses

	(1)	(2)	(3)	(4)	(5)	Obs.
Type B only						
Not in B cv_c	-0.243***	-0.256***	-0.215***	-0.213***	-0.213***	9,858
	(0.020)	(0.022)	(0.023)	(0.023)	(0.021)	
Type B vs Type A						
Not in B cv_c x $TypeB_i$	-0.279***	-0.294***	-0.278***	-0.280***	-0.280***	63,523
	(0.022)	(0.022)	(0.022)	(0.022)	(0.021)	
Student controls	No	Yes	Yes	Yes	Yes	
Course controls	No	No	Yes	Yes	Yes	
Period f.e.	No	No	No	Yes	Yes	
Student f.e.	No	No	No	No	Yes	

Notes: Each cell shows the coefficient from a separate regression. The first row compares the grades of type B students in university courses that were not included in their high school cv with the grades of type B students in other STEM courses (single difference). The second row compares the grade difference between the courses taught in type B schools and the courses not taught in type B schools, for type B students and type A students (difference-in-differences). *Not in B* cv_c is 1 if a course is not included in the high school cv of type B students. Robust standard errors in parentheses, *** p<0.01, ** p<0.05, * p<0.1.
Sources: University transcripts of students who completed high school in Milan, Italy; 1958–1968.

Table D3: Pre-Reform Differences, Courses with High and Low Class Heterogeneity

	Tenured professor	Institute director	Female professor	Grades (18-31)	Student–faculty ratio
	(1)	(2)	(3)	(4)	(5)
Not in B cv_c	0.209**	0.037	0.092	0.184	-25.049**
	(0.097)	(0.095)	(0.073)	(0.276)	(9.931)
Observations	404	404	404	403	404

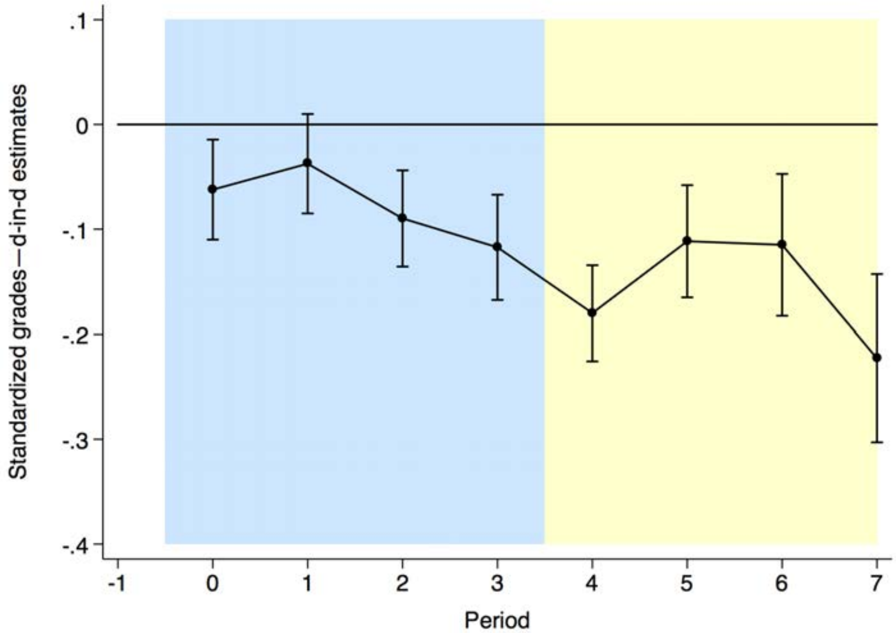
Notes: Data from compulsory courses in STEM majors before the reform. The dummy variable Not in B cv identifies the courses that were not taught in type B schools. The regressions control also for major, period, and curriculum year fixed effects. "Grades" are non-standardized grades. "Tenured", "Institute director", "Female professor" are characteristics of the professor assigned to the course. "Student–faculty ratio" is the student–faculty ratio of each course in each period. Standard errors clustered by course in parentheses, *** p<0.01, ** p<0.05, * p<0.1.
Sources: Annals of Università Statale di Milano, Politecnico di Milano, and Università Cattolica del Sacro Cuore. University transcripts of students who completed high school in Milan, Italy; 1958–1968.

Table D4: Class Heterogeneity, Additional Results and Robustness Checks

	Not in B cv_c x Phase I_p	*Not in B* cv_c x Phase II_p	Obs.
	(1)	(2)	(3)
Course f.e.	-0.128***	-0.208***	53,502
	(0.020)	(0.023)	
Major-specific linear trends	-0.105***	-0.179***	53,502
	(0.019)	(0.022)	
Trends for quartile of precollegiate ability	-0.104***	-0.176***	53,502
	(0.019)	(0.022)	
Major-specific linear and squared trends	-0.103***	-0.180***	53,502
	(0.019)	(0.022)	
High school-specific linear trends	-0.103***	-0.175***	53,502
	(0.019)	(0.022)	
Major–ability linear trends	-0.105***	-0.179***	53,502
	(0.020)	(0.016)	
S.e. clustered at student level	-0.104***	-0.176***	53,502
	(0.023)	(0.025)	
S.e. clustered at exam level	-0.104**	-0.176**	53,502
	(0.047)	(0.086)	
S.e. clustered at high school–graduation year level	-0.104***	-0.176***	53,502
	(0.029)	(0.027)	

Notes: Each row shows the results of a different specification of equation (7). The coefficients reported measure the effect of the increase in class heterogeneity on grades of type A students. Robust (or clustered, where specified) standard errors in parentheses, *** $p<0.01$, ** $p<0.05$, * $p<0.1$.
Sources: University transcripts of students who completed high school in Milan, Italy; 1958–1968.

Figure D1: Class Heterogeneity, Per-Period Effect on Standardized Grades



Notes: Per-period effect of the increase in class heterogeneity on grades of type A students. The coefficients are computed from equation (7). The bars represent 95 percent confidence intervals. The omitted academic periods are -3 to -1.
Sources: University transcripts of students who completed high school in Milan, Italy; 1958–1968.

E The Major Choice of Type A Students

Main Results

The major choice of type A students can be written as a multinomial logit model:

$$\log(\frac{Pr(\text{major}_{ic} = m)}{Pr(\text{no university})}) = \alpha_m + \beta_m X_{it} + \gamma_{mt}, \tag{E1}$$

where the choice is between STEM, restricted, and nonrestricted majors, and no university is the baseline. γ_{mt} is a full set of cohort fixed effects. X_{it} includes gender, high school fixed effects, the high school exit score, the mean score of high school classmates, a dummy for home-schooled students, and a dummy for students who did not repeat a grade in high school. The identifying assumption to estimate cohort effects in equation (E1) is that adjacent cohorts differ only in relation to their exposure to the reform. The threats to identification can be grouped in two classes: changes in the characteristics of type A students and exogenous changes in the returns to different majors.

Initially, I estimate the multinomial logit model in equation (E1) dropping controls for precollegiate ability. The cohort effects do not change (Appendix Figure E1, Panel A). In 1968, type A students were 10 percent less likely to enroll in STEM and 18 percent more likely to enroll in restricted majors, relative to the 1958 cohort.

The increasing education of female students over this time period could explain the shift towards restricted majors, which include fields with high female participation like the humanities. To test this hypothesis, I estimate the model in (E1) using data on male students only. The main findings hold (Table E1 and Appendix Figure E1, Panel B). Male type A students who completed high school from 1965 to 1968 were 10.7 percentage points more likely to enroll in STEM and 16.6 percentage points less likely to choose a restricted major, relative to the 1958 cohort.

In a separate test, I estimate the model in (E1) using only the pre-1961 cohorts. I then use the estimated coefficients to predict the major choice of type A students who completed high school after 1961. If changes in students' characteristics do not drive the diverging trends in major choice, predicted and actual shares should follow different paths. The predicted share of type A students enrolling in STEM majors follows a slightly increasing path after 1961 (Appendix Figure E2, Panel A), while the predicted share in restricted majors is stable (Appendix Figure E2, Panel B).

The second group of robustness checks tests for the role of concurrent and exogenous changes in the returns to higher education.[21] To address this concern, I estimate a conditional multinomial logit model in which I control simultaneously for students' characteristics and returns to different university majors. As a proxy for returns to education, I use the sectoral value added per full-time equivalent worker in the industry, finance, and service sectors (Baffigi, 2011).[22] Controlling for contemporaneous changes in the economy does not affect the path of the marginal cohort effects (Appendix Figure E1, Panel C). In 1968, type A students were 16.7 percentage points less likely to enroll in STEM and 26.5 percentage points more likely to enroll in restricted majors, relative to the 1958 cohort.

Lastly, I estimate the model in (E1) using more disaggregated choices to show that enrollment in fairly different STEM (restricted) majors follow the same decreasing (increasing) trend after 1961 (Appendix Figure E1, Panel D).[23] In 1968, type A students were 2.1 percentage points less likely to enroll in engineering and 5.1 percentage points in physics, relative to the 1958 cohort. At the same time, they were 6.3 percentage points more likely to enroll in the humanities and 13.6 percentage points in medicine.

Movers and Stayers

The coefficients estimated from equation (E1) using pre-1961 data can be used to predict the major choice of students who completed high school after 1961. The resulting predicted and actual major choices of type A students can be compared to identify the students who were more likely to abandon STEM majors (Appendix Table E2). First, the graduates from type A humanities schools (*licei classici*) were more likely to move to restricted majors after 1961, compared with the graduates from type A scientific schools (*licei scientifici*): the predicted decrease in STEM enrollment was equal to 10.2 percentage points among humanities students and only to 3.1 percentage points among scientific students. This finding holds at any level of precollegiate ability and suggests that educational expansion affected disproportionately the students with stronger preferences for STEM disciplines. Second, the type A students who scored in the top quartile of their high school class (high achieving) were more likely to move out of STEM majors after 1961, compared with the students who scored in the bottom quartile (low achieving): the predicted decrease in STEM enrollment was 11.2 percentage

[21] The economic downturn that affected Italy during the 70's could have affected the industry sector more than services and government, therefore inducing more students to abandon STEM majors.

[22] I use the SHIW dataset to show that different majors lead to occupations in different sectors: STEM to industry, restricted majors to services and government, non-restricted majors to banking and finance, and a high school diploma to retail. I, then, assign to each major the corresponding sectoral value added.

[23] I divide majors with restricted access in 4 groups: (1) medicine - medicine, pharmacy, veterinary, (2) humanities - Italian, history, philosophy, foreign languages, (3) law and political science, (4) architecture. Similarly, I divide STEM in: (1) physics, (2) mathematics, (3) sciences - geology, biology, natural science, and chemistry, (4) engineering.

points among high-achieving students and only 4.1 percentage points among low-achieving students. This finding suggests that the high-achieving students had the skills to succeed in different fields and that they potentially suffered the most from the deterioration of the signal attached to STEM degrees.

Estimating equation (E1) with the restricted majors divided into different subcategories make it possible to investigate where type A students moved after 1961 (Appendix Table E3). Humanities students moved in similar proportions towards medicine and humanities majors, while scientific students moved exclusively to medicine. In medicine, enrollment grew monotonically with precollegiate ability: among humanities students, for example, enrollment in medicine increased by 11.7 among high-achieving students and by only 6.1 percentage points among low-achieving students. In the humanities majors, instead, enrollment increased more among low-achieving students (+ 6.8 percentage points), compared with high-achieving students (+5.4 percentage points).

Table E1: Multinomial Logit, Probability of Type A Students Enrolling in STEM and Restricted Majors

	STEM Majors				Restricted Majors			
	All Type A		Males		All Type A		Males	
	Coeff.	Marginal Effects	Coeff.	Marginal Effects	Coeff.	Marginal Effects	Coeff.	Marginal Effects
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Post 61	0.294***	-0.009	0.259***	-0.002	0.449***	0.059***	0.373***	0.043***
	(0.079)	(0.013)	(0.096)	(0.016)	(0.078)	(0.014)	(0.099)	(0.016)
Post 65	0.183**	-0.085***	-0.141	-0.107***	0.803***	0.172***	0.548***	0.166***
	(0.080)	(0.013)	(0.095)	(0.015)	(0.078)	(0.013)	(0.096)	(0.015)
Male	0.429***	0.079***			0.012	-0.087***		
	(0.069)	(0.011)			(0.064)	(0.011)		
HS exit score	0.339***	0.070***	0.336***	0.089***	0.073**	-0.031***	0.011	-0.042***
	(0.035)	(0.005)	(0.043)	(0.006)	(0.035)	(0.005)	(0.045)	(0.006)
HS class score	-0.023	-0.004	-0.082	-0.020	-0.008	0.002	0.012	0.018
	(0.119)	(0.019)	(0.141)	(0.023)	(0.115)	(0.020)	(0.142)	(0.023)
Home schooled	-0.582***	-0.110***	-0.779***	-0.150***	-0.163	0.048**	-0.275**	0.052**
	(0.133)	(0.023)	(0.154)	(0.028)	(0.116)	(0.022)	(0.137)	(0.025)
Non-repeater	0.691***	0.125***	0.843***	0.180***	0.193**	-0.064***	0.145	-0.100***
	(0.104)	(0.019)	(0.121)	(0.022)	(0.093)	(0.018)	(0.110)	(0.020)
Mean, 1958–60	0.383	0.383	0.436	0.436	0.387	0.387	0.340	0.340
HS fixed effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Observations	11,425	11,425	7,632	7,632	11,425	11,425	7,632	7,632

Notes: Coefficients and marginal effects are estimated from a multinomial logit model where the choice is either a STEM, restricted, or nonrestricted major, or no university (baseline). STEM majors are engineering, physics, mathematics, biology, geology, natural science, chemistry, and agricultural science. The restricted majors are medicine, the humanities, political science, law, and architecture. Post 61_t is equal to 1 for the cohorts that completed high school between 1961 and 1964, while Post 65_t is 1 for the cohort that graduated starting in 1965. The omitted category is represented by the cohorts that graduated between 1958 and 1960. Robust standard errors in parentheses, *** p<0.01, ** p<0.05, * p<0.1.
Sources: School data of type A students who completed high school in Milan, Italy; 1958–1968.

Table E2: Type A, Decrease in STEM Enrollment after 1961

	Humanities high schools			Scientific high schools		
	Actual Shares	Predicted Shares	Difference	Actual Shares	Predicted Shares	Difference
	(1)	(2)	(3)	(4)	(5)	(6)
All type A students	0.192	0.294	-0.102***	0.501	0.532	-0.031***
HS exit score						
Quartile 1 (1-25)	0.149	0.218	-0.069***	0.425	0.439	-0.014**
Quartile 2 (26-50)	0.169	0.256	-0.087***	0.470	0.492	-0.022***
Quartile 3 (51-75)	0.196	0.302	-0.106***	0.518	0.554	-0.036***
Quartile 4 (76-100)	0.272	0.435	-0.163***	0.631	0.691	-0.060***

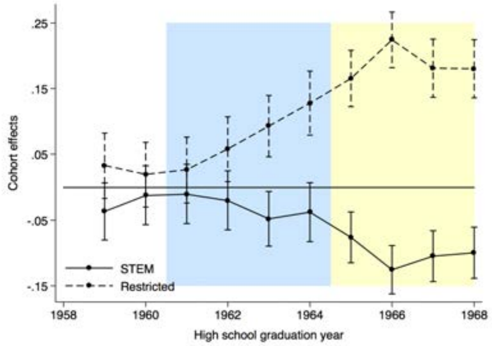
Notes: Columns (1) and (4) show the actual share of type A from the humanistic and scientific high schools choosing STEM fields after 1961. Columns (2) and (5) show the predicted share of type A students from humanistic and scientific schools that would choose a STEM major after 1961, using the coefficients estimated from equation (E1) with data of the pre-1961 cohorts. *** p<0.01, ** p<0.05, * p<0.1.
Sources: School data of type A students who completed high school in Milan, Italy; 1958–1968.

Table E3: Type A, Enrollment Shares in Restricted Majors After 1961

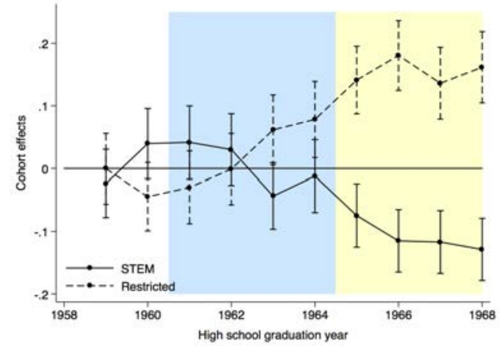
	Humanities high schools			Scientific high schools		
	Actual Shares	Predicted Shares	Difference	Actual Shares	Predicted Shares	Difference
	(1)	(2)	(3)	(4)	(5)	(6)
			Medicine			
HS score - Q1	0.138	0.077	0.061***	0.158	0.069	0.089***
HS score - Q2	0.139	0.055	0.084***	0.156	0.049	0.107***
HS score - Q3	0.141	0.034	0.107***	0.153	0.036	0.117***
HS score - Q4	0.131	0.014	0.117***	0.135	0.012	0.123***
			Humanities			
HS score - Q1	0.300	0.232	0.068***	0.010	0.007	0.003**
HS score - Q2	0.316	0.252	0.064***	0.011	0.010	0.001
HS score - Q3	0.350	0.290	0.060***	0.011	0.015	-0.004**
HS score - Q4	0.372	0.318	0.054***	0.011	0.017	-0.006***

Notes: See table E2. Two groups of restricted majors (Architecture, Law and PoliSci) are not reported. *** p<0.01, ** p<0.05, * p<0.1.

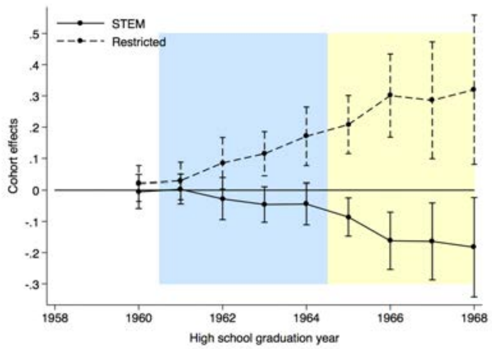
Figure E1: Multinomial Logit, Major Choice of Type A Students



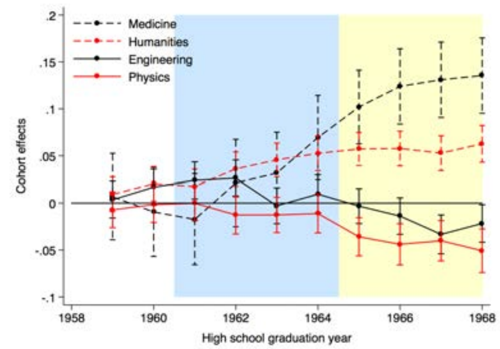
A. No precollegiate ability



B. Males



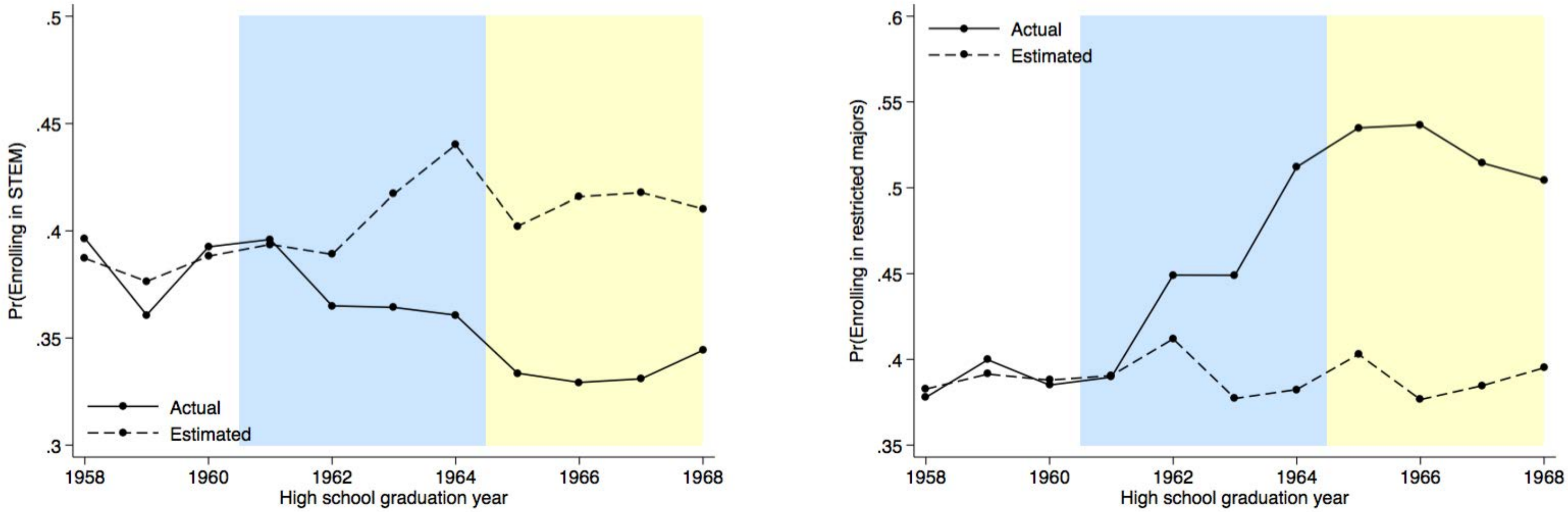
C. Conditional Multinomial Logit



D. Disaggregated Major Choice

Notes: Robustness checks of the marginal cohort effects on the major choice of type A students from equation (E1). The bars represent 95 percent confidence intervals. Panel A shows marginal cohort effects without controlling for measures of precollegiate ability. Panel B includes only males. Panel C shows marginal cohort effects from a conditional multinomial logit model with controls for concurrent returns to different majors. Panel D estimates a multinomial logit model with a more disaggregated major choice.
Sources: School data of type A students who completed high school in Milan, Italy; 1958–1968.

Figure E2: Predicted and Actual Probability of Type A Enrollment



A. STEM Majors

B. Restricted Majors

Notes: Actual and predicted probabilities of type A students enrolling in STEM and restricted majors. The predicted probabilities are constructed estimating the equation (E1), using only the cohorts that completed high school until 1961.
Sources: School data of type A students who completed high school in Milan, Italy; 1958–1968.

F Income Adjustment

The problem of estimating age-cohort effects in a cross-section

After controlling for observable characteristics, the difference between the average income of type B students who completed high school in year b, before the policy implementation, and b', after the policy implementation, is the function of two elements: $E(y_b - y_{b'} \mid X) = f(A\,,\,B)$, where A are age effects, and B are cohort effects.

Since most of the students in the dataset were between 55 and 67 years old in 2005, there are two main age effects (A) at play. Older cohorts were more likely to be retired. In Italy, pensions were computed as a fraction of the last 10 wages. For this reason, pension earners received a lower income, compared with similar individuals that are still in the labor force. In addition, income of younger cohorts could still be on an increasing trajectory.

The cohort effects (B), instead, measure the income change due to the fact that one group of students completed high school in year b (before the policy), while the other group graduated in year b' (after the policy).

Figure F1 (Panel A) visualizes this problem. In one cross section, the difference between two observations at different ages can be caused by age effects only, cohort effects only, or a combination of the two. In this paper, I am interested in isolating cohort effects. To do so, I need to predict the income that individuals from different cohorts would earn at the same age (65 years old, the retirement age for men). As Figure F1 (Panel A) suggests, cohort and age effects cannot be disentangle in a single cross section. To circumvent this problem, I use out-of-sample observations from repeated cross-sections to estimate the age effects. However, moving from one cross-section to repeated cross-sections is not a solution in itself. In fact, this procedure adds a dimension ("period" or "survey year" effects) that is a linear combination of age and cohort (cohort = survey year - age): this means that it is impossible to observe individuals from the same cohorts at different ages in the same survey year. As a consequence, period-cohort-age effects cannot be estimated simultaneously. Figure F1 (Panel B) represents this problem graphically. In this case, repeated cross sections and assumptions on the structure of period-cohort-age effects are necessary.

Repeated cross-sections: description of the sample

I pool 11 waves of the Bank of Italy's Survey of Household Income and Wealth (SHIW) collected between 1991 and 2012. This representative dataset of the Italian population contains information on 245,184 individuals. I keep household heads and their spouses/partners (91,700 observations deleted), individuals born between 1930 and 1955 years old (-72,871 obs), with at least a high school diploma (-57,265 obs), and with positive income (-2,879 obs). This procedure leaves 20,469 observations. Table F1 shows the characteristics of the cohorts born between 1930 and 1950 in the survey years 2004 and 2006, compared with the characteristics of the high school graduates from Milan.[24] The SHIW sample has a higher share of females. This is due to the fact that the SHIW sample contains graduates from all high schools, including the female-oriented education schools, while the sample of high school graduates from Milan is focused on high schools that were either equally split between men and women (type A and C schools) or men-only schools (type B). On average, individuals from the SHIW sample are less likely to have a university degree and earned lower incomes. These differences could be due to the fact that Milan has higher returns to university education, compared with the rest of Italy. Unfortunately, the sample cannot be restricted to individuals living in the north of Italy (the geographical aggregation included in the dataset that is closer to Milan), because of small sample size. The average income of university graduates is not statistically different across the two dataset.

Procedure and assumptions to disentangle age and cohort effects

As stated previously, the age effects are mainly two. Older cohorts had a higher probability of being pension earners. The income of younger cohorts, instead, could still be increasing with age. I find that the first effect is larger than the second.

In the empirical analysis, I estimate these two parts separately, because the SHIW contains detailed information about pensions. The procedure uses several assumptions: (1) period effects can be captured by macroeconomic indicators (the unemployment rate observed in the survey year), (2) age effects are constant across cohorts, (3) age effects have a specific functional form. In the next section, I will show how results would change with different assumptions or just using non-adjusted income.

First, I estimate (1) the probability of being a pension earner, (2) the ratio of pensions to total income, and (3) the replacement rate (the ratio of pensions to last wage) as a function of age, gender, completed education, unemployment rate (u_t), and birth year fixed effects (B_b):

[24] This subsample is close (both in characteristics and in time) to the group of income earners in 2005 that completed high school in Milan between 1958 and 1968. To improve precision, however, the estimating sample is bigger, including also individuals born between 1951 and 1955 and more survey years. The characteristics of the full sample are reported in the last column of Table F1.

$$p_{abt} = F(\text{age}_{abt}\,,\ \text{age}^2_{abt}\,,\ \text{age}^3_{abt}\,,\ \text{male}_{abt}\,,\ \text{college}_{abt}\,,\ \text{age}_{abt}\cdot\text{male}_{abt}\,,\ \text{age}_{abt}\cdot\text{college}_{abt}\,,\ u_t\,,\ B_b), \tag{F1}$$

where F is the logit function for the probability of being retired, and linear for the remaining two dependent variables. I set the probability of retirement equal to 1 above age 70. As expected, the estimated probability of retirement decreases among younger cohorts (Appendix Figure F2). Male individuals with a university degree that completed high school in 1958 have a 90.3 percent probability of being retired, while similar individuals that completed high school in 1968 have a 31.5 percent probability. Conditional on education, women have a higher probability of being retired. In the Italian system, in fact, women could retire at 55 years old, five years before men. Conditional on gender, high school graduates have a higher probability of being retired, relative to university graduates. A higher investment in education induces individuals to stay longer in the labor market to recoup the initial investment in human capital. In addition, it can also increase productivity later in life.

Second, I estimate how income increases with age in the years before retirement. On the subsample of individuals who are not retired and below 65 years old, I estimate the following income equation

$$y_{abt} = \alpha + \beta \text{age}_{abt} + \sum_b \gamma_b B_b + \delta u_t + \epsilon_{abt} \tag{F2}$$

separately for men (1) with and (2) without a university degree, and women (3) with and (4) without a university degree. B_b are cohort fixed effects, while u_t is the unemployment rate in each survey year. The estimated $\hat{\beta}$ varies with gender and completed education. One additional year increases income by €786 for male university graduates and by €524 for female university graduates. The coefficients are smaller for high school graduates: one additional year increases income of male high school graduates by (€202) and of female high school graduates by (€91).

Using these estimates, I adjust the taxable income in 2005 according to the following formula:

$$\tilde{y}_b = \hat{\pi}_b \cdot \left[\frac{\hat{\%}_b \cdot y_b}{\hat{\text{replace}}_b} + (1 - \hat{\%}_b) \cdot y_b \right] + (1 - \hat{\pi}_b) \cdot \left[y_b + \hat{\beta} \text{age}_b \right], \tag{F3}$$

where $\hat{\pi}_b$ is the probability of being retired, $\hat{\%}_b$ is the ratio of pensions to total income, and $\hat{\text{replace}}_b$ is the replacement rate.

To test the validity of these estimates, I split randomly the SHIW dataset into two groups: a portion of the observations (75 percent) is employed to estimate the age effects, while the remaining part is used to validate the results. The predicted and actual means are close. For example, the average predicted probability of retirement is 33.7 percent, compared with an actual 34.5 percent of individuals in the sample being retired. The difference is significant at the 10 percent level, but is small. The average incomes predicted by equation (F2) are above the actual means, but the difference is not significant for men with and without a university degree. For women with a university degree, however, the actual income is on average €3,095 (14 percent) lower than predicted.

I winsorize the adjusted income in 2005 at the 2nd and 98th percentiles to limit the influence of outliers on the analysis. Figure F3 compares the average adjusted and unadjusted income by cohorts, both winsorized. The plot shows that the adjustment increased income for both older and younger cohorts, but the first effect prevailed. The results in the paper can be replicated qualitatively using unadjusted income.

Sensitivity analysis

In this section, I show how different assumptions on the structure of period-cohort-age effects would change the results. For comparison, I use the estimation of equation (9), which measures the effects of the reform on the long-run income of type B students (Table F4). For each set of assumptions, I also plot the average adjusted income (Figure F4). This exercise shows that the different assumptions lead to results that are qualitatively similar.

A. Period effects do not exist

I estimate equations (F1) and (F2) excluding the unemployment rate. The results are virtually unchanged. The average adjusted income follows the same path and the adjustment is larger for older cohorts. The estimators of δ_t are very close to the baseline.

B. Normalization of period effects

Deaton and Paxson (1994) suggested a normalization for the period effects in which period dummies sum up to 1 and are orthogonal to a time trend. In practice, the normalization implies that any growth in income is attributed to age and cohort effects. I estimate equations (F1) and (F2) replacing the unemployment rate with these set of modified survey year dummies. The average adjusted income follows the same path. The estimators of δ_t are very close to the baseline.

C. Quadratic age effects on income

In this case, I estimate equation (F2) including both a linear and quadratic age variable. All the other assumptions are unchanged, which means that I assume that period effects are captured by the unemployment rate and age effects are constant across cohorts. The quadratic age component in equation (F2) lowers significantly the income increase in the years that lead to retirement, compared with the baseline. This implies that the adjustment for pension earners, in this case, is predominant. Age-adjusted income of younger cohorts is very close to the non-adjusted incomes, while there is a large gap among older cohorts. The estimators of δ_t are both negative and do not suggest that type B earned large returns to STEM degrees after 1965.

D. Age effects change with cohorts

I estimate equations (F1) and (F2) for separate cohorts grouped in 5-year bins. This procedure allows age effects to differ across (groups of) cohorts. All the remaining assumptions are unchanged. In this case, income is almost stable in the years that lead to retirement. The estimators of δ_t are both negative and do not suggest that type B earned large returns to STEM degrees after 1965.

Table F1: Summary Statistics of the SHIW Sample

	SHIW 1930-1950	HS Graduates from Milan	Difference	SHIW 1930-1955
	(1)	(2)	(3)	(4)
Birth year	1943.38	1943.20	0.18	1946.10
Males	0.64	0.75	-0.11***	0.61
University graduates	0.25	0.35	-0.10***	0.24
University graduates—males	0.24	0.33	-0.09***	0.24
Income	29,508	35,348	-5,840***	28,824
Income—males	34,703	39,730	-5,027***	34,372
Income—university graduates	42,038	43,403	-1,365	41,584
Income—male university graduates	53,162	49,838	3,326	52,416
Income—HS graduates	25,334	30,650	-5,316***	24,770
Income—male HS graduates	29,017	34,304	-5,827***	28,823

Notes: The table shows the characteristics of the cohorts born between 1930 and 1950 in the survey years 2004 and 2006. They represent a subsample of a larger estimating sample; details on the sample construction can be found in appendix F. Column (2) shows the characteristics of the dataset of students that completed high school in Milan between 1958 and 1968. The last column shows the summary statistics for the whole estimating sample from the SHIW; this includes individuals with a high school diploma born between 1930 and 1955. *** p<0.01, ** p<0.05, * p<0.1. Sources: Bank of Italy's SHIW; 2004–2006 waves.

Table F2: Marginal Age Effect on Income

	University Degree	High School Diploma
Males	786.30***	202.34***
	(191.77)	(58.11)
Females	524.43***	91.25**
	(92.69)	(39.78)

Notes: The table shows the $\hat{\beta}$ from the equation (F2), separately estimated on men (1) with and (2) without a university degree, and women (3) with and (4) without a university degree. *** p<0.01, ** p<0.05, * p<0.1.
Sources: Bank of Italy's SHIW; 1991-2012 waves.

Table F3: Income Adjustment, Actual and Predicted Age Effects

	Actual (1)	Predicted (2)	Difference (1) - (2)
Retired workers			
Probability of being retired ($\hat{\pi}_b$)	0.345	0.337	0.008*
Ratio of pensions to income ($\hat{\%}_b$)	0.742	0.723	0.019***
Replacement rate ($\widehat{\text{replace}}_b$)	0.752	0.751	0.001
Income of employed workers			
Male with university degree	43,892	46,078	-2,186
Male with HS diploma	28,694	28,885	-191
Female with university degree	21,556	24,651	-3,095***
Female with HS diploma	17,114	17,948	-834***

Notes: For a random subsample (25 percent) of the SHIW dataset, the table compares the actual means to those predicted with the coefficients estimated using the remaining 75 percent of the sample. The probability of being retired, the ratio of pensions to total income, and the replacement rate are estimated using equation (F1). Income is estimated using equation (F2). *** p<0.01, ** p<0.05, * p<0.1.

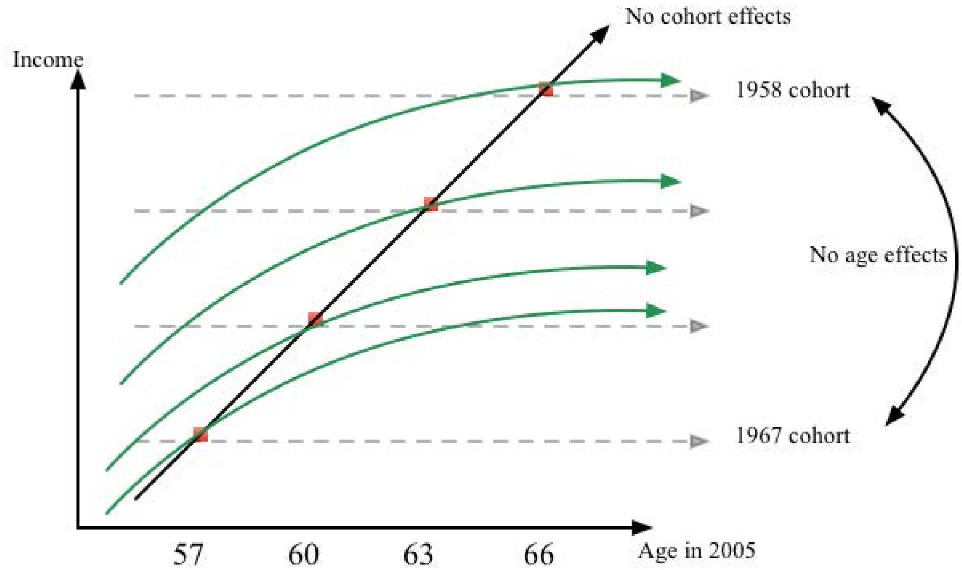
Sources: Bank of Italy's SHIW; 1991-2012 waves.

Table F4: Sensitivity Analysis of Assumptions on Age-Cohort-Period Effects

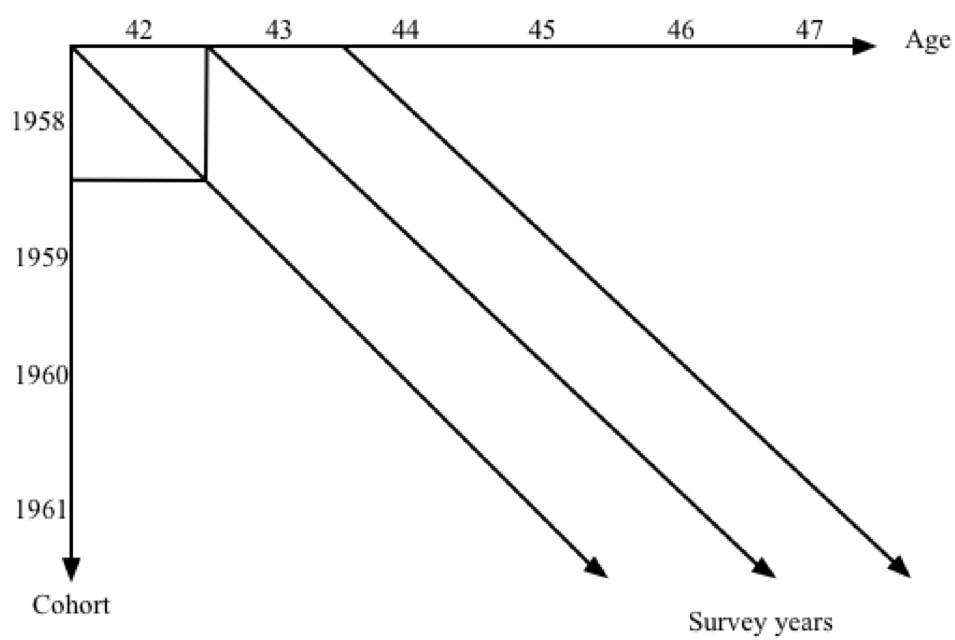
	Type B x Post 61 (1)	Type B x Post 65 (2)
Baseline	0.136**	0.059
	(0.066)	(0.066)
Not adjusted	0.154	0.027
	(0.123)	(0.129)
No period effects	0.136**	0.043
	(0.066)	(0.065)
Deaton and Paxson (1994) normalization	0.138**	0.070
	(0.068)	(0.066)
Quadratic age effects	-0.114	-0.359***
	(0.114)	(0.119)
Cohort-dependent age effects	-0.193***	-0.516***
	(0.054)	(0.055)

Notes: The table shows the estimation of equation (9) under different assumptions about the structure of age - cohort - period effects, as described in section F. *** p<0.01, ** p<0.05, * p<0.1.
Sources: School data of type B students that completed high school in Milan, Italy; 1958-1968.

Figure F1: Identification of Age and Cohort Effects



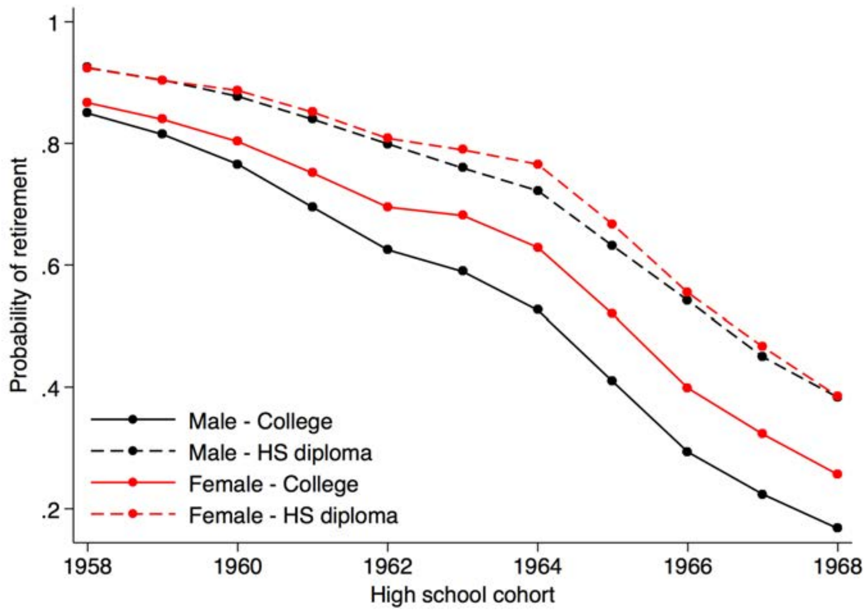
A. Age and Cohort Effects in One Cross Section



B. Age, Cohort and Year Effects in Repeated Cross Sections

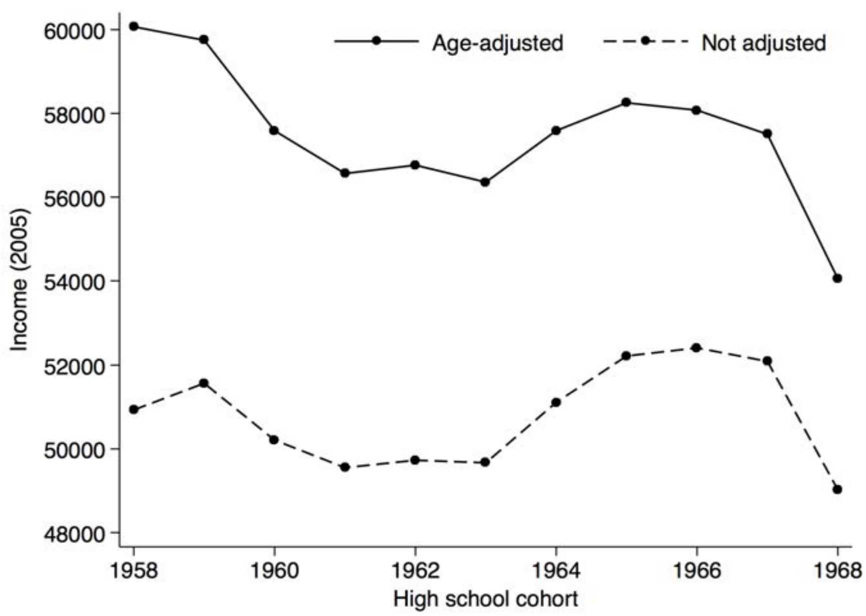
Notes: Panel A shows how age and cohort effects are confounded in a single cross-section. Panel B shows that in repeated cross-sections there is an additional dimension (period or survey year), which is a linear combination of cohort and age. Therefore, the simultaneous estimation of all three is not possible with additional assumptions.

Figure F2: Estimated Probability of Retirement



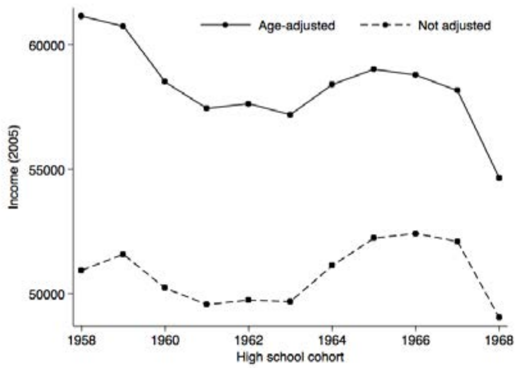
Notes: Estimated probability of retirement by completed education and gender. Probability of retirement is estimated by equation (F1).
Sources: Bank of Italy's SHIW; 1991-2012 waves.

Figure F3: Age-adjusted vs. Unadjusted Income

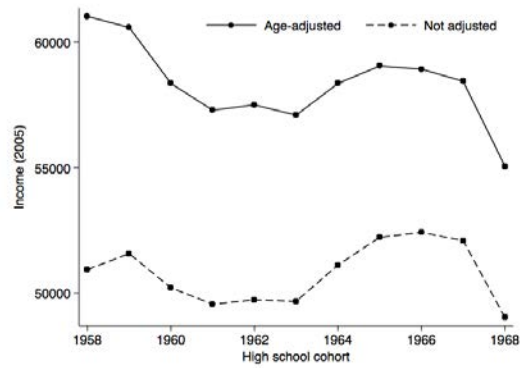


Notes: The figure shows means by high school cohorts. Income is adjusted to account for age effects using equation (F3).
Sources: Bank of Italy's SHIW; 1991-2012 waves.

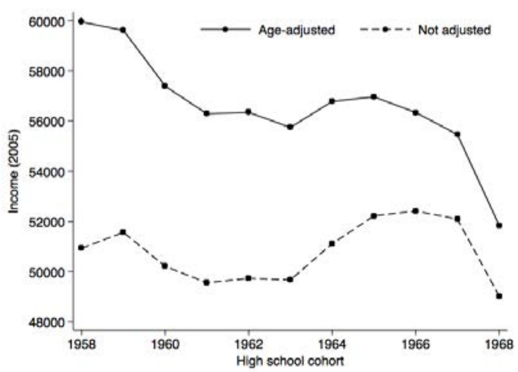
Figure F4: Age-adjustment: Different Assumptions on Age-Cohort-Period Effects



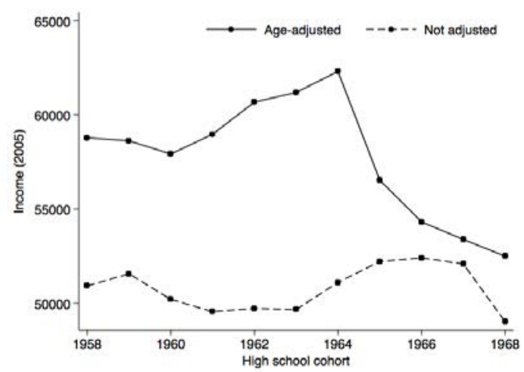
A. No Period Effects



B. Normalized Period Effects



C. Quadratic Age Effects



D. Varying Age Effects

Notes: The figure shows average adjusted and unadjusted income by cohorts, using different assumptions on the structure of age-cohort-period effects. The different assumptions are described in section F.

G Long-Run Income: Additional Results

Type A Students

Table G1: Income of Type A Students, Additional Results and Robustness Checks

	STEM x Post 61	STEM x Post 65	Obs.
	(1)	(2)	(3)
Linear and squared trends for ability levels	0.014	-0.144*	6,398
	(0.081)	(0.074)	
High school-specific linear trends	-0.009	-0.195**	6,398
	(0.083)	(0.079)	
Major–ability linear trends	-0.011	-0.200**	6,398
	(0.083)	(0.080)	
S.e. clustered at high school–graduation year level	0.017	-0.138*	6,398
	(0.069)	(0.070)	

Notes: Each row shows the results of a different specification of equation (8). Robust (or clustered, where specified) standard errors in parentheses, *** p<0.01, ** p<0.05, * p<0.1.
Sources: University transcripts of students who completed high school in Milan, Italy; 1958–1968.

Type B Students

In the text, I compared type B students with type C students. Alternative comparisons lead to similar findings (Appendix Table G2).

First, I can measure how outcomes of type B students who scored in the top half of their high school class (and, therefore, should have benefited more from the opportunity to enroll in STEM majors) changed across cohorts, relative to type B students who scored in the bottom half. As expected, the university graduation rate of high-achieving type B students increased by 4.8 percentage points from 1961 to 1964 and by 6 percentage points from 1965 to 1968. The long-run income of high-achieving type B students, however, decreased by 4.9 percent from 1961 to 1964 and by 0.8 percent from 1965 to 1968. Both coefficients are not statistically different from zero and each other.

Second, I compare type B with type A students. In the pre-reform cohorts, type A students were 56.2 percent more likely to enroll in university, 49.8 percent more likely to receive a university degree, and earned 39 percent more in 2005 relative to type B students. The enrollment gap fell by 15.9 percentage points from 1965 to 1968. Consequently, the graduation gap fell by 3.2 percentage points from 1965 to 1968. The long-run income of high-achieving type B students, however, increased by 10.2 percent from 1961 to 1964 and decreased by 7.4 percent from 1965 to 1968. The second coefficient is not statistically significant. These results are very similar to the baseline.

These results are robust to the inclusion of different trends, to the estimation on the pre-1966 cohorts, and to the use of sampling weights that match the average characteristics of pre-reform and post-reform observations.

Table G2: Income of Type B Students, Additional Difference-in-Differences Specifications

	Baseline		Trends		Pre-1966 Cohorts	Reweighted
	(1)	(2)	(3)	(4)	(5)	(6)
High-achieving vs low-achieving type B students						
High-achieving x Post 61	-0.049	0.030	-0.073	-0.083	-0.053	-0.042
	(0.083)	(0.121)	(0.124)	(0.125)	(0.084)	(0.083)
High-achieving x Post 65	-0.008	0.071	-0.057	-0.061	0.154	-0.012
	(0.079)	(0.117)	(0.216)	(0.217)	(0.115)	(0.079)
Pre-reform trend x High-achieving		0.079				
		(0.084)				
Type B vs Type A students						
Type B x Post 61	0.102*	0.161**	0.103*	0.219**	0.108*	0.116**
	(0.057)	(0.079)	(0.057)	(0.088)	(0.058)	(0.058)
Type B x Post 65	-0.074	-0.015	-0.070	0.151	-0.128*	-0.078
	(0.054)	(0.077)	(0.055)	(0.148)	(0.075)	(0.055)
Pre-reform trend x Type B		0.059				
		(0.056)				
Trends for quartiles of precollegiate ability	No	No	Yes	No	No	No
Trends for ability–high school	No	No	No	Yes	No	No

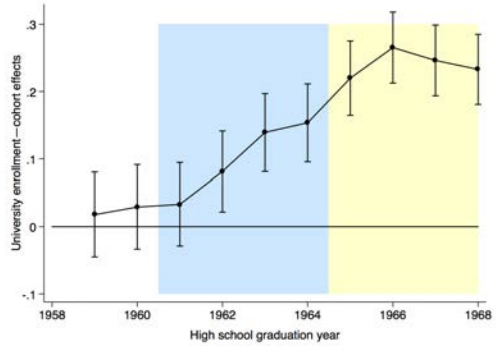
Notes: The coefficients are estimated from different specifications of equation (9). "High-achieving vs Low-achieving type B students" compares type B in the top half and bottom half of the ability distribution. Robust standard errors in parentheses, *** p<0.01, ** p<0.05, * p<0.1.
Sources: School data of type B students who completed high school in Milan, Italy; 1958–1968. Income tax returns in 2005.

Table G3: Cohort Effects, Education and Income of Type B Students

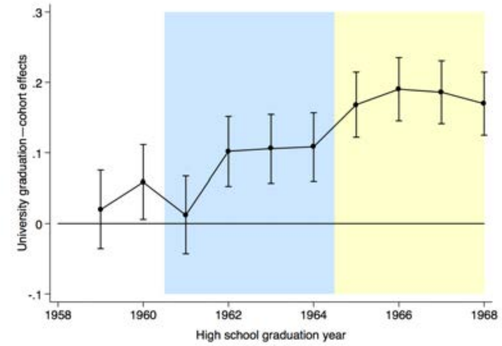
	University Enrollment		University Graduation		Log Income
	Coeff.	Marginal Effects	Coeff.	Marginal Effects	Coeff.
	(1)	(2)	(3)	(4)	(5)
Post 61	0.386***	0.090***	0.499***	0.059***	0.091**
	(0.070)	(0.016)	(0.111)	(0.013)	(0.042)
Post 65	0.965***	0.225***	1.271***	0.150***	0.015
	(0.065)	(0.014)	(0.101)	(0.012)	(0.040)
Male	0.444**	0.103**	0.317	0.037	0.952***
	(0.190)	(0.044)	(0.227)	(0.027)	(0.166)
HS exit score	0.324***	0.075***	0.438***	0.052***	0.112***
	(0.024)	(0.006)	(0.028)	(0.003)	(0.013)
HS class score	0.196**	0.046**	0.040	0.005	0.030
	(0.079)	(0.018)	(0.106)	(0.012)	(0.046)
Home schooled	-0.269**	-0.063**	-0.662***	-0.078***	-0.131*
	(0.105)	(0.025)	(0.171)	(0.020)	(0.068)
Non-repeater	0.499***	0.116***	1.004***	0.118***	0.192***
	(0.092)	(0.021)	(0.172)	(0.020)	(0.052)
Mean, 1958-60	0.292	0.292	0.082	0.082	10.447
HS fixed effects	Yes	Yes	Yes	Yes	Yes
Observations	8,791	8,791	8,791	8,791	7,381

Notes: Coefficients and marginal effects are estimated from $\text{outcome}_{it} = F(\alpha + \beta_1 \text{Post } 61_t + \beta_2 \text{Post } 65_t + \gamma X_{it})$, where outcome_{it} is a dummy for university enrollment in columns (1) and (2), a dummy for university graduation in columns (3) and (4), and log income in column (5). The function F is logit for university enrollment and graduation and linear for log income. Post 61_t is equal to 1 for the cohorts that completed high school between 1961 and 1964, while Post 65_t is 1 for the cohort that graduated between 1965 and 1968. The omitted category is represented by the cohorts that graduated between 1958 and 1960. X_{it} is a set of student characteristics. Robust standard errors in parentheses, *** p<0.01, ** p<0.05, * p<0.1.
Sources: School data of type B students who completed high school in Milan, Italy; 1958–1968. Income tax returns in 2005.

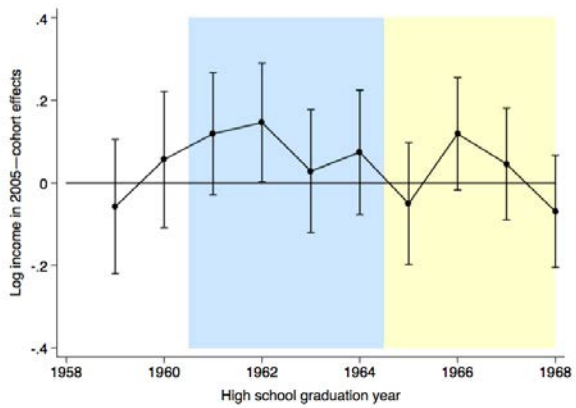
Figure G1: Cohort Effects, Education and Income of Type B Students



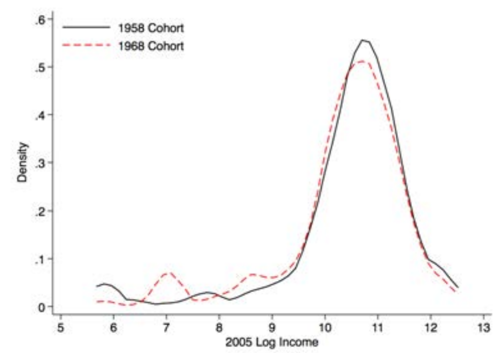
A. University Enrollment



B. University Graduation



C. Income in 2005



D. Income Distribution

Notes: The marginal effects (Panels A–C) are computed from outcome$_{it} = F(\alpha + \sum_t \beta_t Y_t + \gamma X_{it})$, where outcome$_{it}$ is either a dummy for university enrollment, a dummy for university graduation, or log income. The function F is logit for university enrollment and graduation and linear for log income. Y_t is a set of year of high school graduation fixed effects with 1958 as omitted category. X_{it} is a set of student characteristics. The bars represent 95 percent confidence intervals. Panel D shows the income distribution for the 1958 and 1968 cohorts of type B students.

Sources: School data of type B students who completed high school in Milan, Italy; 1958–1968.

H Cost-Benefit Analysis

In this section, I examine whether the costs of keeping quality of education and class heterogeneity at their pre-reform level would have surpassed the benefits for STEM students. Initially, I compute the discounted present value at age 25 (age of university graduation) of the income losses caused by lower human capital. The average long-run income of STEM students in pre-reform cohorts was equal to €57,632. I assume that the effects of lower quality of education (-12.9 percent) and higher class heterogeneity (-12.6 percent) do not change with age, that income increases by 1.8 percent every year, and that the discounting rate is equal to 10 percent.[25] Based on these assumptions, the yearly income loss at age 25 was €3,642 due to lower quality of education and €3,557 due to higher class heterogeneity. Their discounted present values over the lifetime (in this case, until retirement at age 65) were €42,412 and €41,425 respectively.

At this point, I estimate the costs of hypothetical actions that the Italian government could have taken to prevent lower human capital. I only consider the costs of hiring more teaching fellows (professors and assistants). Although this procedure involves less arbitrary assumptions, it could underestimate the total costs. In addition, I will use the empirical model described in section 4 to base my analysis. This will lead to some simplifications. For example, the model suggests that the only way to keep class heterogeneity fixed is to establish a strict tracking system in which type A and B students do not interact. This, however, ignores any potential benefit of more diverse classes. Moreover, I will assume that quality of education depends linearly on the student-faculty ratio.

First, I consider a scenario in which the government intends to keep class heterogeneity unchanged but allows the quality of education to vary. For each STEM course, the government creates a separate section for type B students and assigns only one professor to it. In doing so, the government hires the minimum amount of faculty to keep type A and B students in separate classes. For a degree with 25 university courses, this plan costs €690,875 or €5,074 per student.[26] In this case, the benefits of keeping class heterogeneity fixed (€41,425) are far greater than the costs (€5,074) for the government.

Second, I consider an alternative scenario in which the government wants to keep the quality of education constant but allows class heterogeneity to increase. Specifically, type B students join the lectures attended by type A, but the government hires new teaching

[25] The yearly increase of income is estimated using SHIW data on males with a university degree (appendix F). The discounting rate is computed as the average discount rate in Italy between 1958 and 2003 (available from the Bank of Italy at this link).

[26] All the figures are in 2005 €. The average annual salary of a professor in 1965 was equal to €27,635 (Supplemento alla Gazzetta Ufficiale 108, 30/04/1965, Tabella C). In Milan, the average STEM major had 136.16 students enrolled in each academic year.

assistants to keep the student-faculty ratio fixed. In Milan, the average STEM major had 4.81 teaching fellows per course, 71.91 students per year before 1965, and 64.25 new type B students per year after 1965. To keep the student–faculty ratio constant, the government needs to hire 4.3 assistants per course. In a degree with 25 courses, this plan costs €1,666,406 or €12,239 per student.[27] Also in this case, the benefits of keeping quality of education fixed (€42,412) are far greater than the costs (€12,239) for the government.

Third, I consider the case in which the government intends to keep both the quality of education and class heterogeneity fixed. The plan consists in creating two sections for each course, one for type A and one for type B students, in which the student–faculty ratios do not exceed the pre-reform level. The government still needs to hire 4.3 new teaching fellows per course: in this case, however, 1 professor responsible for the teaching and 3.3 new assistants per course. In a major with 25 courses, this plan costs €1,969,531 or €14,465 per student. In this last case, the benefits of keeping both the quality of education and class heterogeneity fixed (€83,837) are higher than the costs (€14,465) for the government.

References

Baffigi, Alberto. 2011. "Italian National Accounts, 1861-2011." Bank of Italy, Economic History Working Paper 18.

Deaton, Angus, and Christina Paxson. 1994. "Intertemporal Choice and Inequality." *Journal of Political Economy*, 102(3): 437–467.

Marbach, Giorgio, Alfredo Rizzi, and Tommaso Salvemini. 1969. *Gli assistenti universitari e i liberi docenti in Italia.* Roma:Istituto di statistica, Facoltà di scienze statistiche, demografiche ed attuariali, Università degli studi di Roma.

[27] The average annual salary of an assistant in 1965 was equal to €15,510 (Supplemento alla Gazzetta Ufficiale 108, 30/04/1965, Tabella C; Marbach, Rizzi and Salvemini 1969).