The General Equilibrium Effects of Educational Expansion*

Job Market Paper

Nicola Bianchi, Stanford University

January 12, 2015

Abstract

In an effort to raise skills or promote equality, states sometimes engage in sweeping reforms that rapidly increase access to education for a significant share of their population. Such reforms are hard to evaluate because they may alter more than the outcomes of marginal students induced to enroll. They may change returns to skill, school quality, peer effects, and the educational choices of apparently inframarginal students (those who would have enrolled in the absence of the reform). I identify such general equilibrium effects by examining a dramatic 1961 Italian reform that increased university enrollment in science, technology, engineering, and math (STEM) fields by more than 200 percent in a few years. The peculiar features of the reform allow me to identify students who were unaffected, directly affected, and indirectly affected. They also allow me to identify key channels through which the effects ran. Using data I collected from tax returns and hand-written transcripts on more than 27,000 students, I show that the direct effects of the reform were as intended: many more students enrolled and many more obtained degrees. However, I also find that those induced to enroll earned no more than students in earlier cohorts who were denied access to university. I reconcile these surprising results by showing that the education expansion reduced returns to skill and lowered university learning through congestion and peer effects. I also demonstrate that apparently inframarginal students were significantly affected: the most able of them abandoned STEM majors rather than accept lower returns and lower human capital.

JEL Classification: I21, I24, I28, J24, N34.

Keywords: educational expansion, education policy, higher education, human capital, quality of education, class heterogeneity.

*Email: nbianchi@stanford.edu. Mailing address: 579 Serra Mall, Stanford, CA 94305. The latest version of this paper may be found at http://www.stanford.edu/∼nbianchi. I thank Ran Abramitzky, Caroline Hoxby, Petra Moser, and Luigi Pistaferri for ongoing support and discussion. I also thank Michael Dinerstein, Pascaline Dupas, Michela Giorcelli, Guzman Gonzalez-Torres, Daniel Grodzicki, Davide Malacrino, Giovanni Peri, Pietro Tebaldi, Stephen Terry, Constantine Yannelis, and seminar participants at Stanford, UC Davis, UC Berkeley BEHL, 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. Acknowledgements for the data collection can be found in appendix B.

I study the effects of an Italian reform that led to a 216 percent increase in enrollment in university STEM (Science, Technology, Engineering, and Mathematics) programs over a mere eight years. Like many reforms being tried or considered today, the policy was intended to increase the share of workers with a university education, to increase the share with STEM skills specifically, and to raise incomes. These intended effects are what economics would lead us to expect if the reform had had only partial equilibrium effects.

However, reforms of this magnitude can have substantial general equilibrium effects so that the actual consequences differ substantially from the intended ones. Reforms that expand university enrollment can increase the number of workers with a postsecondary degree. The consequent higher supply of university-educated labor may drive down its wage (Heckman, Lochner and Taber, 1998*a*). Enrollment expansions can also alter the education production function, reducing returns before individuals reach the labor market.[1] For example, enrollment expansions can cause university resources to be overcrowded, reducing the quality of education (Bound and Turner, 2007). It can increase the degree of heterogeneity of university students, especially when the incremental students tend to be less well prepared. In the resulting classrooms, learning will be difficult if teaching is less effective with heterogeneous students (Duflo, Dupas and Kremer, 2011) or if the less prepared students exert negative peer effects (Lavy, Paserman and Schlosser, 2012). Finally, talented students who would otherwise have enrolled in the expanded programs (in this case, STEM) may shift away toward other programs and activities, depriving the target area of some of its best learners. In short, the general equilibrium effects of an enrollment expansion can play out through several channels: skill prices, educational resources per student, class heterogeneity and other peer effects, changes in the enrollment of students who were apparently inframarginal to the policy. In this paper, I analyze the role of each of these channels in explaining why Italy's reform failed to raise incomes (as shown below).

In any study of the general equilibrium effects of a reform, econometric identification poses the main obstacle. Large-scale policies often affect all individuals simultaneously so that the analysis must rely entirely on cross-cohort comparisons. For the reform I study, however, this problem does not exist. Rather, Italy's reform, although massive, made only a specific type of high school graduates eligible for university and allowed them to enroll in STEM majors only–not in the "restricted" majors which included law, medicine, architecture, social sciences,[2] and the humanities. These features of the reform allow me to employ a variety of empirical strategies that exploit cross-sectional as well as cross-cohort variation.

[1] Since these effects are not market-based, they are not strictly general equilibrium effects. For example, Abbring and Heckman (2007) refer to them as social interactions. For sake of simplicity, I will call general equilibrium effect any effect generated by higher university enrollment.

[2] Business economics is not included in the social sciences.

Italian high schools differ in their curricula. Up until 1960, a student who graduated from a university-prep high school (hereafter "type A") could enroll in university in any major. A student who graduated from a high school for industry-sector professionals (type B) could enroll in only a few majors and most often did not enroll in university at all (there were also other professional and vocational schools that were not directly affected by the reform).[3] In 1961, the Italian government allowed type B graduates to enroll in STEM majors at universities for the first time. Although temporary caps somewhat limited the growth of type B enrollment through 1964, freshman enrollment in STEM programs had increased by 216 percent by 1968. After this, enrollment stabilized substantially. Since type B students were, on average, less prepared for university and from less advantaged backgrounds, they changed not just the size of universities but also the composition of their student bodies. Throughout the period, the restricted majors remained inaccessible to type B students, a fact that is important to my empirical strategies. Moreover, there is no evidence that students changed how they chose high schools after 1961.

To analyze the reform, I collected and digitized administrative data for the population of 27,236 students who completed high school in Milan between 1958 and 1968. I chose Milan because it is Italy's commercial capital and second largest city. 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. I collected data from high school registries, which include the grade students receive on the exit exam. I also collected university transcripts, which contain detailed information on each course students attend, each exam they take, and the degrees they attain. I merge these data with the (former) students' 2005 income information from their tax returns. Since the individuals were between 56 and 67 years old in 2005, the income I observe is a measure of long-run earnings potential or permanent income.

I first estimate the direct effect of the policy on type B students and find that, compared with pre-1961 cohorts, they were 9 percentage points more likely to enroll from 1961 to 1964 and 22.5 percentage points more likely from 1965 to 1968.[4] Type B students were 5.9 percentage points more likely to receive a university degree from 1961 to 1964 and 15 percentage points more likely from 1965 to 1968. These estimates imply a 72 to 183 percent increase in university graduation, compared with a baseline probability of 8.2 percent.

Despite their increased university enrollment and degrees, I find little evidence that type B students gained positive returns to university education. Especially after 1965, when the

[3] More details can be found in section I.A.

[4] A cohort is a group of students who completed high school in the same year. I will refer to cohorts simply with the corresponding school year. For example, the phrase "from 1961 to 1964" refers to the 1961 through 1964 cohorts of B type high school graduates.

increase in education was at its highest, I find no evidence that type B students earned higher incomes than the pre-1961 cohorts. This is an important result because it suggests that the general equilibrium effects were very different than what partial equilibrium analysis would lead us to expect. I therefore subject this finding to a variety of specification tests including difference-in-differences estimates among type B students who were more and less likely to enroll and between type B and type A students. The finding of no effect on earnings is robust.

To explain these results, I lay out a simple model in which enrollment expansions affect returns through three main channels. First, they potentially decrease the educational resources per student, a proxy for the quality of education. This may occur if resources do not adjust elastically so that higher enrollment induces crowding. The result may be lower human capital accumulation per unit of education attained. Second, enrollment expansions may increase the heterogeneity of students' preparation and family background. This could make learning more difficult through mechanisms such as student-to-student interactions and disruption of teaching practices. Third, enrollment expansions may depress skill prices because they increase the supply of human capital, assuming that demand is downward sloping. My model also predicts that some type A students who would otherwise enroll in STEM programs would be induced to enroll elsewhere in order to avoid the negative indirect effects from resource crowding, peers, and skill prices.

Testing the last prediction, I find that the probability of STEM enrollment among type A students decreased by 0.9 percentage points from 1961 to 1964 and by 8.5 percentage points from 1965 to 1968. Their probability of enrolling in restricted majors increased by 5.9 percentage points from 1961 to 1964 and by 17.2 percentage points from 1965 to 1968. These findings cannot be explained by changes in the students' characteristics. Moreover, the best prepared type A students were the most likely to switch from STEM to restricted programs, suggesting that their greater aptitude made them able to do well in a wider array of fields or that they suffered the most from the deterioration of the signal (of high aptitude) attached to a university STEM degree. Whatever the mechanism, the result was an especially pronounced decrease in talented type A students with STEM education.

Next, I show how human capital accumulation changed in STEM majors after the entry of type B students. The unit of observation is a student - university course - academic year combination. I measure human capital using the grade (based on an absolute grading scale) received in each course.

Initially, I estimate that educational expansion severely crowded resources per student, a measure of quality of education, and led to lower human capital. I exploit within compulsory courses - major - year variation in the congestion of university inputs generated by increased enrollment. In the Italian universities, in fact, courses could have different amounts of resources (number of assistants to the teaching faculty), in spite of similar enrollment. These imbalances depended on the seniority of the faculty[5], but did not necessarily reflected a differential need for inputs. Educational expansion, combined with constant university inputs, led to more congestion (and, consequently, lower quality of education) in courses that had fewer resources in 1961. In the empirical analysis, I estimate how grades changed after 1961 in compulsory STEM courses with high pre-existing student-faculty ratios (large expected quality decrease), relative to courses in the same major with low student-faculty ratios (smaller quality decrease). I find that, due to lower quality of education, grades decreased by 0.11σ from 1961 to 1964 and by 0.10σ from 1965 to 1968 in the average STEM course. These results should be interpreted as a lower bound of the true crowding effect of educational expansion. Quality of education, in fact, can decrease with crowding on many dimensions other than the student-faculty ratio.

In addition to crowding university resources, educational expansion dramatically changed the composition of students in university STEM programs. To estimate the effect of higher class heterogeneity, I exploit within compulsory courses - major - year variation in the preparedness of type B students. Some fields, in fact, were taught in type B schools (mainly the applied disciplines), while others were not (math, physics, and the formal sciences). Courses in STEM majors, however, were tailored specifically to type A students and their syllabus did not change after 1961. The entry of type B students, then, was more likely to generate negative peer effects and disrupt teaching practices in the STEM courses that were not taught in type B schools. For example, engineering students had to take "Technical drawing" (taught in type B schools) and "Math I" (not taught in type B schools) during their freshman year. Within each STEM major, I study how grades of type A students changed after 1961 in courses like "Math I", relative to courses like "Technical drawing". I find that grades of type A students decreased by 0.07σ from 1961 to 1964 and by 0.09σ from 1965 to 1968 in courses not included in the pre-collegiate curriculum of type B students.

After finding that human capital decreased in STEM majors, I show that, among type A students, returns to a STEM education declined after 1961. This result could explain why type B students did not gain positive returns to STEM degrees and why the best type A students left for majors with restricted access. Initially, I examine how the long-run income of type A graduates with a STEM education changed after 1961, compared with other type A graduates. I find that the positive income premium associated with a STEM university education, which was equal to 28 percent in the pre-reform cohorts, decreased by 12.6 percentage points from 1965 to 1968. The estimated decline is larger if I control for the

[5] Senior faculty could hire more assistants. Section V.B describes how resources were assigned to courses.

positive selection of type A students out of STEM majors. This last result might suggest that the reform increased the mismatch between students' skills and university majors.

As a summary of the main findings, I use the model to relate changes in human capital to changes in long-run income. I estimate that lower human capital could account for 42.9 percent of the total decline in STEM returns. Specifically, crowding of university resources could explain 24.6 percent of the decline, while higher class heterogeneity could explain another 18.3 percent. The remaining share can be attributed to lower skill prices or possibly other minor channels of general equilibrium effects. I compare these income losses with the likely costs (salary of new hires) that the government would have incurred to keep quality of education, class heterogeneity, or both fixed. In the last two cases, the costs of preventing lower human capital would have been smaller than the potential benefits.

This paper is related to three strands of literature. Heckman, Lochner and Taber (1998*a*) build a life-cycle model, in which educational expansion affects skill prices through changes in the relative supply of different types of human capital.[6] Similarly, Duflo (2004) shows how a school construction program in Indonesia decreased the wages of older and untreated individuals by increasing the supply of educated workers in the economy. My paper adds to these findings, highlighting how educational expansion can decrease the returns to education also through lower human capital accumulation.

A few papers explore how cohort size affects academic outcomes (Stapleton and Young, 1988; Card and Lemieux, 2001 Bound and Turner, 2007).[7] Bound and Turner (2007) show that larger cohorts within US states overcrowd public postsecondary institutions' resources and, therefore, have lower graduation rates. My paper relates to this literature by attempting to estimate how educational expansion congests university resources. However, by combining a shock to the number of university students together with pre-existing differences of school inputs across university courses within a major, I do not rely primarily on cross-cohort variation. I am able to identify the effect of crowding based solely on within-major, within-cohort, between-course variation.

A large literature examines how the composition of a class affects students' outcomes.[8] Most of these papers exploit natural variations in class composition across the units of

[6] Heckman, Lochner and Taber (1998*b*) use the same model to simulate the effects of a tuition subsidy on university enrollment. Few papers have extended these results: Lee (2005) assumes perfect substitutability of different types of human capital, Lee and Wolpin (2006) introduce adaptive expectations and a multi-sectoral economy with aggregate shocks, Abbott et al. (2013) model endogenous parental transfers.

[7] My paper is also related to the larger literature that examines the effects of school inputs on educational outcomes (for example, Card and Krueger, 1992; Loeb and Bound, 1996; Case and Deaton, 1999).

[8] For example, papers explore the effect of gender (Hoxby, 2000; Lavy and Schlosser, 2011; Anelli and Peri, 2013), ability (Figlio and Page, 2002; Lefgren, 2004; Duflo, Dupas and Kremer, 2011; Lavy, Paserman and Schlosser, 2012), and racial (Hoxby, 2000; Hoxby and Weingarth, 2006; Card and Rothstein, 2007; Cooley, 2010) composition of a class on academic outcomes.

observation (classes, schools, neighborhoods, or coarser aggregations) or time, while a few rely on policy changes (Hoxby and Weingarth, 2006; Cooley, 2010) or randomized control trials (Duflo, Dupas and Kremer, 2011). This literature focuses primarily on pre-collegiate education, in part because self-selection into different universities or majors makes identification more problematic. De Giorgi, Pellizzari and Woolston (2012), however, estimate the effects of class composition in university in a context with random assignment of students across classes within the same major.[9] My identification strategy is novel because I can rely on within-major, within-cohort, between-preparedness variation.

The rest of the paper is organized as follows. Section I describes the policy change. Section II describes the data. Section III analyzes the direct effect of the policy on type B students. Section IV presents the general equilibrium model. Section V shows evidence that educational expansion lowered returns to a STEM education. Section VI relates changes of human capital to income changes. Section VII concludes.

I Institutional Details

I.A The Italian High School System

Italian high schools offer different diplomas. General education schools (*licei*; for sake of simplicity, type A schools) focus on either the humanities (*licei classici*) or the sciences (*licei scientifici*) and traditionally prepare students for university. Their curricula range from philosophy, Latin, and Ancient Greek to mathematics and physics (Appendix Table A.1) and largely overlap across the two categories of type A school. In addition, there are technical high schools, which train professionals for specific economic sectors.[10] Two of these schools are industrial schools (*istituti industriali*; type B), which prepare students for jobs such as chemists and surveyors, and commercial schools (*istituti commerciali*; type C), which prepare students for jobs in accounting and the service sector. The technical curricula, therefore, focus on applied disciplines.

Students choose a high school at age 14, after completing compulsory schooling. 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. Type A schools provide a better preparation for most university majors, but are characterized by a heavy workload

[9] The literature of peer effects in university is mainly based on random housing assignment. However, this empirical strategy does not identify the effect of one's classmates, but the effects of one's roommates or acquaintances.

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

and are more challenging. 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.[11] In Italy, family background is a strong predictor of the high school choice. For instance, using a representative sample of Italians[12] born between 1931 and 1950 and with at least a high school diploma to measure the correlation between the high school attended and parental characteristics, I find that 74.9 percent of technical graduates had a father with 8 or less years of completed education, compared with only 53 percent of type A (Appendix Table D.1). Similarly, 40.8 percent of type B students had a father with a blue-collar job or unemployed, compared with 27.5 percent of type A. These differences are large and statistically significant. Maternal background shows the same pattern. 80.1 percent of technical graduates had a mother with 8 or less years of completed education, compared with 66 percent of type A.

I.B Expanding Access to University STEM Programs

After the end of World War II, the demand for workers with technical education probably increased significantly. In 1950, type B students were only 1.5 percent of the population between 14 and 18 years old, while type A students were 4.1 percent (Appendix Figure A.1). In the following 10 years, enrollment in type B schools grew at a much faster rate than enrollment in type A schools. By 1960, type B students were 4.7 percent and type A 5.7 percent of the relevant population . Despite these changes in the number and composition of high school graduates, access to university was still regulated by laws enacted during the Fascist regime. Type A graduates could enroll in any university major, while type B could enroll only in business economics and statistics.[13] These restrictions made the Italian university accessible to only a small share of high school graduates. In 1960, high school students were 18.6 percent of the population between 14 and 18 years old, while university students were only 4.1 percent of the population between 19 and 24 years old.

In July 1960, a shift of the coalition government towards center-left parties put university reform on the political agenda. In July 1961, law 685 allowed type B students to enroll in university STEM programs for the first time.[14] Specifically, the majors that became

[11] The school location suggests that they cater to different populations of students. Type A schools tend to be located near central and more well-off areas, while type B schools are closer to industrial zones. Appendix Figure B.1 shows the location of schools in Milan in 1958.

[12] Data come from the Survey of Household Income and Wealth (SHIW), administered biannually by the Bank of Italy. Appendix D provides more details about the construction of the sample.

[13] 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 had access to this major, even though only type C graduates had received a pre-collegiate preparation in accounting and, therefore, were interested in this major.

[14] Legge 685/61, http://www.edizionieuropee.it/data/html/31/zn57_11_044.html.

accessible to type B students were engineering, mathematics, physics, chemistry, biology, geology, natural science, and agricultural science. The reform applied retroactively: all students with a type B diploma, even if received before 1961, were allowed to enroll in STEM majors. Until 1964, the number of type B students was capped and applicants were selected with an admission test. However, as mandated by the original law, the enrollment cap was 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 the 19-year-old population, STEM freshmen increased from 1.5 percent in 1960 to 2.4 percent in 1964 (Appendix Figure A.2). 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. The time series shows two discontinuities generated by the policy: the 16.4 percent yearly increase in 1961 and the 64.4 percent yearly increase in 1965. Data that categorize freshmen by high school diploma indicate that the increase was driven by the entry of type B students after 1961. Type B freshmen in STEM majors rose from 2,713 in 1963 to 17,310 in 1968 with an overall 6.4-fold increase. This led to a major compositional change of the student population. In 1964, there were 0.28 type B for every type A freshman in STEM fields (Figure 1, Panel B). By 1967, type B freshmen had become more numerous than type A. The policy was not only associated with an increase in the incoming flow of STEM students. After 1961, the stock of enrolled students and the number of graduates rose accordingly. The total number of students enrolled in STEM fields increased from 43,251 in 1960 to 105,083 in 1968, a 143 percent change. The number of STEM graduates moved from 5,362 in 1960 to 10,196 in 1968, a 90.2 percent increase.

The policy left many majors accessible only to type A students: law, medicine, the humanities, the social sciences (with the exclusion of business economics, accessible to type B before 1961), and architecture. Freshmen in majors with restricted access increased by 24 percent from 20,382 students in 1960 to 25,280 in 1964. After 1965, they increased by 64.3 percent to 41,533 students in 1968. The total number of enrolled students rose by 83.8 percent, from 79,169 in 1960 to 145,477 in 1968. Graduates increased from 11,668 in 1960 to 20,670 in 1968, a 77.2 percent increase.

I.C Effects on High School Choice

The reform increased the value of a type B diploma and gave students an alternative path to pursue a STEM degree. Therefore, the policy might have affected how students sorted into type A and B schools: more specifically, some students who would have enrolled in a type A school without the policy might have decided to enroll in a type B school. Aggregate data, however, do not show any discontinuity in high school enrollment after 1961. The number of type B students started increasing steadily from 1952 and the growth rate slowed down after 1964 (Appendix Figure A.1). On the other hand, the number of type A students grew at a slower rate from 1950 to 1962, but started increasing at a faster pace afterwards.

In addition, I use the Bank of Italy's SHIW data to test whether paternal characteristics (education and occupation) of students who enrolled in high school after the policy implementation changed systematically, compared with earlier cohorts (Appendix Table D.2). With high school lasting 5 years and transfers across school types not admissible, the 1966 cohort was the first who knew about the policy at the time they enrolled in high school. Technical students (graduates of type B, type C, and other technical schools) who graduated after 1965 were 6.3 percent less likely to have a father with a high school diploma, relative to earlier cohorts. The effect is significant at the 10 percent level (p-value 0.095), but is not matched by an offsetting increase in the same probability for type A students, suggesting that all students in later cohorts had more educated fathers (rather than students with more educated fathers switching to type B schools). Furthermore, the average paternal occupation did not change after 1965 for type A and type B students. Overall, looking at changes in parental characteristics of enrolled students, there is no evidence of a significant enrollment shift from type A to type B schools.

It might surprise some readers that after 1961 students did not increase their enrollment in type B schools, relative to type A. The two schools, however, differed on many aspects that were not affected by the reform. First, type A schools provided a solid preparation for university, while type B schools did not teach important disciplines like math and physics. Second, type A schools were still granting access to a wider array of majors. For a high-achieving student, especially one who was not sure about what major to pursue in university, a type A school was clearly the best choice even after 1961. Low-achieving low-income students were less likely to attend a type A school to begin with. On the other hand, low-achieving high-income students were probably attending type A schools primarily for non-academic reasons (parental pressure, the social prestige of a type A diploma) and were not inclined to switch to type B schools, regardless of the option value of a type B diploma.

II Data

To analyze the effects of the policy, I collected administrative data for the 27,236 students who completed high school between 1958 and 1968 in Milan, Italy. Milan is the second largest city in Italy and the capital of Lombardy, a region in the Northwest of Italy. Traditionally an industrial powerhouse, Milan is also a major financial and commercial center. Disposable income per-capita is the highest in Italy at €25,866, 49.2 percent higher than the national average of €17,336 (Unioncamere, 2013). Similarly, 38 percent of its adult population has at least a high school diploma, compared with 33 percent for Italy (2001 Census).[15] This section describes the data collected from three sources: high school registries, university transcripts, and tax returns.

II.A 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 B.1). Two schools did not participate and in few isolated cases the registries of single academic years could not be located in the archives of participating schools: for this reason, the coverage of the dataset is equal to 74 percent. 66.8 percent of type A students are male, compared with 98.6 percent of type B (Table 1). Type B students are one average one year older than type A.

At the end of senior year, students had to pass a national examination (*maturità*) in order to graduate. The registries report the outcome of the exit exam for each discipline in the curriculum: a numerical score from 0 to 10 with 6 as passing grade. On average, type A students graduated with a 6.48 GPA and type B with a 6.36 GPA. For each student, I compute the GPA and I standardize it by year and school. The registries identify home schooled students (7.2 percent of type A and 8.4 percent of type B) and repeaters (9.7 percent of type A and 8.2 percent of type B). I use the standardized high school score and the binary indicators for home schooled and repeaters as measures of pre-collegiate ability.

In each high school, cohorts were divided randomly into classes of 20-30 students at the beginning of freshman year. Usually, the initial assignment remained unchanged in the following years, except for students who moved to new classes after failing a grade. Students in the same class interacted on a daily basis for five years: they attended the same lectures, had the same homework, and studied for the same tests. For each student, I compute his other classmates' average score and I use it as a measure of pre-collegiate peer effects.

[15] Due to the relevance of Milan in the Italian economy, the dataset used in this study is comparable to the individual-level data of New York City students (NYC Basic Educational Data System) that has often been used in the economics of education literature.

II.B 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).

The transcripts are an incredibly rich source of information: they contain the title of each course attended, the exam date and the grade received. Grades (from 0 to 30 cum laude with 18 as passing grade) are assigned using an absolute scale and are correlated to labor market outcomes. The degree, in fact, is awarded with a final mark (*voto di laurea*; a function of the GPA and the quality of a final thesis) that is very salient to potential employers during a job search. Based on these considerations, I use university grades to model human capital as a function of knowledge acquired in single courses. The transcripts contain also the start and end date of each student's university career, together with a description of the final outcome (graduation, dropout, or transfer). 86.7 percent of type A students enrolled in university and 63.7 percent graduated (Table 1). In comparison, only 40.8 percent of type B students enrolled and 16.1 percent graduated.

The fact that I have transcripts only from the universities located in Milan could raise some concerns about the representativity of the sample. If, for example, a large share of high school students chose to attend university in another city or country, limiting the data collection to local universities would return a selected group of students. High school graduates in Milan, however, did not travel out of the area to receive a university education. In 1967, for instance, 93.5 percent of university freshmen who had attended high school in Milan enrolled in a local university (ISTAT).[16]

II.C Income Tax Returns in 2005

I use taxable income from personal income tax returns in 2005 as a long-run labor market outcome. This measure is the sum of all individual earnings that are taxable under the Italian personal income tax after allowed deductions. It includes labor earnings for employees, profits for the self-employed, pensions, rents, and interests. The main excluded categories

[16] 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 limited admission and high tuition fees.

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.[17]

Using name and birthdate, I uniquely match 83 percent of the high school graduates to one income earner in 2005. The matching rate is balanced by gender (81.8 percent for women and 83.3 for men) because couples file separately and married women are required to use their maiden name. The difference between the survival and matching rate is slightly increasing in time.[18] The matching rate of the 1958 cohort is 81 percent, 6 percentage point lower than the survival rate. At the other extreme, the matching rate of the 1968 cohort is 85.5 percent, 9.5 lower than the survival rate. This suggests that some individuals who were alive in 2005 were not included in the list of income earners, because they either had a disposable income below the taxable threshold (€7,500) or were living and working abroad. The comparison of pre-collegiate characteristics of students with and without an income observation suggests that matched individuals were on average 0.8 years younger (Appendix Table B.1). This difference is significant, although small, and is consistent with the fact the mortality rate is higher for older cohorts. Moreover, matched individuals were more likely to be men and to have higher high school grades. These findings suggest that attrition might be primarily driven by individuals (women and low-achieving students) whose incomes were below the minimum threshold, and not by high-skilled expatriates. Importantly, selective attrition does not change across cohorts (Appendix Table B.2).

In 2005, 95 percent of the students in the dataset were between 56 and 67 years old. Therefore, I have one observation at - or just after - the peak of the earning potential, when the returns of education were realized in full. In addition, this measure is less likely to be affected by temporary shocks, as is more common with earnings measured at a younger age. However, the fact that I have only one cross section limits the analysis in two dimensions. First, I cannot examine the effect of the policy on income dynamics. Second, in a single cross-section, any comparison between students from different cohorts combines age and cohort effects: cohort plus age equals year. 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). In my particular application, one would be concerned that older individuals were more likely to be retired and that younger individuals could still be

[17]In March 2008, the Italian Ministry of the Treasury published online the 2005 income tax data (with identifiers like name, birthdate, and city of residence). The original goal was to fight tax evasion by allowing every citizen to check the income reported by acquaintances, coworkers, and neighbors. The Italian public strongly opposed this way of disseminating income data and, therefore, the data files remained online for less than 24 hours. The researchers that downloaded the data in March 2008 can now use the income observations for research purposes.

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

on a rising part of their income trajectory. It is not obvious what bias would result. On the one hand, retirees in Italy tend to have pay that is lower than, though closely related to, their pre-retirement pay. On the other hand, Italian workers who are still short of retirement likely have pay that is lower than that they will receive in their final work year. What is certainly true, however, is that Italian workers in the 56 to 67 age range face pay and pension systems that are rigid and display little year-to-year fluctuation.[19]

To measure the effect of the policy, then, I need to eliminate age effects and retain the cohort effects. To deal with this problem, I use a dataset in which separating age and cohort effects is a reasonable exercise. Specifically, I use repeated cross-sections of the SHIW data. In this dataset, I estimate age effects as a function of observable characteristics, allowing for cohort and year effects. Of course, owing to the fact that cohort plus age equals year, I cannot allow fully free year effects. I therefore assume that year effects are smooth. This assumption is very reasonable because I am not using data from a period that contains sharp year events (like a war). The results, however, are robust to different specifications of the age, cohort, and year effects.

My out-of-sample estimates of age effects serve two purposes. First, I use them to predict income at age 65 (the age of retirement for men) for all my in-sample observations.[20] Then, throughout all of my analysis, I use this age-adjusted income to measure the cohort effects of the policy. Second, I use the out-of-sample estimates to gauge the importance of age effects among Italian individuals aged 56 to 67. If I find that the age effects are trivial in the SHIW dataset, they should also be trivial in my cross-section based on income tax returns.

In fact, I do find that age effects for older Italian workers are unimportant. This is both because their pre-retirement earnings are hardly rising with age and because their post-retirement 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 is useful. Indeed, the age effects are so trivial that I find that using age-adjusted income leads to results that are similar to results based on raw income (Appendix Table C.4).[21]

---

[19] Italy displays high real and nominal wage rigidity, even relative to other European countries with rigid labor markets (Grubb, Jackman and Layard 1983, Devicienti, Maida and Sestito 2007, Holden and Wulfsberg 2008, Fabiani and Sabbatini 2011). This characteristic could be due to high unionization rates and string employment protection laws.

[20] Appendix C describes this procedure in greater detail.

[21] To further show that the age effects are small, I use the panel component (a smaller sample) of the SHIW dataset. The advantage of the longitudinal data is that the cohort is literally the same people year after year. This makes their cohort effect constant by definition (apart from attrition). However, the sample size is too small to precisely estimate the age effects. With this caveat in mind, I use the small longitudinal component of the SHIW dataset to show that the post-retirement earnings are highly correlated with the pre-retirement earnings. This result is another piece of evidence about the irrelevance of age effects for older Italian workers.

On average, type A students earned €58,657 in 2005, which increases to €65,749 after adjusting for age effects (Table 1). Type B students, instead, earned €47,628 in 2005 or €53,812 after adjusting for age effects

III Education and Income of Type B Students

In this section, I study the direct effects of the reform on the students who were allowed to enroll in STEM. I find that type B students were more likely to enroll in university and receive a university degree after 1961. The reform, then, was very successful in expanding access to university among high school graduates who would have not not enrolled otherwise. However, type B students did not earn higher income in the long-run, relative to earlier cohorts who were denied access to STEM majors.

III.A Inter-Cohort Changes in Education and Income

In the baseline specification, I compare outcomes of adjacent cohorts of type B students. I estimate the following regression:

$$\text{outcome}_{it} = F(\alpha + \sum_t \beta_t Y_t + \gamma X_{it}) \tag{1}$$

where outcome_{it} is either university enrollment, university graduation, or log income in 2005 of student i in cohort t. $F()$ is the logistic function when the dependent variable is binary and a linear function otherwise. Y_t is either a full set of cohort fixed effects or two binary variables for phase 1 (Post 61_t for 1961-to-1964 cohorts) and 2 (Post 65_t for 1965-to-1968 cohorts) of the policy. 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 effect on cohort t is $\text{E}[\text{outcome}_{it} \,|\, Y_t = 1\,,\, Y_j = 0\,,\, X_{it} = x] - \text{E}[\text{outcome}_{it} \,|\, Y_t = 0\,,\, Y_j = 0\,,\, X_{it} = x]$, with $j \neq t$.

Enrollment rates of type B students increased by 9 percentage points from 1961 to 1964 and by 22.5 percentage points from 1965 to 1968, relative to the baseline cohorts (Table 2, Columns 1 and 2). The reform, then, did not merely move type B students across majors (from non-restricted to STEM), but induced more students to enroll in university. Estimates for single cohorts show that the effect becomes statistically significant after 1961, increases steadily until 1966, and then stabilizes at this permanently higher level (Figure 2, Panel A).

Higher university enrollment translated into more completed education. 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 (Table 3, Columns 3 and 4). Compared with an average graduation rate of 8.2 percent in the pre-reform cohorts, these estimates imply a 72 percent and 183 percent increase. The estimates for single cohort show two positive discontinuities in 1962 and 1965, which correspond to the two phases of the policy (Figure 2, Panel B).

Type B students who completed high school from 1961 to 1964 earned 9.5 percent more in 2005, relative to earlier cohorts (Table 3, Column 5). However, the following cohorts, for whom enrollment and graduation rates increased the most, did not earn higher income, relative to the baseline (coefficient 0.015, p-value 0.700). Similarly, the estimates of single cohort effects are not statistically significant and close to zero after 1965 (Figure 2, Panel C).

In the next step, I measure more directly the effect of the policy on the incomes of type B students who enrolled in university after 1961. The starting point is a standard Mincerian equation that relates log income to a binary variable for completed university education:

$$\log(\text{income}_{it}) = \beta_0 + \beta_1 \text{degree}_{it} + \beta_2 X_{it} + \eta_{it} \tag{2}$$

where degree_{it} is equal to one for university graduates and X_{it} is the usual set of individual characteristics. I instrument degree_{it} in (2) with two binary variables: Post 61_t which is 1 for the 1961-to-1964 cohorts and Post 65_t which is 1 for the 1965-to-1968 cohorts. These variables measure the increase in completed education among type B students after the policy implementation.

The OLS estimator of β_1 is equal to 0.337 and statistically significant (Appendix Table A.2). If we assume that each year of university contributes equally, the OLS estimator implies that one additional year of university increased long-run earnings by 8 percent in case of 5-year degrees (or by 10 percent in case of 4-year degrees).

Next, I instrument university graduation (degree_{it}) with the two binary variables (Post 61_t and Post 65_t) that identify the cohorts of type B students who were affected by the policy. Earlier in this section, I described the results of the first stage regression, which captures the increase in university graduation generated by the reform: 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 the pre-reform cohorts (Table 3, Columns 3 and 4). The IV estimator of β_1 in equation (2) measures the effect of the policy on the incomes of type B students who received a university degree after 1961. The IV estimator is lower than its OLS counterpart and not statistically significant (-0.225, p-value 0.343), but is imprecisely estimated. In spite of a large confidence interval, I can rule out returns to university education above 5.4 percent per academic year.[22] If I include controls for the region of residence in

[22] These upper-bound estimates are much lower than what found by Maurin and McNally (2008). They use a temporary increase of high school graduation rates in France to estimate the returns to university

2005, I can rule out returns to education above 4.5 percent per academic year.

To confirm the fact that type B students did not earn positive returns to a STEM education after 1965, I estimate equation (2) using only Post 65_t as instrument. The IV estimator of β_1, then, measures the effect of the policy on the incomes of type B students who enrolled in university from 1965 to 1968. The estimate is negative and statistically significant at the 10 percent level. Using the confidence interval, I can rule out returns to education above 1.9 percent per academic year.

The policy did not increase the average income of type B students, but could have modified the income dispersion. However, I find that the policy left the whole income distribution unchanged. For instance, I can compare the income distribution of the 1958 cohort to the income distribution of the 1968 cohorts, the cohort in my sample for whom university education increased the most (Figure 2, Panel D). Based on the results of a Kolmogorov-Smirnov test, I fail to reject the hypothesis that the two distribution are statistically equivalent at the 5 percent level (p-value 0.064). Moreover, a visual inspection reveals that the share of type B students on the two tails of the income distribution did not increase significantly in the 1968 cohort, relative to the 1958 cohort. For example, 5.2 percent of type B students in the 1958 cohort and 3.7 percent in the 1968 cohort earned more than €148,747. On the opposite tail of the distribution, 5.7 percent of type B students in the 1958 cohort and 6 percent in the 1968 cohort earned less than €2,981.

These findings do not change if I estimate equation (2) on different quartiles of the ability distribution. The OLS estimator is negative and significant for the bottom quartile and, for these students, I can rule out positive returns to a STEM education after 1961. For the remaining quartiles, the IV estimators are not statistically different from zero, but are less precisely estimated.[23]

III.B Robustness Checks and Additional Results

The identifying assumption for the estimation of β_t in equation (1) is that adjacent cohorts differ only in the exposure to the reform. One major threat is changing selection into type B schools over time. In section I.C, I addressed this concern using aggregate high school enrollment and out-of-sample survey data. In addition, the previous estimates showed that

for cohorts born between 1947 and 1950. They find that one year of university increased earnings by 14 percent.

[23] Results are robust to using separate cohort dummies as instruments. In addition, I can restrict the sample to type B in the bottom and top quartile of the ability distribution and use as an instrument the interaction of Post 61_t and Post 65_t with an indicator for being in the top quartile. Alternatively, I replace degree$_{it}$ in (2) with STEM degree$_{it}$, which is 1 if a type B students received a university STEM degree. The estimates do not change.

enrollment and graduation of type B students increased before 1966: selection cannot bias these estimates because these cohorts enrolled in high school before the policy implementation. As an additional test, I estimate equation (1) without any control for pre-collegiate ability.[24] The cohort effects do not change significantly. University enrollment increases from 1962 to 1965 and, then, stabilizes (Appendix Figure A.3, Panel A). University graduation shows two jumps in 1962 and 1965 (Appendix Figure A.3, Panel B). Despite a substantial increase in completed education, cohorts who completed high school after 1965 did not earn higher incomes in 2005 (Appendix Figure A.3, Panel C).

Even if selection into high schools stayed constant, younger cohorts could be different for a variety of unrelated reasons. Instead of relying exclusively on inter-cohort comparisons, I can adopt a difference-in-differences approach:

$$\text{outcome}_{it} = \alpha + \beta \text{B}_i + \sum_t \gamma_t Y_t + \sum_t \delta_t [\text{B}_i \times Y_t] + \zeta X_{it} + \eta_{it} \tag{3}$$

where B_i is a binary variable that identifies students who are substantially more "treated" within each cohort. There are multiple ways to identify more and less treated individuals within a cohort.

First, I compare type B students who scored in the top quartile of their high school class with type B students who scored in the bottom quartile ($B_i = 1$ for high-achieving type B). In the pre-reform cohorts, high-achieving type B students were 15.1 percent more likely to enroll in university, 9.3 percent more likely to receive a university degree, and earned 33.8 percent more in 2005. The education gap widened significantly after 1961. The gap in university enrollment (in favor of high-achieving students) increased by 7.1 percentage points from 1961 to 1964 and by 6 percentage points from 1965 to 1968 (Table 3). Similarly, the gap in university graduation increased by 7.9 percentage points from 1961 to 1964 and by 9.7 percentage points from 1965 to 1968. The income gap, however, did not change after 1961. These results suggest that within each post-reform cohort the type B students with higher aptitude completed more education relative to the type B students with lower aptitude, but did not earn more in the long-run. This is an important finding because it does not rely on the strong assumption that adjacent cohorts differ only in their exposure to the reform (it compares changes in education and income across time and within cohorts), but confirms that the policy did not increase the incomes of type B students (as simpler cross-

[24] I drop the high school exit score, average score of high school classmates, a dummy for home schooled, and a dummy for non-repeaters. If selection into type B school changes across cohorts and depends on pre-collegiate ability, dropping these variables would change the cohort effects.

cohort comparisons showed).[25] Moreover, the difference-in-differences setting should capture remaining age effects (all the effects invariant to education), if the procedure described in section II.C and appendix C failed to eliminate all of them.

In addition, I can 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 pre-existing income gap, however, remained unchanged. After the policy implementation, then, type B students invested more in education relative to type A students, but did not earn higher incomes in the long-run. This finding is another piece of evidence, which does not simply rely on inter-cohort comparisons, suggesting that the policy did not increase the earnings of type B students.

Furthermore, I can compare type B and C students, the graduates from commercial schools. These two groups of students had access to the same university programs up until the policy implementation (mainly business economics and statistics). Then, the reform allowed type B students to enroll in university STEM majors, but did not have any direct effect on type C students. This comparison leads to similar conclusions. In the pre-reform cohorts, type C students were 27.2 percent more likely to enroll in university, 4.8 percent more likely to graduate, and earned 22.5 percent more in 2005. The enrollment gap fell by 22 percentage points from 1965 to 1968. The graduation gap decreased by 12.5 percentage points from 1965 to 1968. Nevertheless, the income gap remained unchanged. Although surprising, these results are consistent with the previous finding of no income effect on type B students. After the policy implementation, type B students completed much more education (relative to type C students) to the extent that the pre-reform gap in university enrollment had reversed by 1968. However, they did not earn higher incomes in the long-run.[26]

[25] Alternatively, I use cohorts who completed high school after 1961 to predict the propensity to enroll in STEM majors as a function of observable characteristics. I adopt a leave-one-out estimator in order to avoid over-fitting bias (Abadie, Chingos and West, 2014). Then, I can compare the type B students who have a high propensity to enroll in STEM (top third) with the type B who have a low propensity (bottom third). "High-propensity" type B students completed more education after 1961, but did not earn more.

[26] If I restrict the sample to male students, the main findings hold. In only one case (male type B vs male type C students), I find that the income gap fell (in favor of type B students), but this result is driven exclusively by a marked income decrease among type C students.

III.C Explanations for Zero Returns to a STEM Education

After 1961, type B students were more likely to enroll in university and to receive a university degree. However, the returns to a STEM education were either non-existent or small and concentrated among the cohorts who graduated immediately after 1961. Within a partial equilibrium framework, there could be different non-mutually exclusive explanations for these findings.

It could be the case that the Italian economy did not value STEM skills. In the pre-reform cohorts, however, STEM students earned 65.2 percent more in 2005, relative to high school graduates. If the labor market did not value STEM skills, this large income premium would depend exclusively on ability sorting. This is contradicted by previous research: using US data, Arcidiacono (2004) shows that the STEM premium cannot be fully explained by selection into different majors.

Alternatively, the lack of returns to a university education could depend on the type of high school attended. For example, type A students could find high-income jobs through personal connections, grown spending 5 years of high school with other type A students. Similarly, a type B diploma could bear a negative stigma in the labor market that would persist after acquiring a university degree. Indeed, a type A diploma is associated with 36.1 percent higher income, compared with a type B diploma (p-value < 0.001). However, a type A diploma loses its predictive power after controlling for university graduation. The coefficient becomes small and not statistically significant (0.028, p-value 0.457).

In addition, older cohorts started their work careers in a better economic environment. While the 1958 cohort entered the labor market in a period of economic growth, the 1968 cohort completed university around the 1973 oil crisis. The literature that examines the costs of graduating in a recession found that the negative effects fade out after few years from graduation.[27] In addition, graduates from high-return majors do not seem to suffer at all from entering the labor market in a period of high unemployment (Altonji, Kahn and Speer, 2013). Based on the existing evidence, it is unlikely that graduating in a recession could generate negative income effects large enough to wipe out returns to a STEM education 30 years after graduation.

In conclusion, none of these hypotheses is particularly convincing in explaining the initial findings. In the next section, I will propose a model in which educational expansion can lower returns through various channels.

[27] Oreopoulos, von Wachter and Heisz (2012) use Canadian data and find that the effect on earnings disappears after 10 years from graduation. In the US context, Altonji, Kahn and Speer (2013) find that on average the negative effect on earnings fades out after only 7 years.

IV General Equilibrium Model

In the model described in this section, educational expansion can affect returns to education in three ways. First, higher enrollment can congest university resources and lead to a decrease in the quality of education. Second, the admission of less well prepared students might increase the degree of class heterogeneity, generating negative peer effects and making teaching more difficult. Both of these effects contribute to weaken human capital accumulation. Lastly, higher enrollment leads to a higher supply of human capital in the labor market, which in turn decreases the price of STEM skills.

IV.A Model

At the end of high school, students choose between a university major or work (HS). The majors are divided in 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 non-restricted majors). Aggregate production in the economy is determined by a CES production function that uses the four types of human capital:

$$Y = (S_{HS}H_{HS}^{\rho} + S_{STEM}H_{STEM}^{\rho} + S_R H_R^{\rho} + S_{NR}H_{NR}^{\rho})^{\frac{1}{\rho}} \tag{4}$$

where H_k is the aggregate supply of human capital with education k, S_k 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 unit price of each skill equals its marginal productivity in equilibrium. This implies that the skill price of human capital $k = \{HS,\ STEM,\ R,\ NR\}$ is:

$$w_k = Y^{1-\rho} S_k H_k^{\rho-1} \tag{5}$$

Skill prices depend on the aggregate supply of different skills in the economy. An increase in the supply of STEM skills, relative to other skills, might lead to a decrease in its price. The wage of individual i with education k is the product of skill prices, which are the same for all workers with the same education, and human capital, which varies across individuals ($W_i^k = w_k \cdot h_i^k$). Individual human capital is a function of the knowledge acquired in each university course:

$$h_i^k = \sum_{c \in N_k} \mu_c \cdot k_{ic}^k \tag{6}$$

where μ_c are normalized weights and N_k is the set of courses in major k. Knowledge in course c is a function of individual and course characteristics:

$$k_{ic}^k = \gamma_0^k + \gamma_1^k X_i + \gamma_2^k C_{ic} + \gamma_3^k Q_c + \gamma_4^k \mathrm{CH}_{ic} + u_{ic}^k \tag{7}$$

where X_i are individual characteristics and C_c course characteristics, Q_c is quality of education in course c, and CH_{ic} measures class heterogeneity. The errors u_{ic}^k are i.i.d. random shocks normally distributed with mean 0 and standard deviation σ. The utility of choosing a major k is the sum of net non-monetary preferences and the log of monetary returns:

$$u_i^k = \alpha^k X_i + \log(W_i^k) + \varepsilon_i^k \tag{8}$$

where W_i^k is wage with a degree in major k and ε_i^k is a random idiosyncratic shock. The parameters α^k depend on individual characteristics and measure net preferences with respect to the work option. The random shocks (ε_{it}^k) follow a type I extreme value distribution and, therefore, the choice probabilities take the form of a multinomial logit model (Train, 2009).[28] An equilibrium is defined as a collection of individual schooling decisions $\{d_i^k\}$, skill prices $\{w_k\}$, aggregate human capital $\{H_k\}$ such that (1) the schooling decisions solve the individual maximization problem, (2) skill prices equate the marginal products of human capital, (3) and the aggregate demand of human capital equates the aggregate supply.

IV.B General Equilibrium Effects

In the remaining part of this section, I describe how returns to education respond to the entry of type B students in STEM majors.

IV.B.1 Congestion of Public Resources

Knowledge acquired in course c is a function of 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). Resources do not vary with enrollment. This fits the Italian scenario, where tertiary education was heavily subsidized by lump-sum transfers and tuition

[28] The main limitation of the multinomial logit model (ML) is the assumption of independence from irrelevant alternatives (IIA). As a robustness check, I allow the random shocks to follow a generalized extreme value distribution. The resulting nested logit (NL) has two nests: the three schooling options in one nest and the work option in the second. In this case, the ML is just a NL with an additional parameter constraint (the dissimilarity parameter of the schooling nest, λ_S, equal to 1). I run a LR test and I fail to reject the ML ($\lambda_S = 0.925$, p-value 0.573).

fees were low.[29] However, this assumption could be relaxed to allow for cases in which university resources respond to enrollment, although not elastically (for example, a context in which tuition fees represent a large source, but not the only one, of revenues).[30] Following a marginal increase in type B enrollment, quality of education varies according to:

$$ge_{Q_c} = \frac{\mathrm{d}Q_c}{\mathrm{d}E_{B,STEM}} = \frac{\partial Q_c}{\partial E_{STEM}} \cdot (1 + \frac{\partial E_{A,STEM}}{\partial E_{B,STEM}}) \tag{9}$$

The total effect is a function of (1) the crowding of public resources ($\frac{\partial Q_c}{\partial E_{STEM}}$), and (2) the response of type A students to type B entry into STEM majors ($\frac{\partial E_{A,STEM}}{\partial E_{B,STEM}}$). The element $\frac{\partial Q_c}{\partial E_{STEM}}$ describes the fact that, as total enrollment increases, resources are shared among a larger amount of students, access to university inputs become more difficult, and quality of education decreases. The sign of equation (9) hinges on how total enrollment changes. On one hand, the entry of one type B student increases STEM enrollment by one unit. On the other hand, type A students might decide to move out of STEM fields as type B move in. This second effect is due to the fact that type A students can enroll in a group of majors with restricted access after 1961, where quality of education is not affected by the entry of type B. The overall effect is negative only if the first effect prevails and total enrollment in STEM majors increases ($\frac{\partial E_{A,STEM}}{\partial E_{B,STEM}} > -1$).

IV.B.2 Increase of Class Heterogeneity

In addition to quality of education, knowledge acquired in course c is a function of class heterogeneity. In this particular context, class heterogeneity depends primarily on the skill set that university students acquired during high school. If students have a homogeneous set of pre-collegiate skills, teaching can be easier and negative peer effects less likely (for example, repeated questions from students lagging behind or disruptive behavior of disengaged students).

Educational expansion can modify the composition of skills in a classroom. 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 of the number of students of a different type ($E_{-i,STEM}$). Following a marginal increase in type B enrollment, class heterogeneity varies according to:

[29] In Milan, tuition fees were equal to 20 percent *(Università Statale)* and 11 percent *(Politecnico)* of total revenues.

[30] If higher enrollment generates congestion costs, then quality of education could be affected even in a context in which resources adjust elastically to the number of enrolled students.

$$ge_{\mathrm{CH}_{ic}} = \frac{\mathrm{dCH}_{ic}}{\mathrm{d}E_{B,STEM}} = \frac{\partial CH_{i,c}}{\partial E_{B,STEM}} + \frac{\partial CH_{i,c}}{\partial E_{A,STEM}} \cdot \frac{\partial E_{A,STEM}}{\partial E_{B,STEM}} \tag{10}$$

The total effect is a function of (1) how class heterogeneity varies with a change in type A ($\frac{\partial CH_{i,c}}{\partial E_{A,STEM}}$) and type B enrollment ($\frac{\partial CH_{i,c}}{\partial E_{B,STEM}}$), and (2) how many type A students move out as one type B moves in ($\frac{\partial E_{A,STEM}}{\partial E_{B,STEM}}$). For type A students in STEM majors, class heterogeneity increases after the reform. In fact, type B students enter and make classes more diverse ($\frac{\partial CH_{i,c}}{\partial E_{B,STEM}}$). At the same time, some type A students who would have chosen STEM might decide to enroll elsewhere to avoid the negative consequences of educational expansion ($\frac{\partial CH_{i,c}}{\partial E_{A,STEM}} \cdot \frac{\partial E_{A,STEM}}{\partial E_{B,STEM}}$).

IV.B.3 Change of Skill Prices

An increase in STEM enrollment drives up the aggregate supply of STEM skills in the economy. Under the assumption of imperfect substitutability between different types of human capital, this decreases prices for STEM skills. From equation (5), the change in prices of STEM skills following a marginal increase of type B enrollment is:

$$ge_{w_{STEM}} = \frac{\mathrm{d}log(w_{STEM})}{\mathrm{d}E_{B,STEM}} = -\phi^{-1} \cdot \frac{1}{H_{STEM}} \cdot (\frac{\partial H_{STEM}}{\partial E_{B,STEM}} + \frac{\partial H_{STEM}}{\partial E_{A,STEM}} \cdot \frac{\partial E_{A,STEM}}{\partial E_{B,STEM}}) \tag{11}$$

with aggregate production Y fixed. The total effect depends on (1) the elasticity of substitution (ϕ), (2) how the supply of human capital changes with type A ($\frac{\partial H_{STEM}}{\partial E_{A,STEM}}$) and type B ($\frac{\partial H_{STEM}}{\partial E_{B,STEM}}$) enrollment, and (3) how many type A students move out as type B move into STEM ($\frac{\partial E_{A,STEM}}{\partial E_{B,STEM}}$). As seen above, the sign depends on two offsetting changes. On 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 ($\frac{\partial H_{STEM}}{\partial E_{A,STEM}} \cdot \frac{\partial E_{A,STEM}}{\partial E_{B,STEM}}$). The change in skill prices is negative if the entry of type B students is not compensated by a decrease in type A enrollment ($\frac{\partial H_{STEM}}{\partial E_{A,STEM}} \cdot \frac{\partial E_{A,STEM}}{\partial E_{B,STEM}} > -\frac{\partial H_{STEM}}{\partial E_{B,STEM}}$).

V The Effects of Educational Expansion

In this section, I show different effects of educational expansion. I find that the type A students with higher pre-university achievement abandoned STEM after 1961 in favor of restricted majors. Then, I show that human capital, measured by the grades received in university courses, decreased due to lower quality of education and higher class heterogeneity.

Finally, I show that income of STEM students with a type A diploma decreased after 1961 to the extent of almost erasing the pre-existing STEM premium.

V.A Major Choice of Type A Students

Following a marginal increase in type B enrollment, the probability of type A students choosing STEM varies according to:

$$\frac{\partial P_{iA}^{STEM}}{\partial E_{B,STEM}} = P_{iA}^{STEM}(1 - P_{iA}^{STEM})(GE_i^{STEM} - GE_i^R) \tag{12}$$

where GE_i^k is the marginal decrease in the returns to major k due to educational expansion and P_{iA}^k is the ex-ante probability of type A choosing major k.[31] For ease of exposition only, I assumed that type A students choose only between STEM and restricted majors.

I will test two prediction stemming from equation (12). First, if the returns decrease more in STEM ($\left|GE_i^{STEM}\right| > \left|GE_i^R\right|$), some type A students move from STEM to restricted majors.[32] Second, the change in probability is larger among the students who are ex-ante uncertain between STEM and restricted majors.[33]

In Milan, the share of type A students enrolling in STEM and restricted majors follow a diverging trend after the reform (Figure 3, Panel A). In 1958, 39.6 percent of type A enrolled in STEM, while 37.7 percent enrolled in restricted majors. The two shares stay constant until 1961. After 1961, however, the portion of type A students enrolling in STEM starts decreasing and displays two negative discontinuities in 1962 (-3.1 percentage points from 1961) and 1965 (-2.7 percentage points from 1964). The share of type A students choosing a restricted major, instead, follows the opposite path. Enrollment in restricted majors increased to 44.9 percent in 1962 and 53.5 percent in 1965.

To control for individual characteristics, I estimate a multinomial logit model:

$$\log(\frac{Pr(\text{major}_{it} = k)}{Pr(\text{no college})}) = \alpha_k + \sum_t \beta_{kt} Y_t + \gamma_k X_{it} \tag{13}$$

[31] GE_i^k is the sum of the three effects described in the previous section: decrease of quality of education, increase of class heterogeneity, and decrease of skill prices. It can be re-written as:

$$GE_i^k = ge_{w_k} + \frac{\sum_{c \in N_k} \mu_c(\gamma_3^k \cdot ge_{Q_c} + \gamma_4^k \cdot ge_{\text{CH}_{ic}})}{h_i^k}$$

[32] Type A students do not respond if either the general equilibrium effects are not existent ($GE_i^k = 0$), or the two fields are affected equally ($GE_i^{STEM} = GE_i^R$). The first scenario corresponds to a partial equilibrium model.

[33] For given values of GE_i^k, the change in probability is maximum at $P_{iA}^{STEM} = 0.5$.

where the choice is between STEM, restricted, non-restricted majors and no university as baseline. Y_t is either a full set of cohort fixed effects or two binary variables for phase 1 (Post 61_t for 1961-to-1964 cohorts) and 2 (Post 65_t for 1965-to-1968 cohorts) of the policy. X_{it} includes the usual student controls. The effect for cohort t is $\Pr[\text{major}_{it} = k \,|\, Y_t = 1\,,\, Y_j = 0\,,\, X_{it} = x] - \Pr[\text{major}_{it} = k \,|\, Y_t = 0\,,\, Y_j = 0\,,\, X_{it} = x]$, where $j \neq t$.

Type A students who completed high school between 1961 and 1964 were 0.9 percentage points less likely to choose a STEM major and 5.9 percentage points more likely to choose a restricted major, compared with the baseline cohorts (Table 4). The change in the probability of enrolling in STEM is not statistically significant (p-value 0.468). Type A students who completed high school between 1965 and 1968 were 8.5 percentage points less likely to choose STEM and 17.2 percentage points more likely to choose a restricted major. Both effects are statistically significant. These estimates indicate a 22.2 percent decrease of enrollment in STEM and a 44.4 percent increase of enrollment in restricted majors, relative to the baseline probabilities.

The paths of cohort effects follow the same diverging trends shown by the raw probabilities (Figure 3, Panel B). Cohort effects are not statistically significant until 1961 and do not show any pre-existing trend. In 1963, however, type A students were 3.7 percentage points less likely to enroll in STEM and 8.8 percentage points more likely to enroll in a restricted major, compared with the 1958 cohort. The cohort effects keep diverging until 1966 and, then, reach a plateau. In 1968, type A students were 9.6 percentage points less likely to choose STEM and 17.8 percentage points more likely to enroll in restricted majors, relative to the 1958 cohort.

V.A.1 Movers and Stayers

I estimate equation (13) using data on the pre-reform cohorts. I use these coefficients to predict major choices among the cohorts who completed high school after 1961 (post-cohorts). I, then, compare predicted and actual shares of type A students in different majors. I find that the graduates from the humanities schools and the students with the higher high school grades were more likely to abandon STEM majors after 1961 (Appendix Table A.3). These results suggest that educational expansion affected disproportionately the students with stronger preferences for STEM disciplines and with lower pre-collegiate achievement.

First, I find that 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*). In fact, 19.2 percent of type A humanities students enrolled in STEM after 1961, 10.2 percentage points less than predicted. In comparison, 50.1 percent of type A scientific students chose STEM after 1961, only 3.1 percentage points less than predicted. This finding holds at any level of pre-collegiate ability. Since scientific students had both a stronger interest (by revealed preference) and better preparation in STEM fields, they were less inclined to abandon STEM majors when the returns started declining.

Second, I find that the type A students who scored in the top quartile of their high school class were more likely to move out of STEM majors after 1961, compared with the lower achieving students. This finding holds among the graduates from both the humanities and scientific schools. For example, 43.5 percent of the humanities students who scored in the top quartile of their high school class enrolled in STEM after 1961, 16.3 percentage points less than predicted. By comparison, 21.8 percent of the humanities students who scored in the bottom quartile enrolled in STEM after 1961, only 6.9 percentage points less than predicted. The high-achieving students were more well prepared in different disciplines and, therefore, were more flexible in their major choice. When the returns to a STEM education started decreasing, the high-achieving students were the first to move elsewhere.

To investigate where type A students moved after 1961, I estimate equation (13) splitting the restricted majors in different categories (Appendix Table A.4). I find that the humanities students moved in similar proportions towards medicine and humanities majors, while the scientific students moved exclusively to medicine.

In medicine, enrollment grew monotonically with pre-collegiate ability. 13.1 percent of the humanities students who scored in the top quartile of their class enrolled in medicine after 1961, 11.7 percentage points above predictions. In comparison, 13.8 percent of the humanities students who scored in the bottom quartile of their class enrolled in medicine, only 6.1 percentage points above predictions. I find the same pattern among scientific students. In the humanities majors, instead, enrollment increased more among the students with lower pre-collegiate achievement. For example, 37.2 percent of the humanities students who scored in the top quartile of their class enrolled in a humanities major, only 5.4 percentage points above predictions. In comparison, 30 percent of the humanities students who scored in the bottom quartile of their class enrolled in a humanities major, 6.8 percentage points above predictions.

V.A.2 Robustness checks

The identifying assumption to estimate cohort effects in equation (13) is that adjacent cohorts differ only in relation to their exposure to the reform. I already ruled out the hypothesis that the reform changed the choice of high school in section I.C. The remaining 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.

As done in section III.B, I estimate the multinomial logit model in equation (13) dropping controls for pre-collegiate ability. The cohort effects do not change (Appendix Figure A.4, 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 (13) using data on male students only. The main findings hold (Table 4 and Appendix Figure A.4, 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 (13) 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 A.5, Panel A), while the predicted share in restricted majors is stable (Appendix Figure A.5, Panel B).

The second group of robustness checks tests for the role of concurrent and exogenous changes in the returns to higher education.[34] 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).[35] Controlling for contemporaneous changes in the economy does not affect the path of the marginal cohort effects (Appendix Figure A.4, 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 (13) using more disaggregated choices to show that enrollment in fairly different STEM (restricted) majors follow the same decreasing (increasing) trend after 1961 (Appendix Figure A.4, Panel D).[36] In 1968, type A students were 2.1 percentage points less likely to enroll in engineering and 5.1 percentage points in physics,

[34] 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.

[35] 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.

[36] 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.

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.

V.B Quality of Education and Congestion of Public Resources

In a setting in which university resources are fixed (or do not fully adjust to enrollment changes), educational expansion can congest access to university inputs. Consider a practical example in which a course has one professor who delivers the main lectures and one assistant who holds weekly office hours. If enrollment doubles and the number of assistants stays fixed, the office hours might not be enough to accommodate a larger class and, therefore, learning becomes more difficult. In this example, educational expansion causes a higher number of students to compete for a fixed university input (office hours), which leads to lower quality of education and, in turn, lower human capital. If courses differ in the amount of university resources assigned to them, the congestion that follows an enrollment increase should be more severe in courses with fewer resources. Continuing the previous example, consider a second course in the same major with equal enrollment and 10 assistants. Compared to the course with 1 assistant, this course is better equipped to handle an enrollment increase and, therefore, educational expansion prevents fewer students from attending office hours. Then, even though enrollment increases equally across all courses in the same major, quality of education decreases more in courses with fewer resources.

In the empirical analysis, I assume that quality of education in course c is a function of the student-faculty ratio, $Q_c = f\left(\frac{E_c}{fac_c}\right)$, where E_c is the number of students enrolled in c and fac_c is the number of teaching fellows (professors and teaching assistants) assigned to the course. Unlike more direct metrics of university spending, the student-faculty ratio is observed at the course level. However, as long as quality of education depends on other uncorrelated factors, the student-faculty ratio might under-measure the effect of educational expansion on academic outcomes.[37]

The function f is decreasing ($f^{'} < 0$) and either linear or concave ($f^{''} \leq 0$). As discussed above, this parametrization implies that the decrease in quality following a marginal increase in enrollment is larger among courses with a higher pre-existing student-faculty ratio.[38] I test

Similarly, I divide STEM in: (1) physics, (2) mathematics, (3) sciences - geology, biology, natural science, and chemistry, (4) engineering.

[37] Crowding could have happened on many dimensions other than the student-faculty ratio. For example, a few universities (*Università Statale di Milano* included) were forced to move the courses with higher enrollment into movie theaters, because their infrastructures were undersized. Since these venues were not ideal classrooms, human capital accumulation in the displaced courses probably decreased. I, however, do not examine this type of crowding in the empirical analysis.

[38] The idea of congestion costs in the education production function has been formalized by Lazear (2001, p. 778). Appendix E discusses the choice of a decreasing concave/linear function f.

this prediction using grades in compulsory STEM courses as a proxy for acquired knowledge. The sample is restricted to compulsory courses to avoid self-selection into electives, which could potentially change after 1961. I estimate regressions with the following specification:

$$g_{cmt} = \alpha + \beta \frac{E_{cm}^{pre}}{fac_{cm}^{pre}} + \sum_t \gamma_t A_t + \sum_t \delta_t \left(\frac{E_{cm}^{pre}}{fac_{cm}^{pre}} \times A_t \right) + \zeta Z_{cmt} + \psi_m + t_{mt} + u_{cmt} \qquad (14)$$

where I omitted the individual subscript i for ease of notation. g_{cmt} is the standardized grade received in the compulsory course c, which belongs to institute m, in the academic year t. An institute is a bureaucratic entity with a director, an administrative staff, and a dedicated budget that groups courses within a major in the same field of study. $\frac{E_{cm}^{pre}}{fac_{cm}^{pre}}$ is the pre-existing student-faculty ratio in course c, measured as the average between 1958 and 1964. A_t are academic-year fixed effects: either a full set from 1958 to 1968, or just two binary variables (Post 61_t for years between 1961 and 1964 and Post 65_t for years between 1965 and 1968). Z_{cmt} are student and course characteristics. The student variables are gender, a quadratic polynomial of age in year t, high school fixed effects, measures of pre-collegiate ability, major and university fixed effects. Course variables are the tenure and gender of the professor, and an indicator variable that identifies the institute directors. The regression includes institute fixed effects (ψ_m) and institute-specific linear and quadratic time trends that allow for a trend break after 1961 (t_{mt}). Negative and significant estimates of δ_t would imply that grades decreased more after 1961 in courses with higher pre-existing student-faculty ratio, relative to other compulsory courses in the same major.

V.B.1 Professors and Assistants in the Italian University

The identification of the key parameters δ_t in equation (14) hinges on two elements: a large educational expansion after 1961 and sufficient pre-existing variation of student-faculty ratios across courses within a major.

The entry of type B students led to significant crowding of university resources in STEM fields. In 1961, the average student-faculty ratio in first-year compulsory courses was equal to 19.6 in STEM and 18.8 in restricted majors (Appendix Figure A.6, Panel A). In 1968, however, the average student-faculty ratio was equal to 43.9 in STEM and to 21.6 in restricted fields. The increase of student-faculty ratio in STEM was driven by a surge of enrolled students that was not met by an increase in teaching fellows. The average number of teaching fellows per first-year course, instead, slightly decreased over the period under consideration, moving from 5.3 in 1961 to 4.1 in 1968 (Appendix Figure A.6, Panel B).

The variation in student-faculty ratios across compulsory courses within a major (therefore, courses with similar enrollment) was driven primarily by differences in the number of teaching fellows. In Milan, 35 percent of STEM compulsory courses had only 1 teaching fellow, 39 percent had between 2 and 5, and 20 percent had between 6 and 13 (Appendix Figure A.6, Panel C). STEM compulsory courses had on average 4.4 teaching fellows (with a standard deviation of 4.8). Anecdotally, this variation depended primarily on the tenure status of the professor assigned to each course (Clark, 1977). Assistants were a coveted resource[39] and tenured professors were better able to hire them for two main reasons. First, since they sat on the hiring committees, they were more likely to have their requests approved. Second, unlike juniors, senior professors were permanently assigned to the same course and, therefore, had a much larger pool of former students to hire as assistants. Using data from universities in Milan, I find that a compulsory STEM course assigned to a tenured professor had on average 3.1 additional teaching fellows, compared with other compulsory STEM courses (Appendix Table A.5, Column 1). Similarly, a course taught by an institute director had on average 2.5 additional teaching fellows. Since an institute received funds to organize teaching and research, a director usually had the budget to hire assistants for her or his own courses. Other variables are of secondary importance. For example, increasing by 10 the number of enrolled students correlated to only 0.29 additional teaching fellows. These findings hold among all STEM courses (Appendix Table A.5, Column 3) and after the inclusion of professor fixed effects (Appendix Table A.5, Columns 2 and 4).

There are indications that, prior to 1961, student-faculty ratios were already correlated with student learning. Before 1961, in fact, grades in STEM courses with a student-faculty ratio below the 25th percentile were 6.4 percent higher, relative to grades in courses with a student-faculty ratio above the 75th percentile (Appendix Table A.6). Prior to 1961, however, these differences were driven primarily by arbitrary variation in the number of teaching assistants, not by variation in incoming student preparation (which did not vary much at that time because all university STEM students had a type A diploma)

V.B.2 Results

The results show that grades decreased after 1961 in courses with a higher pre-existing student-faculty ratio (Table 5, Columns 1 and 2). A marginal increase in the student-faculty

[39] Assistants were assigned to a specific course and were often former students of the course. Unlike American TAs who are usually in charge of only office hours and sections, Italian teaching assistants had more tasks and greater responsibility. On top of holding office hours and teaching sections, Italian assistants helped with oral examinations, delivered the main lectures when the professor was unavailable, supervised undergraduate theses (all students had to write a thesis in order to graduate), and carried out administrative work (Marbach, Rizzi and Salvemini, 1969).

ratio decreased grades by 0.004σ between 1961 and 1964 and by 0.003σ between 1965 and 1968. Since the average student-faculty ratio increased by 24.3 between 1961 and 1968, the average effect is much larger. In the average course, grades decreased by 0.090σ from 1961 to 1964 and by 0.082σ from 1965 to 1968. Including institute-specific linear and quadratic trends does not modify the baseline estimates. Grades decreased by 0.107σ from 1961 to 1964 and by 0.098σ from 1965 to 1968 in the average STEM course (Table 5, Columns 3 and 4). Difference-in-differences estimates for each academic year confirm that the effect is negative after 1961 (Figure 4). In the average course, grades decreased by 0.110σ in 1961, by 0.153σ in 1962, and by 0.078σ in 1963. After 1963, estimates keep decreasing. In 1968, grades decreased by 0.111σ in the average STEM course.

These results are robust to a variety of robustness checks. Including student fixed effects, grades decreased by 0.083σ from 1961 to 1964 and by 0.110σ from 1965 to 1968 in the average STEM course (Table 5, Columns 5 and 6). Replacing institute fixed effects in equation (14) with course fixed effects, grades decreased by 0.091σ from 1961 to 1964 and by 0.084σ from 1965 to 1968 (Table 5, Columns 7 and 8). In both cases, the estimates are statistically significant and very close to the baseline.

If tenured and untenured professors had a different ability to deal with higher enrollment, the estimates of δ_t would confound two effects: the increase of student-faculty ratios and the unpreparedness of junior professors to handle bigger classes. To address this concern, I estimate equation (14) with the inclusion of professor fixed effects. Since tenure was linked to a specific course, a professor could be tenured faculty in one course and untenured faculty in a second course during the same academic year.[40] Then, professor fixed effects control for any unobservable that is constant across all courses taught by the same faculty member. Including professor fixed effects, grades decreased by 0.098σ from 1961 to 1964 and by 0.087σ from 1965 to 1968 in the average STEM course (Table 5, Columns 9 and 10).

The effect is not driven by the academic outcomes of type B students. Restricting the sample to type A students only, the estimates do not change significantly. Grades decreased by 0.108σ from 1961 to 1964 and by 0.100σ from 1965 to 1968 (Appendix Table A.7).

In the previous regressions, I computed the student-faculty ratios using the number of enrolled students who attended high school in Milan, because data on total enrollment by course and academic year (including students who attended high school in other cities) are not available. To address this measurement error, I adjust enrollment in each course

[40]Universities had a fixed number of chairs to be filled by tenured professors. A chair was linked to a course, whose teaching was a responsibility of the chair holder. The courses that were not linked to a chair were taught by untenured faculty members appointed for a short term among professors with or without a chair.

using information about the coverage of the dataset in different university majors.[41] The resulting augmented student-faculty ratio followed the same increasing trend after 1961, but the average increase was equal to 84.3 from 71.5 in 1961 to 155.8 in 1968 (Appendix Figure A.6, Panel D). Estimating equation (14) with the augmented student-faculty ratio, grades decreased by 0.098σ from 1961 to 1964 and by 0.084σ from 1965 to 1968 in the average STEM course (Appendix Table A.7).

As a placebo test, I estimate equation (14) using data from restricted and non-restricted majors. The average student-faculty ratio increased only by 2.8 in restricted majors and decreased by 61 in non-restricted majors. In these fields, grades did not decrease after 1961 in courses with a higher pre-existing student-faculty ratio (Appendix Table A.7).

V.C Class Heterogeneity and High School Curricula

At the time of enrollment, type A and B students differed greatly in their scientific knowledge (Appendix Table A.1). Type A students, in fact, had studied formal, physical, and life sciences (for example, math, physics, chemistry, and biology) during high school. On the contrary, type B students had studied only applied sciences. In the empirical analysis, then, I exploit 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 curricula of type B schools. Since type B students were less well prepared in these areas, their entry was more likely to disrupt teaching practices and student-to-student interactions. A practical example should further clarify this point. Engineering students had to pass "Technical drawing" and "Mathematics I" in their freshman year, both mandatory to receive a degree. Type B students were proficient in technical drawing, but not prepared for university-level math. Unlike technical drawing, in fact, math was not part of the curricula of type B schools. To isolate the effect of increased class heterogeneity, then, I measure how grades of type A students changed after 1961 in courses where type B were not prepared (like "Mathematics I"), relative to other courses in the same major (like "Technical drawing").

I estimate the following regression, using the academic outcomes of type A students in STEM compulsory courses between 1958 and 1968:

$$g_{cmt} = \alpha + \beta Not\ in\ B\ cv_{cm} + \sum_t \gamma_t A_t + \sum_t \delta_t \left(Not\ in\ B\ cv_{cm} \times A_t\right) + \zeta Z_{cmt} + \psi_m + u_{cmt} \quad (15)$$

[41] For example, on average the number of engineering freshmen who completed high school in Milan were 25.5 percent of the total. I divide enrollment in each engineering course by 0.255 to estimate total enrollment.

As seen in section V.B, g_{cmt} is the standardized grade received in the compulsory course c of institute m during the academic year t. *Not in B* cv_{cm} is equal to 1 if course c was not included in the high school curriculum of type B students. I create this variable using 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 pre-collegiate curriculum of type B students.[42] The remaining variables in equation (15) have been described in section V.B.[43] Negative and significant estimate of δ_t would suggest that grades of type A students decreased after 1961 in courses not included in the high school curriculum of type B, relative to other courses in the same majors. If I did not restrict the sample to type A students, the estimates of δ_t would overmeasure the effect of increased class heterogeneity, due to worse outcomes of type B students in courses not in their pre-collegiate curriculum.

Data on grades of type B students validate the procedure used to create the variable *Not in B* cv_{cm}. Grades of type B students in STEM courses that were not in their pre-collegiate curriculum were 0.243σ lower, compared with other compulsory courses (Appendix Table A.9). This result is robust to the inclusion of student controls (-0.256σ), course controls (-0.212σ), academic year fixed effects (-0.202σ), and student fixed effects (-0.210σ). In these same courses, grades of type B students were 0.279σ lower than grades of type A. Results are robust to the inclusion of student controls (-0.294σ), course controls (-0.293σ), academic year fixed effects (-0.277σ), and student fixed effects (-0.279σ).

Before 1961, the average grade of type A students in courses that were not in the high school curriculum of type B was 2.5 percent higher, compared with other STEM courses (Appendix Table A.10). This difference is small, but precisely estimated. The average student-faculty ratio, however, was not statistically different across the two groups of courses. This suggests that equation (15) is not exploiting the same source of variation used to estimate the effect of decreased quality of education.[44] Characteristics of the professors were balanced across the two groups with the exception of the share of tenured faculty members:

[42] Appendix Table A.8 lists the STEM institutes for which *Not in B* cv_m is 1.

[43] Equation (15) does not include institute-specific trends (as equation (14)), because they are collinear with the treatment variable (*Not in B* $cv_{cm} \times A_t$)

[44] The opposite is also true: the share of courses in the high school curriculum of type B is not statistically different between courses with a high and low pre-existing student-faculty ratio (Appendix Table A.6).

before 1961, courses not included in the pre-collegiate curriculum of type B students were 17.2 percent more likely to be assigned to junior professors.

V.C.1 Results

The estimates of δ_t are negative and significant. In courses not in the pre-collegiate curriculum of type B students, grades of type A decreased by 0.100σ between 1961 and 1964 and by 0.115σ between 1965 and 1968, relative to other compulsory courses (Table 6, Column 1).

The inclusion of institute fixed effects leads to slightly smaller estimates: grades decreased by 0.067σ between 1961 and 1964 and by 0.086σ between 1965 and 1968 (Table 6, Column 2). The difference-in-differences estimates for single academic years show that the effect becomes more negative over time (Figure 5). Grades decreased by 0.056σ in 1961, by 0.077σ in 1963, and by 0.021σ in 1965, even though the last estimate is not statistically significant (p-value 0.533). After 1965, the effect of increased class heterogeneity becomes larger: grades decreased by 0.062σ in 1966, by 0.097σ in 1967, and by 0.170σ in 1968.

The results are robust to different robustness checks. Including student fixed effects, grades decreased by 0.104σ from 1961 to 1964 and by 0.106σ from 1965 to 1968 (Table 6, Column 3). Replacing institute fixed effects with course fixed effects, estimates do not change significantly. In courses not in the pre-collegiate curriculum of type B students, grades of type A decreased by 0.072σ from 1961 to 1964 and by 0.090σ from 1965 to 1968 (Table 6, Column 4). The inclusion of professor fixed effects, however, leads to slightly smaller estimates: grades of type A decreased by 0.050σ from 1961 to 1964 and by 0.062σ from 1965 to 1968 (Table 6, Column 5).

V.D Lower Returns to STEM Degrees

This section aims at quantifying the total effect of educational expansion on the returns to STEM degrees.

First, I examine how the long-run income of STEM students with a type A diploma changed after 1961, relative to type A students enrolled in different university majors or with just a high school diploma. This differential income change is a function of the effects of educational expansion and the changing selection into STEM majors. Moreover, if type A students decreased returns to restricted majors after 1961 by abandoning STEM, these estimates should be interpreted as a lower bound of the true decrease in STEM returns. I will later control for the positive selection out of STEM.

I estimate the following equation:

$$log(\text{income}_{ft}) = \alpha + \sum_f \beta_f \mathrm{C}_f + \sum_t \gamma_t Y_t + \sum_t \delta_t[\mathrm{STEM}_f \times Y_t] + \zeta X_{ft} + u_{ft} \qquad (16)$$

where the unit of observation is a student (i, omitted) in a field of study (f) and in an high school cohort (t). Y_t is a set of fixed effects for high school graduation year. C_f are major fixed effects: separate binary variables for STEM, restricted majors, non-restricted majors, and no university as the omitted category. STEM_f is equal to 1 if the student enrolled in a STEM major. X_{ft} is the usual set of student characteristics. A negative estimate of δ_t would suggest that income of STEM students who completed high school in year t decreased , relative to income of type A students in the same cohort and with different education.

In pre-reform cohorts, type A students who enrolled in a STEM major on average earned 28 percent more in 2005, relative to other type A students (Table 7, column 1). The income gap between STEM and other fields remained unchanged from 1961 to 1964, but decreased by 12.6 percentage points from 1965 to 1968. Difference-in-differences estimates for single cohorts confirm that the negative effect on the STEM premium is larger after 1965. Before 1964, the estimates are constant and not statistically significant (Figure 6). After 1964, the STEM premium decreased by 28.9 percentage points for the 1966 cohort, 27.3 percentage points for the 1967 cohort, and 33.9 percentage points for the 1968 cohort. Estimating equation (16) on the subsample of type A students with a university degree leads to the same findings (Table 7, column 2). The STEM premium, which was equal to 17.4 percent in the pre-reform cohorts, decreased by 12.9 percentage points between 1965 and 1968.

To control for selection, I need to repeat the analysis on a subsample of students who did not modify their major choice after 1961. Section V.A.1 indicated that high-achieving students were more likely to leave STEM fields after 1961. For this reason, I restrict the sample to type A students who scored in the bottom quartile of their high school class.

In pre-reform cohorts, low-achieving type A students who enrolled in STEM earned 24 percent more in 2005, compared with other low-achieving type A (Table 7, column 3). The STEM premium remained unchanged between 1961 and 1964, but decreased by 21.8 percentage points from 1965 to 1968. The analysis on low-achieving type A students with a university degree leads to similar results (Table 7, column 4). The STEM premium, which was equal to 35.1 percent for pre-reform cohorts, decreased by 34.4 percentage points from 1965 to 1968. After controlling for selection, the estimated decrease of STEM returns from 1965 to 1968 almost completely erased the pre-existing STEM premium. This result suggests that the reform might have increased the mismatch between abilities and chosen major among the high-achieving type A students.

Within the group of low-achieving students, I compare graduates from the humanities schools with graduates from the scientific schools. The latter had a relatively high probability of enrolling in STEM (42.3 percent), but did not switch to other majors after 1961. The share that enrolled in STEM decreased by only 2.3 percent and the effect is not statistically significant (not reported, p-value 0.419); in comparison, the share of high achieving type A from the same schools decreased by 11.6 percent after 1965 (not reported, p-value 0.001). On the other hand, low-achieving humanities students in the pre-reform cohorts were not likely to enroll in STEM (16.9 percent).

First, I directly compare income of low-achieving scientific and humanities students. The results can be interpreted as intent-to-treat estimates of the effect of educational expansion on returns to STEM education. In pre-reform cohorts, low-achieving scientific students earned 30.3 percent more than low-achieving humanities students (Table 7, column 5). The scientific premium did not change until 1964, but decreased by 26.7 percentage points from 1965 to 1968. The effects are larger if the sample is further restricted to students with a university degree (Table 7, column 6). The scientific premium, which was equal to 43.2 percent in pre-reform cohorts, decreased by 25.1 percentage points from 1961 to 1964 and by 25.4 percentage points from 1965 to 1968.

The last estimates do not take into account that only some low-achieving scientific students enrolled in STEM fields and were, therefore, affected by educational expansion. Similarly, a minor share of low-achieving humanities students enrolled in STEM. To address this point, I estimate equation (16) instrumenting enrollment in STEM fields with an indicator variable that identifies graduates from the scientific schools. These IV estimates of δ_t measure the effect of educational expansion on income of low-achieving scientific students who enrolled in STEM majors after 1961, compared with low-achieving humanities students who enrolled elsewhere (Table 7, column 7). The STEM premium, which was equal to 40.6 percent in pre-reform cohorts, decreased by 67.5 percentage points from 1965 to 1968. For university graduates, the STEM premium decreased by 53.1 percentage points from 1961 to 1964 and by 55.3 percentage points from 1965 to 1968 (Table 7, column 8), relative to a baseline of 57.3 percent.

This analysis provided two main findings. First, returns to STEM degrees declined after 1961 and the decrease was large enough to either erase or reduce significantly the pre-existing STEM premium. Second, positive selection of type A students out of STEM after 1961 led to an underestimation of the decline in the returns to a STEM education.

VI The Costs of Lower Human Capital

In this section, I relate the decrease of human capital in STEM majors to income losses. In addition, I perform a basic cost-benefit analysis of different plans that the Italian government could have implemented to prevent the unintended consequences of the reform.

Using the model described in section IV, I can relate a decline in human capital to an income loss. Specifically, log income responds to a marginal increase in the STEM enrollment of type B students according to:

$$\frac{\partial \log(W_i^k)}{\partial E_B^k} = \frac{\partial log(w^k)}{\partial E_B^k} + \frac{\partial log(h_i^k)}{\partial E_B^k} \tag{17}$$

where human capital of student i is the sum of skills (measured by the final grade) acquired in each university course.

Section V.B showed that grades decreased by 0.004σ from 1961 to 1968 after a marginal increase in the student-faculty ratio (Table 5, Column 3). I multiply these estimates by the actual student-faculty ratio increase in each university course to compute the grade change per student and course (leave-one-out estimator; Abadie, Chingos and West, 2014). For each student, then, I compute the total decrease in human capital as the sum of these grade changes over all attended university courses. On average, crowding of university resources in STEM majors decreased human capital and income by 3.1 percentage points. Using a more conservative specification of equation (14) with course fixed effects, I estimate a lower bound of 2.0 percentage points.

I repeat the same steps to gauge the effect of higher class heterogeneity. Grades of type A students decreased by 0.067σ from 1961 to 1964 and by 0.086σ from 1965 to 1968 in the courses not taught in type B schools (Table 6, Column 2). On average, higher class heterogeneity decreased human capital and income by 2.3 percentage points (in this case, the baseline coincides with the lower bound).

The combined change of crowding of university resources and higher class heterogeneity decreased income by 5.4 percentage points. In section V.D, I showed that returns to a STEM education declined by 12.6 percentage points between 1965 and 1968 (Table 7, Column 1). Then, lower human capital can account for 42.9 percent of the total income decrease (34.1 percent at the lower bound).[45] In the absence of other general equilibrium effects, lower skill prices (due to higher supply of STEM-educated workers) would account for the remaining share.

[45] Among low-achieving type A students, lower human capital would account for 36.7 percent of the total decline in STEM returns.

In the last part of 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 (-3.1 percentage points) and higher class heterogeneity (-2.3 percentage points) do not change with age, that income increases by 1.8 percent every year, and that the discounting rate is equal to 10 percent.[46] Based on these assumptions, the yearly income loss at age 25 was €875 due to lower quality of education and €649 due to higher class heterogeneity. Their discounted present values over the lifetime (in this case, until retirement at age 65) were €10,192 and €7,562 respectively.

At this point, I estimate the costs of hypothetical actions that the Italian government could have taken to prevent lower human capital. I will only compute the costs of hiring more teaching fellows (professors and assistants). This procedure involves less arbitrary assumptions, but it can underestimate the total costs. In addition, I will use the empirical model described in section IV 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 aims at keeping class heterogeneity unchanged, but allows 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 separated. In a major with 25 courses, this plan costs €690,875 or €5,074 per student.[47] In this case, the benefits of keeping class heterogeneity fixed (€7,562) are greater than the costs (€5,074) for the government.

Second, I consider an alternative scenario in which the government wants to keep 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 assistants to

[46] The yearly increase of income is estimated using SHIW data on males with a university degree (Appendix C). 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).

[47] 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.

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 major with 25 courses, this plan costs €1,666,406 or €12,239 per student.[48] In this case, the benefits of keeping quality of education fixed (€10,192) are slightly lower than the costs (€12,239) for the government.

Third, I consider the case in which the government intends to keep both 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 quality of education and class heterogeneity fixed (€17,754) are higher than the costs (€14,465) for the government.

VII Conclusions

In this paper, I study a dramatic 1961 Italian reform that increased university enrollment in STEM majors by more than 200 percent in only a few years. To do so, I collected and digitized high school registries, university transcripts, and long-run income for 27,236 Italian students who completed high school in Milan between 1958 and 1968.

I show that, after 1961, many more students enrolled in STEM majors and completed their degrees. These are the effects that policy makers intended. However, the new university-educated individuals did not earn higher incomes than people from earlier cohorts who were denied access to STEM majors. I propose three mechanisms to explain this surprising result: higher enrollment decreased the returns to a STEM education through congestion of university resources; class composition changed in a way that produced negative peer effects or that disrupted teaching; the greater supply of STEM-educated workers reduced the returns to university-level STEM skills.

I found evidence that human capital in STEM majors decreased due to crowding of university inputs and to a decrease in the average preparedness of enrolled students. In post-reform cohorts, the returns to a STEM education declined to the point of erasing the pre-existing STEM premium. In addition, I found that the students with higher pre-

[48] 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).

collegiate achievement (not directly affected by the reform) moved out of STEM programs towards majors in which returns were not affected by educational expansion.

This paper studied the consequences of educational expansion on the students who enrolled in university immediately after the policy implementation. Although long-lasting for these students, the negative effects on human capital and income may have dwindled for individuals in the cohorts that followed them. The government, for example, could have increased the resources of public universities so that there was less crowding. Or, it could have improved the curriculum of type B schools so that their graduates were better prepared for university-level programs. Employers may have adjusted to the greater supply of STEM-educated workers by developing businesses that relied more on such skills. However, some of the adverse effects of creating such an abrupt expansion in access are likely to have lingered. For example, hiring new faculty would not necessarily restore the pre-policy level of education quality because the new professors would, as a logical matter, have to have been trained in the "expanded" university system and might, therefore, have lower human capital. In addition, if the entry of type B students in STEM fields caused the (aptitude or other) signal provided by STEM degrees to deteriorate, especially talented type A students would still have deserted STEM fields and gone into restricted fields (like medicine) even if the expansion had been gradual. Moreover, although employers saw an increased supply of STEM-educated workers, they also saw a decrease in the skills of each such worker. Therefore, it is not obvious that they would have responded in ways that ultimately raised labor demand substantially for STEM workers.

This paper exploited few features of the Italian system to achieve identification, but the estimated effects apply more broadly. For example, open-door admission into university was pivotal in this context, but it is not necessary to cause abrupt enrollment increases. China, for example, increased the number of admission slots in its competitive university system in 1999 and enrollment increased by 200 percent in a few years. The reform likely decreased the quality and, therefore, the value of university education.[49] Nor is the US necessarily immune to quick educational expansions. To meet future demand, presidential advisors indicated that the US should take immediate actions to increase by 34 percent annually the number of STEM graduates (PCAST, 2012). More recently, California (Senate Bill 850, 2013) has been considering to allow community colleges to grant applied baccalaureate degrees. The long-term goal is to increase by 40 percent the number of college degrees awarded in state.[50]

[49]Some commentators have worried openly about falling returns to university education in China: http://www.cnn.com/2014/09/09/opinion/china-education-opinion/ and http://business.time.com/2013/06/27/china-just-as-desperate-for-education-reform-as-the-u-s/.

[50]In the first half of the twentieth century, the US education system went through a similar reorganization. Most US states converted their normal schools (a type of two-year school intended only to train people

All these proposed policy changes have the potential to affect the value of a college degree. Moreover, few results in this paper do not even require an increase in university enrollment. The effect of class heterogeneity, in fact, applies to any context in which an education policy changes the ability composition of university students. The results, for example, may provide useful information about the effects of large-scale affirmative action plans swiftly implemented in countries like India during the colonial era or Malaysia and Sri Lanka in 1971 (Sowell, 2005).

What are the broader implications of my findings for investments in human capital? Economic logic suggests that public intervention is potentially needed to overcome three market failures that may lead to sub-optimal human capital investment: (i) liquidity constraints that prevent people from investing in their own education, (ii) the spillovers associated with certain types of education (inventiveness is often cited for university-level STEM education), and (iii) information failures that prevent people from understanding their likely returns to human capital investment. However, economic logic also suggests that the remedies for these failures are things like (i) means-tested financial aid, (ii) scholarships for people whose aptitude suggests that they may be especially likely to generate positive spillovers (scholarships for top science minds, for instance), and (iii) information campaigns. (In fact, at least for the US, the evidence suggests that students are already highly elastic with respect to economic returns in STEM fields.[51] Thus, it is not obvious that greater information campaigns are required.) What economics does not suggest is that the remedy to human capital investment failures is greatly expanded education provision by state-controlled universities that attempt to force students into certain fields selected by the government. Indeed, the disconnect between what economic logic suggests and what states often do has been noted for many years starting with the seminal work of Peltzman (1973). He notes that "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.[52] Peltzman and others writing about this

in elementary education) into four-year universities without actually improving their resources (Labaree, 2008).

[51] Freeman (1971), Freeman (1975), and Ryoo and Rosen (2004) show that engineering students respond strongly to changes in the price of engineering skills and to variation in career prospects. Training delays and cobweb expectations can explain periods of over-supply and under-supply of engineers in the US market. More generally, the fact that students adjust quickly to changes in the returns to education is not surprising. For example, Abramitzky and Lavy (2014) study a change in the redistribution schemes of Israeli kibbutzim that increased returns to education and find that affected students invested more in education.

[52] Peltzman (1973) suggests that in-kind subsidies to public institutions (for example, tuition waivers to in-state students) might decrease the accumulation of human capital. 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 might lead to an overall decrease in quality-adjusted human capital. Several

logical disconnect 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 distortions they introduce via in-kind provision are unlikely to be offset by or forced out by private universities. In short, countries that currently aspire to increase university-level STEM human capital (presumably because they believe it generates positive spillovers[53]) may be reminded by Italy's example that dramatic, field-specific expansions of access are not what most economic models would indicate.

References

Abadie, Alberto, Matthew M. Chingos, and Martin R. West. 2014. "Endogenous Stratification in Randomized Experiments." working paper.

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

Abbring, Jaap H., and James J. Heckman. 2007. "Econometric evaluation of social programs, Part III." In *Handbook of Econometrics*. Vol. 6, , ed. James J. Heckman and Edward E. Leamer, Chapter 72, 5145–5303.

Abramitzky, Ran, and Victor Lavy. 2014. "How Responsive Is Investment in Schooling to Changes in Redistributive Policies and in Returns?" *Econometrica*, 82(4): 1241–1272.

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. 2013. "Cashier or Consultant? Entry Labor Market Conditions, Field of Study, and Career Success." working paper.

papers (Ganderton 1992, Long 2004, Kane 2007, Cellini 2009, Cohodes and Goodman 2014) test and confirm Peltzman's theoretical prediction.

[53] In the macroeconomic literature, Barro (1991) finds that the growth rate of GDP is positively correlated to the level of human capital. Hall and Jones (1999) and Bils and Klenow (2000) find that different education attainments can explain less than one third of the cross-country variation in growth and output per worker, while other factors like institutional and government policies might be responsible for the bulk of the observed variation. By considering differences in both the level and the quality of education, however, Schoellman (2011) and Hanushek (2013) find that human capital can account for a larger share of the international variation in growth and output per worker.

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

Arcidiacono, Peter. 2004. "Ability sorting and the returns to college major." *Journal of Econometrics*, 121(1-2): 343–375.

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

Barro, Robert J. 1991. "Economic Growth in a Cross Section of Countries." *The Quarterly Journal of Economics*, 106(2): 407–443.

Bils, Mark, and Peter J. Klenow. 2000. "Does Schooling Cause Growth?" *American Economic Review*, 5(1968): 1160–1183.

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

Card, David, and Alan B. Krueger. 1992. "Does School Quality Matter? Returns to Education and the Characteristics of Public Schools in the United States." *Journal of Political Economy*, 100(1): 1–40.

Card, David, and Jesse Rothstein. 2007. "Racial Segregation and the Black-White Test Score Gap." *Journal of Public Economics*, 91(11-12): 2158–2184.

Card, David, and Thomas Lemieux. 2001. "Risky Behavior among Youths: An Economic Analysis." In *Risky Behavior among Youths: An Economic Analysis.* , ed. Jonathan Gruber, Chapter 9, 439–482. University of Chicago Press.

Case, Anne, and Angus Deaton. 1999. "School inputs and educational outcomes in South Africa." *The Quarterly Journal of Economics*, , (August): 1047–1084.

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.

Devicienti, Francesco, Agata Maida, and Paolo Sestito. 2007. "Downward Wage Rigidity in Italy: Micro-Based Measures and Implications." *The Economic Journal*, 117(524): F530–F552.

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(August): 1739–1774.

Fabiani, Silvia, and Roberto Sabbatini. 2011. "Wage Adjustment by Italian Firms: Any Difference during the Crisis? A survey-based analysis." *Questioni di Economia e Finanza*, 94.

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.

Freeman, Richard B. 1971. *The Market for College Trained Manpower.* Harvard University Press.

Freeman, Richard B. 1975. "A Cobweb model of the supply and starting salary of new engineers." *Industrial and Labor Relations Review*, 29(2): 236–248.

Ganderton, Philip T. 1992. "The Effect of Subsidies in Kind on the Choice of a College." *Journal of Public Economics*, 48(3): 269–292.

Grubb, Dennis, Richard Jackman, and Richard Layard. 1983. "Wage Rigidity and Unemployment in OECD Countries." *European Economic Review*, 21(1-2): 11–39.

Hall, Robert E., and Charles I. Jones. 1999. "Why do Some Countries Produce so Much More Output per Worker than Others?" *The Quarterly Journal of Economics*, 114(1): 83–117.

Hanushek, Eric A. 2013. "Economic growth in developing countries: The role of human capital." *Economics of Education Review*, 37(C): 204–212.

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. 2000. "Peer Effects in the Classroom: Learning from Gender and Race Variation." NBER Working Paper 7867.

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

Kane, Thomas J. 2007. "Evaluating the Impact of the D.C. Tuition Assistance Grant Program." *The Journal of Human Resources*, XLII(3): 555–582.

Labaree, David F. 2008. "An uneasy relationship: the history of teacher education in the university." In *Handbook of Research on Teacher Education: Enduring Issues in Changing Contexts*. , ed. Marilyn Cochran-Smith, Sharon Feiman Nemser and John McIntyre, Chapter 18. Washington, DC:Association of Teacher Educators.

Lavy, Victor, and Analia Schlosser. 2011. "Mechanisms and Impacts of Gender Peer Effects at School." *American Economic Journal: Applied Economics*, 3(2): 1–33.

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(March): 208–237.

Lazear, Edward P. 2001. "Educational Production." *The 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.

Lefgren, Lars. 2004. "Educational peer effects and the Chicago public schools." *Journal of Urban Economics*, 56(2): 169–191.

Loeb, Susanna, and John Bound. 1996. "The effect of measured school inputs on academic achievement: Evidence from the 1920s, 1930s and 1940s birth cohorts." *The Review of Economics and Statistics*, 78(4): 653–664.

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.

Maurin, Eric, and Sandra McNally. 2008. "Vive la Révolution! Long-Term Educational Returns of 1968 to the Angry Students." *Journal of Labor Economics*, 26(1): 1–33.

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.

President's Council of the Advisors on Science and Technology. 2012. "Engage to Excel: Producing One Million Additional College Graduates with Degrees in Science, Technology, Engineering, and Mathematics."

Ryoo, Jaewoo, and Sherwin Rosen. 2004. "The Engineering Labor Market." *Journal of Political Economy*, 112(1).

Schoellman, Todd. 2011. "Education Quality and Development Accounting." *Review of Economic Studies*, 79(1): 388–417.

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.

Train, Kenneth E. 2009. *Discrete choice methods with simulation.* . Second ed., New York:Cambridge University Press.

Unioncamere. 2013. *Atlante della competitività delle province e delle regioni.* Accessible at http://www.unioncamere.gov.it/Atlante/.

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
No 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 - non-restricted major	0.043	0.159	0.363
College degree	0.637	0.161	0.105
College degree - STEM major	0.269	0.101	0.001
College degree - restricted major	0.353	0.029	0.016
College degree - non-restricted 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) - non-restricted 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 by 11,433 type A, 8,813 type B, and 6,690 type C. The number of course-student combinations from college transcripts are 144,572 for type A, 15,493 for type B, and 12,456 for type C. 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. Non-restricted majors are business, economics and statistics. Income is winsorized at the 2nd and 98th percentile. *Adjusted income* is taxable income in 2005 adjusted for age effects. Details on this procedure can be found in appendix C.
Sources: High school archives, college transcripts, and income tax returns in 2005.

Table 2: Cohort Effects, Education and Income of Type B Students

	College Enrollment		College 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)
No 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 college enrollment in columns (1) and (2), a dummy for college graduation in columns (3) and (4), and log income in column (5). The function F is logit for college 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 from 1965 to 1968. The omitted category is represented by the cohorts that graduated between 1958 and 1960. X_{it} is a set of student characteristics that include gender, the high school exit score, the average exit score of the high school classmates, a dummy for home schooled students, a dummy for students that did not repeat a grade in high school, and high school fixed effects. Robust standard errors in parentheses, *** 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. Income tax returns in 2005.

Table 3: Difference-in-Differences, Education and Income of Type B

	College Enrollment		College Graduation		Log Income	
	1961-64	1965-68	1961-64	1965-68	1961-64	1965-68
	(1)	(2)	(3)	(4)	(5)	(6)
Type B						
High- vs Low-achieving	0.071*	0.060	0.079***	0.097***	-0.000	0.079
	(0.040)	(0.038)	(0.029)	(0.028)	(0.108)	(0.103)
High- vs Low-propensity	0.001	0.121***	0.031	0.136***	0.054	0.160
	(0.037)	(0.035)	(0.024)	(0.024)	(0.102)	(0.098)
Type B vs Type A						
All	0.036**	0.159***	-0.000	0.032**	0.102*	-0.074
	(0.017)	(0.016)	(0.015)	(0.015)	(0.057)	(0.054)
Males	0.049***	0.199***	0.009	0.061***	0.150**	-0.055
	(0.018)	(0.017)	(0.017)	(0.017)	(0.061)	(0.058)
Type B vs Type C						
All	0.062***	0.220***	0.046***	0.125***	0.136**	0.059
	(0.020)	(0.019)	(0.013)	(0.013)	(0.066)	(0.066)
Males	0.086***	0.260***	0.053***	0.131***	0.147**	0.284***
	(0.024)	(0.024)	(0.016)	(0.016)	(0.067)	(0.070)
Mean, 1958-60	0.292	0.292	0.082	0.082	10.447	10.447
HS fixed effects	Yes	Yes	Yes	Yes	Yes	Yes

Notes: Each row contains difference-in-differences estimators from three regressions, using different subsamples of students. Coefficients are estimated from outcome$_{it} = \alpha + \sum_t \beta_t Y_t + \gamma \mathrm{B}_i + \delta_1[\mathrm{B}_i \times \text{Post } 61_t] + \delta_2[\mathrm{B}_i \times \text{Post } 65_t] + \zeta X_{it} + u_{it}$. The dependent variable is college enrollment in columns (1) and (2), college graduation in columns (3) and (4), and log income in columns (5) and (6). Odd columns show estimates of δ_1, while even columns of δ_2. Y_t is a set of cohort effects. Post 61 is equal to 1 for the cohorts that completed high school between 1961 and 1964, while Post 65 is 1 for the cohort that graduated from 1965 to 1968. The omitted category is represented by the cohorts that graduated between 1958 and 1960. X_{it} is a set of student characteristics that include gender, the high school exit score, the average exit score of the high school classmates, a dummy for home schooled students, a dummy for students that did not repeat a grade in high school, and high school fixed effects. "High- vs Low-achieving" compares type B in the top quartile and bottom quartile of the ability distribution. "High- vs Low-propensity" compares type B with highest (top third) and lowest propensity to enroll in STEM majors. The predicted propensity is computed from observable characteristics and the enrollment choices of post-1961 cohorts. Robust standard errors in parentheses, *** 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. Income tax returns in 2005.

Table 4: 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)
No 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, non-restricted major, or no college (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 that completed high school in Milan, Italy; 1958-1968.

Table 5: Quality of Education, Effect on Standardized Grades

	Baseline				Student FEs		Course FEs		Professor FEs	
	Coeff.	Average Effect	Coeff.	Average Effect	Coeff.	Average Effect	Coeff.	Average Effect	Coeff.	Average Effect
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
$\frac{E^{pre}_{cm}}{fac^{pre}_{cm}}$ x $Post61_t$	-0.004***	-0.090	-0.004***	-0.107	-0.003***	-0.083	-0.004***	-0.091	-0.004***	-0.098
	(0.001)	(0.024)	(0.001)	(0.031)	(0.001)	(0.026)	(0.001)	(0.032)	(0.001)	(0.034)
$\frac{E^{pre}_{cm}}{fac^{pre}_{cm}}$ x $Post65_t$	-0.003***	-0.082	-0.004***	-0.098	-0.005***	-0.110	-0.003***	-0.084	-0.004**	-0.087
	(0.001)	(0.024)	(0.001)	(0.030)	(0.001)	(0.026)	(0.001)	(0.031)	(0.002)	(0.038)
Student controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Course controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Institute FEs	Yes	Yes	Yes	Yes	Yes	Yes	No	No	Yes	Yes
Institute trends	No	No	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Observations	62,417	62,417	62,417	62,417	62,418	62,418	62,417	62,417	62,417	62,417

Notes: Effect on STEM grades of a decrease in the quality of education. The "Average Effect" is the effect of an average increase in the student-faculty ratio (+24.3). The coefficients are computed from $g_{cmt} = \alpha + \beta \frac{E^{pre}_{cm}}{fac^{pre}_{cm}} + \sum_t \gamma_t A_t + \delta_1 \left(\frac{E^{pre}_{cm}}{fac^{pre}_{cm}} \times \text{Post } 61_t \right) + \delta_2 \left(\frac{E^{pre}_{cm}}{fac^{pre}_{cm}} \times \text{Post } 65_t \right) + \zeta Z_{cmt} + \psi_m + t_{mt} + u_{cmt}$, using data from STEM compulsory courses. g_{cmt} are standardized grades, and $\frac{E^{pre}_{cm}}{fac^{pre}_{cm}}$ is the pre-existing student-faculty ratio (average over 1958-64). A_t is a set of academic year fixed effects. Post 61 is equal to 1 for years between 1961 and 1964, while Post 65 is 1 academic years between 1965 and 1968. Z_{cmt} is a set of student and course characteristics. The student variables include gender, high school fixed effects, a quadratic polynomial of age, high school exit score, the average score of high school classmates, a dummy for home-schooled students, a dummy for students that did not repeat a grade in high school, major and university fixed effects. The course characteristics are the tenure and gender of the professor, a binary variable that identifies professors that are institute directors, institute fixed effects, and institute-specific linear and quadratic trends. An institute is a group of homogeneous courses within a major. Robust standard errors in parentheses, *** p<0.01, ** p<0.05, * p<0.1.
Sources: College transcripts of students that completed high school in Milan, Italy; 1958-1968.

Table 6: Class Heterogeneity, Effect on Standardized Grades of Type A Students

	Baseline		Student FEs	Course FEs	Professor FEs
	(1)	(2)	(3)	(4)	(5)
Not in B cv_c x $Post61_t$	-0.100***	-0.067***	-0.104***	-0.072***	-0.050*
	(0.025)	(0.026)	(0.024)	(0.026)	(0.029)
Not in B cv_c x $Post65_t$	-0.115***	-0.086***	-0.106***	-0.090***	-0.062*
	(0.025)	(0.026)	(0.025)	(0.027)	(0.033)
Student controls	Yes	Yes	Yes	Yes	Yes
Course controls	Yes	Yes	Yes	Yes	Yes
Academic year FEs	Yes	Yes	Yes	Yes	Yes
Institute FEs	No	Yes	Yes	No	Yes
Observations	53,651	53,651	53,651	53,651	53,651

Notes: Effect of an increase of class heterogeneity on STEM grades. The coefficients are computed from $g_{ct} = \alpha + \beta Not\ in\ B\ cv_c + \sum_t \gamma_t A_t + \delta_1 (Not\ in\ B\ cv_c \times \text{Post } 61_t) + \delta_2 (Not\ in\ B\ cv_c \times \text{Post } 65_t) + \zeta Z_{cmt} + \psi_m + u_{cmt}$, using data from type A students enrolled in STEM compulsory courses. g_{cmt} are standardized grades and *Not in B* cv_c is 1 if course c was not included in the curricula of type B schools. A_t is a set of academic year fixed effects. Post 61_t is equal to 1 for years between 1961 and 1964, while Post 65_t is 1 academic years between 1965 and 1968. Z_{cmt} is a set of student and course characteristics. The student variables include gender, high school fixed effects, a quadratic polynomial of age, high school exit score, the average score of high school classmates, a dummy for home-schooled students, a dummy for students that did not repeat a grade in high school, major and university fixed effects. The course characteristics are the tenure and gender of the professor, a binary variable that identifies professors that are institute directors, and institute fixed effects. An institute is a group of homogeneous courses within a major. Robust standard errors in parentheses, *** p<0.01, ** p<0.05, * p<0.1.
Sources: College transcripts of students that completed high school in Milan, Italy; 1958-1968.

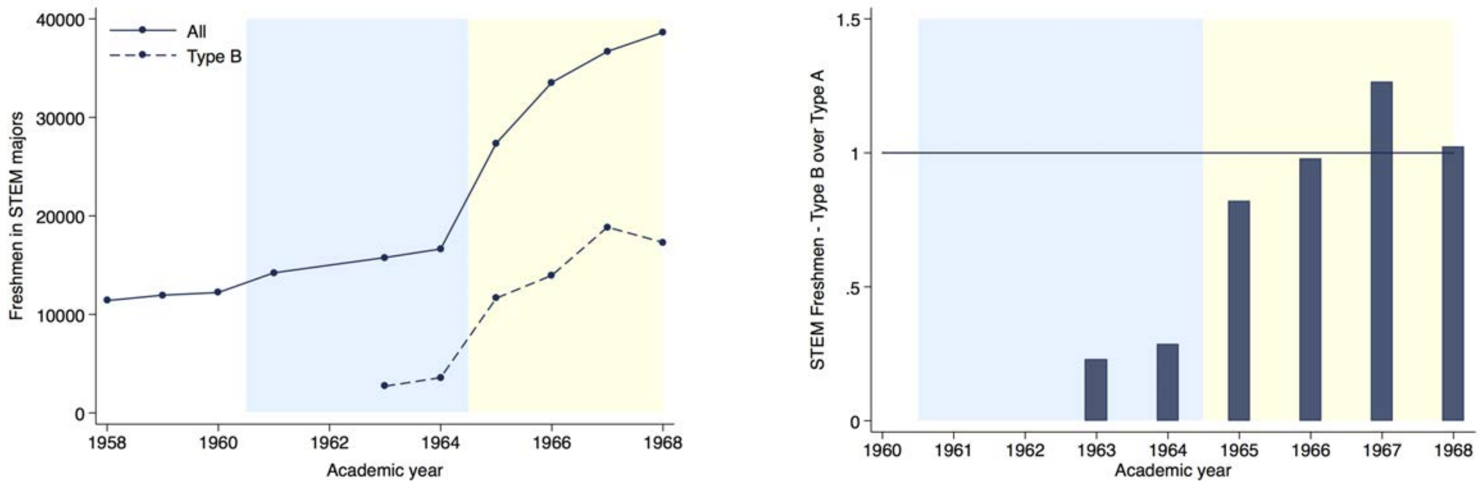
Table 7: Type A, Total Effect on STEM Incomes

	All Type A		Low-Achieving Type A					
			OLS		Intent-To-Treat		IV	
	All	Grads	All	Grads	All	Grads	All	Grads
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
STEM x Post 61	-0.006	0.017	0.021	-0.196			-0.437	-0.758*
	(0.076)	(0.081)	(0.147)	(0.155)			(0.544)	(0.439)
STEM x Post 65	-0.135*	-0.138*	-0.246*	-0.421***			-1.125**	-0.806**
	(0.070)	(0.075)	(0.140)	(0.141)			(0.497)	(0.400)
Scientific x Post 61					-0.123	-0.289*		
					(0.145)	(0.164)		
Scientific x Post 65					-0.311**	-0.293**		
					(0.135)	(0.146)		
STEM/Scientific premium (logs)	0.247	0.160	0.215	0.301	0.265	0.359	0.341	0.453
Student controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
HS fixed effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
HS graduation year FEs	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Major FEs	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Observations	9,495	6,398	3,072	1,731	3,072	1,731	3,072	1,731

Notes: The coefficients are computed from $log(\text{income}_{ct}) = \alpha + \sum_c \beta_c \text{C}_c + \sum_t \gamma_t Y_t + \delta_1[\text{STEM}_c \times \text{Post } 61_t] + \delta_2[\text{STEM}_c \times \text{Post } 65_t] + \zeta X_{ct} + u_{ct}$. Y_t is a set of fixed effects for high school graduation year, and C_c are major fixed effects (STEM, restricted major, unrestricted major, and no college as omitted category). STEM_c is equal to 1 if a student enrolled in STEM. Post 61_t is one for cohorts that graduated between 1961 and 1964, and Post 65_t is 1 for cohorts that graduated between 1965 and 1968. X_i is a set of student characteristics. Columns (1) and (2) use the whole sample of type A students. Columns (3) to (8) use only type A students in the bottom quarter of the ability distribution. Columns (5) and (6) compare low-achieving graduates from the scientific schools to low-achieving graduates from the humanistic schools. Columns (7) and (8) show IV estimates: the instruments for $\text{STEM}_c \times$ Post 61_t and $\text{STEM}_c \times$ Post 65_t are $\text{Scientific}_c \times$ Post 61_t and $\text{Scientific}_c \times$ Post 61_t, where Scientific_c is 1 for low-achieving graduates from the scientific schools. "Grads" restricts the sample to college graduates. The STEM/Scientific premium is the income premium in 2005 (in logs) associated with a STEM degree/Scientific diploma for different sub-groups of students in pre-1961 cohorts. Robust standard errors in parentheses, *** $p<0.01$, ** $p<0.05$, * $p<0.1$.

Sources: College transcripts and income in 2005 of type A that completed high school in Milan, Italy; 1958-1968.

Figure 1: Freshmen in STEM Majors



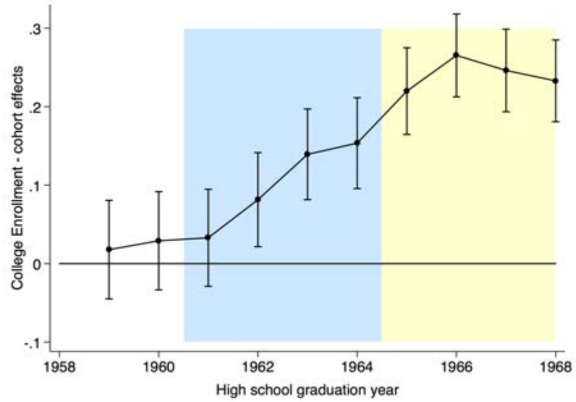
A. Number of Freshmen in STEM Majors

B. Ratio of Type B to Type A

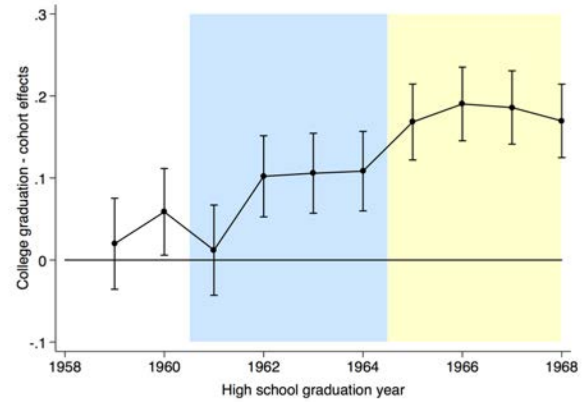
Notes: Panel A: “All” counts the total number of students enrolled in the freshman year of a STEM major for each academic year; “Type B” is the number of students with a type B diploma enrolled in the freshman year of a STEM major. STEM majors are engineering, mathematics, physics, chemistry, biology, geology, natural science, and agricultural sciences. Observations in 1961 and 1962 are missing. The blue shaded area denotes the first phase 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.

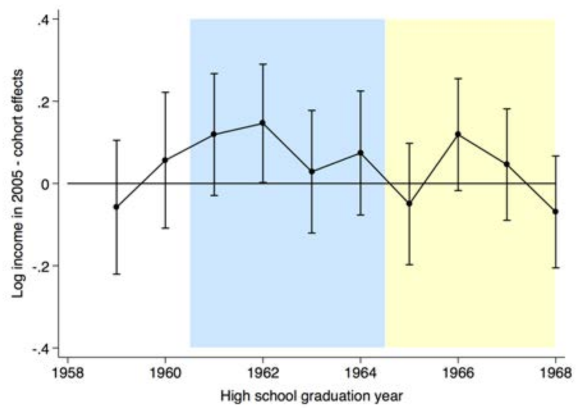
Figure 2: Cohort Effects, Education and Income of Type B Students



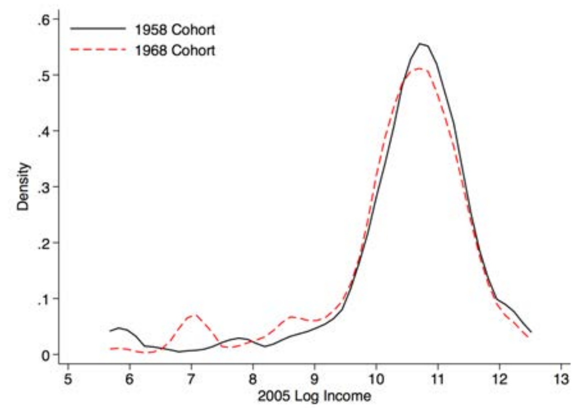
A. College Enrollment



B. College Graduation



C. Income in 2005

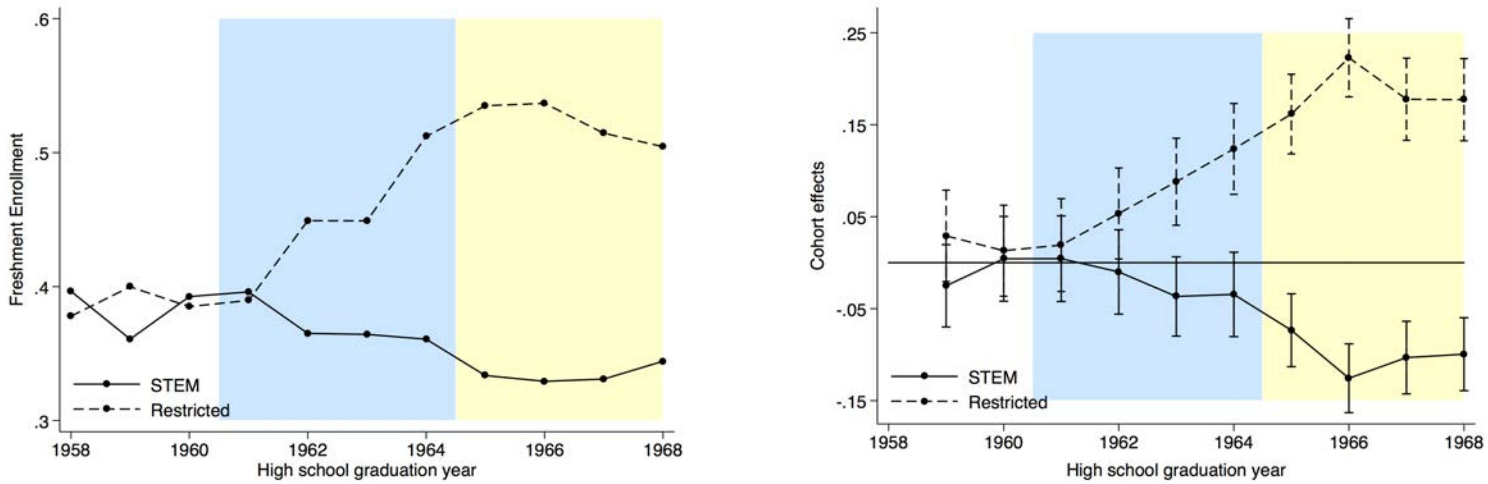


D. Income Distribution

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 college enrollment, a dummy for college graduation, or log income. The function F is logit for college 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 that include gender, the high school exit score, the average exit score of the high school classmates, a dummy for home schooled students, a dummy for students that did not repeat a grade in high school, and high school fixed effects. The blue shaded area denotes the first phase the the reform: between 1961 and 1964, enrollment of type B students in STEM majors was capped. The yellow shaded area denotes the second phase of the reform: after 1965, the enrollment cap was lifted. 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 that completed high school in Milan, Italy; 1958-1968.

Figure 3: Type A, Enrollment in STEM and Restricted Majors

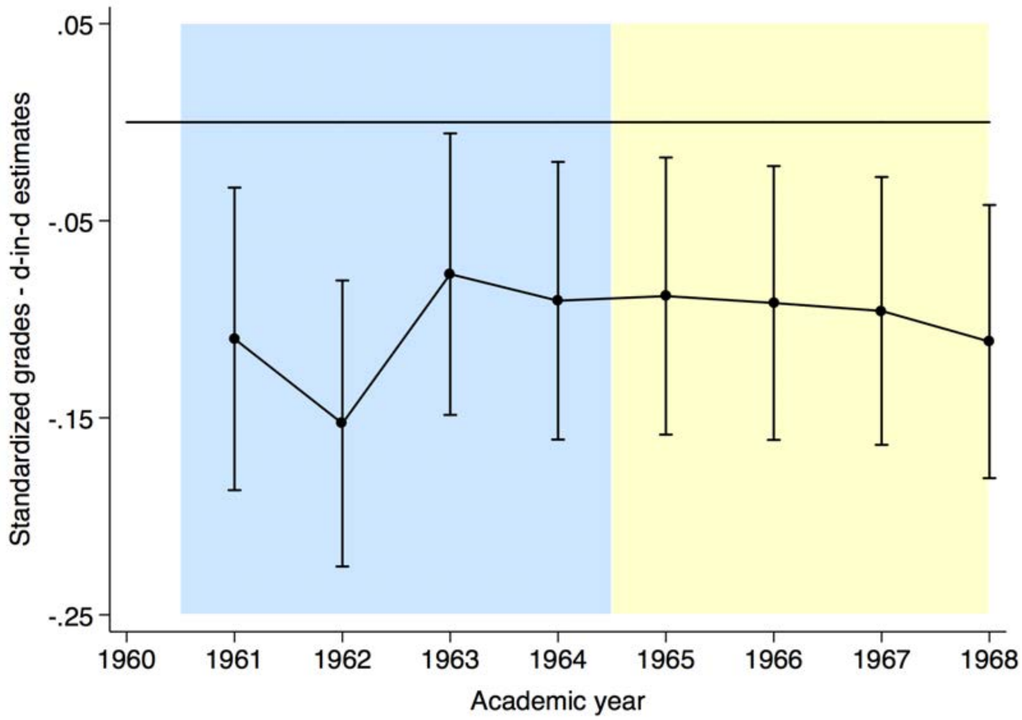


A. Raw Shares

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 from the multinomial logit $ln(\frac{Pr(\text{major}_{it}=k)}{Pr(\text{no college})}) = \alpha_k + \beta_k X_{it} + \sum_t \gamma_{kt} \cdot Y_t$, where k is either a STEM, restricted, non-restricted major, or no college (baseline). Y_t is a set of year of high school graduation fixed effects with 1958 as omitted category. X_{it} 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. The bars represent 95 percent confidence intervals. 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 blue shaded area denotes the first phase the the reform: between 1961 and 1964, enrollment of type B students in STEM majors was capped. The yellow shaded area denotes the second phase of the reform: after 1965, the enrollment cap was lifted. Sources: School data of type A students that completed high school in Milan, Italy; 1958-1968.

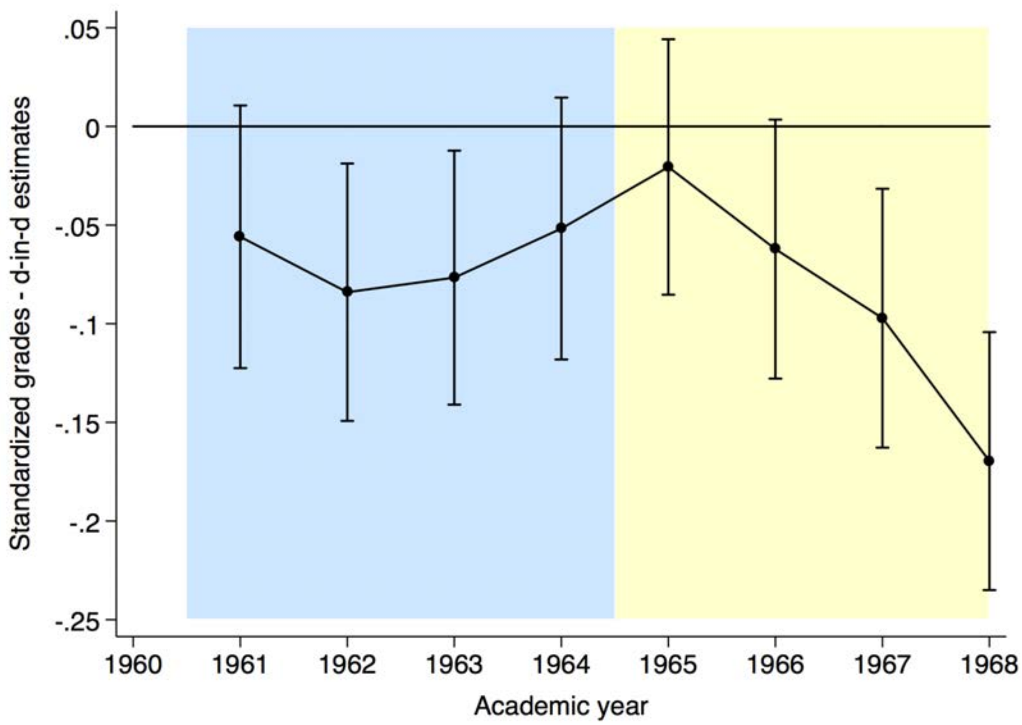
Figure 4: Quality of Education, Average Effect on Grades in STEM Majors



Notes: Effect on grades in STEM compulsory courses of an average decrease in the quality of education (+24.3 student-faculty ratio). The coefficients are computed from $g_{cmt} = \alpha + \beta \frac{E^{pre}_{cm}}{fac^{pre}_{cm}} + \sum_t \gamma_t A_t + \delta_t \left(\frac{E^{pre}_{cm}}{fac^{pre}_{cm}} \times A_t \right) + \zeta Z_{cmt} + \psi_m + t_{mt} + u_{cmt}$, using data from STEM compulsory courses. g_{cmt} are standardized grades, and $\frac{E^{pre}_{cm}}{fac^{pre}_{cm}}$ is the pre-existing student-faculty ratio (average over 1958-64). A_t is a set of academic year fixed effects. Z_{cmt} is a set of student and course characteristics. The student variables include gender, high school fixed effects, a quadratic polynomial of age, high school exit score, the average score of high school classmates, a dummy for home-schooled students, a dummy for students that did not repeat a grade in high school, major and university fixed effects. The course characteristics are the tenure and gender of the professor, a binary variable that identifies professors that are institute directors, institute fixed effects, and institute-specific linear and quadratic trends. An institute is a group of homogeneous courses within a major. The blue shaded area denotes the first phase the the reform: between 1961 and 1964, enrollment of type B students in STEM majors was capped. The yellow shaded area denotes the second phase of the reform: after 1965, the enrollment cap was lifted. The bars represent 95 percent confidence intervals. The omitted academic years are 1958-1960.

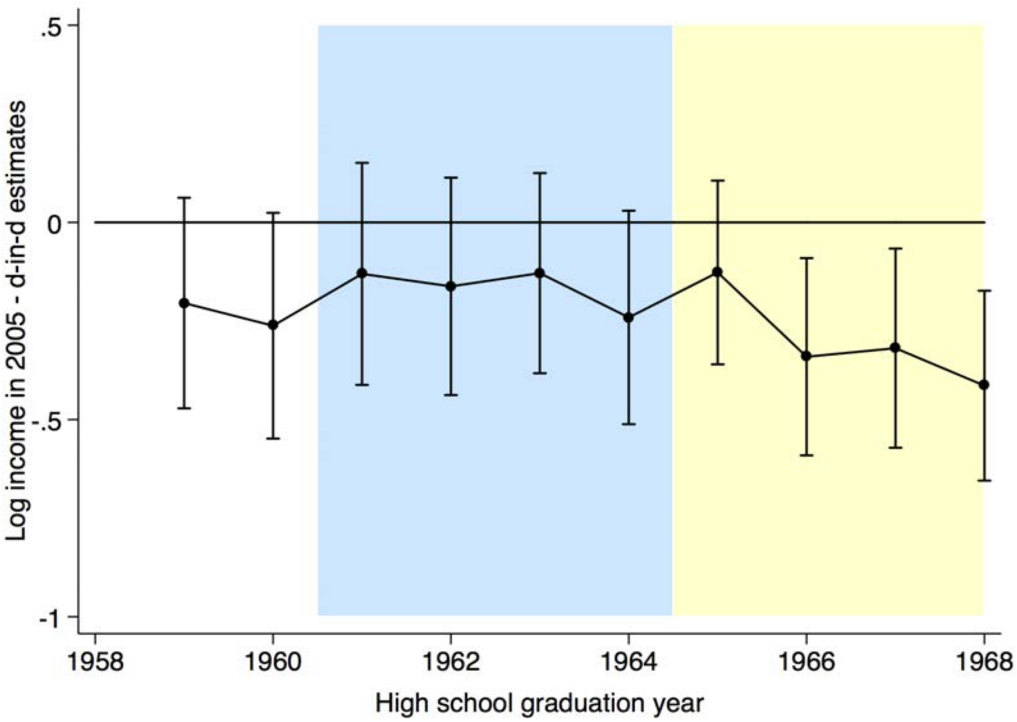
Sources: College transcripts of students that completed high school in Milan, Italy; 1958-1968.

Figure 5: Class Heterogeneity, Effect on Grades of Type A in STEM Majors



Notes: Effect of an increase in the degree of class heterogeneity on STEM grades. The coefficients are computed from $g_{ct} = \alpha + \beta Not\ in\ B\ cv_c + \sum_t \gamma_t A_t + \sum_t \delta_t\,(Not\ in\ B\ cv_c \times A_t) + \zeta Z_{cmt} + \psi_m + u_{cmt}$, using data from type A students enrolled in STEM compulsory courses. g_{cmt} are standardized grades and $Not\ in\ B\ cv_c$ is 1 if course c was not included in the curricula of type B schools. A_t is a set of academic year fixed effects. Z_{cmt} is a set of student and course characteristics. The student variables include gender, high school fixed effects, a quadratic polynomial of age, high school exit score, the average score of high school classmates, a dummy for home-schooled students, a dummy for students that did not repeat a grade in high school, major and university fixed effects. The course characteristics are the tenure and gender of the professor, a binary variable that identifies professors that are institute directors, and institute fixed effects. The blue shaded area denotes the first phase the the reform: between 1961 and 1964, enrollment of type B students in STEM majors was capped. The yellow shaded area denotes the second phase of the reform: after 1965, the enrollment cap was lifted. The bars represent 95 percent confidence intervals. The omitted academic years are 1958-1960.
Sources: College transcripts of students that completed high school in Milan, Italy; 1958-1968.

Figure 6: Type A, Effect on Income Premium of STEM Education



Notes: The marginal cohort effects are estimated from $log(\text{income}_{ct}) = \alpha + \sum_c \beta_c \mathrm{C}_c + \sum_t \gamma_t Y_t + \sum_t \delta_t[\mathrm{STEM}_c \times Y_t] + \zeta X_{ct} + u_{ct}$ on all type A students.Y_t is a set of fixed effects for high school graduation year, and C_c are major fixed effects (STEM, restricted major, unrestricted major, and no college as omitted category). STEM_c is equal to 1 if a student enrolled in STEM. X_i is a set of student characteristics that includes gender, high score exit score, the average score of high school classmates, dummy for home schooled, and dummy for non-repeaters. STEM majors are engineering, physics, mathematics, biology, geology, natural science, chemistry, and agricultural science. The blue shaded area denotes the first phase the the reform: between 1961 and 1964, enrollment of type B students in STEM majors was capped. The yellow shaded area denotes the second phase of the reform: after 1965, the enrollment cap was lifted. The bars represent 95 percent confidence intervals. The omitted academic years are 1958-1960.
Sources: School data and income in year 2005 of type A students that completed high school in Milan, Italy; 1958-1968.

APPENDIX

A Additional Tables and Figures

Table A.1: High school curricula

	Type A	Type B	Type C
	(1)	(2)	(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 college-prep high schools that focus on either the humanities or science. Type B and C are technical high schools that 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 A.2: IV, Returns to College Education for Type B Students

	OLS	IV	Confidence Interval (IV)	F-Stat	Obs.
	(1)	(2)	(3)	(4)	(5)
IV: Post 61_t, Post 65_t					
All Type B	0.337***	-0.225	[-0.691, 0.240]	101.96	7,381
	(0.032)	(0.237)			
Controls for region of residence in 2005	0.330***	-0.262	[-0.724, 0.201]	101.93	7,381
	(0.032)	(0.236)			
1^{st} Quartile of Ability	0.305***	-1.349**	[-2.582, -0.116]	23.39	2,561
	(0.070)	(0.629)			
2^{nd} Quartile of Ability	0.221***	0.189	[-0.714, 1.092]	26.65	1,807
	(0.068)	(0.461)			
3^{rd} Quartile of Ability	0.473***	-0.046	[-0.876, 0.784]	28.54	1,512
	(0.065)	(0.424)			
4^{th} Quartile of Ability	0.347***	0.079	[-0.636, 0.795]	26.11	1,501
	(0.058)	(0.365)			
IV: Post 65_t					
All Type B	0.282***	-0.458*	[-0.872, 0.089]	177.85	7,381
	(0.034)	(0.247)			
IV: Separate cohort dummies (1961-1968)					
All Type B	0.337***	-0.109	[-0.532, 0.314]	30.35	7,381
	(0.032)	(0.216)			
IV: Post $61_t \times 4^{th}$Q, Post $65_t \times 4^{th}$Q					
High- vs Low-achieving	0.356***	0.726	[-1.306, 2.758]	5.05	4,062
	(0.046)	(1.037)			
STEM degrees; IV: Post 61_t, Post 65_t					
All Type B	0.337***	-0.231	[-0.666, 0.205]	191.80	7,381
	(0.037)	(0.222)			

Notes: The table shows estimations of equation (2). Column (1) shows the OLS estimator of β_1. Column (2) shows the IV estimator of β_1 and column (3) the corresponding 95 percent confidence interval. College education is instrumented by two dummy variables: Post 61_t is 1 for the cohorts that completed high school between 1961 and 1965, while Post 65_t is 1 for cohorts that completed high school from 1965 to 1968. The quartiles of ability are computed from the distribution of high school scores. The next-to-last section restricts the sample to type B in the top and bottom quartile of the ability distribution. The instrumental variable is the interaction of Post 61_t and Post 65_t with a dummy that is 1 for students in the top quartile of the ability distribution. The last section replaces degree$_{it}$ with STEM degree$_{it}$, which is 1 for students that received a STEM college degree. The F-Statistic tests for the joint significance of the instruments. Robust standard errors in parentheses, *** 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. Income tax returns in 2005.

Table A.3: 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 (13) with data of the pre-1961 cohorts. *** $p<0.01$, ** $p<0.05$, * $p<0.1$.
Sources: School data of type A students that completed high school in Milan, Italy; 1958-1968.

Table A.4: 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 A.3. Two groups of restricted majors (Architecture, Law and PoliSci) are not reported. *** $p<0.01$, ** $p<0.05$, * $p<0.1$.

Table A.5: STEM, Determinants of the Number of Teaching Fellows

	Compulsory Courses		All Courses	
	(1)	(2)	(3)	(4)
Tenured professor	3.096***	2.646***	3.814***	3.902***
	(0.533)	(0.599)	(0.476)	(0.506)
Institute director	2.518***	1.405	2.172***	0.899
	(0.611)	(1.324)	(0.422)	(0.736)
Female professor	-0.528*	0.468	-0.571**	0.114
	(0.309)	(1.124)	(0.243)	(0.943)
Compulsory course			-0.009	0.326
			(0.334)	(0.466)
Number of students	0.029***	0.025***	0.030***	0.024***
	(0.006)	(0.006)	(0.005)	(0.005)
Lagged average grade	0.019	0.296**	0.018	0.150*
	(0.163)	(0.115)	(0.101)	(0.081)
Mean, 1958-68	4.38	4.38	3.39	3.39
Major-university FEs	Yes	Yes	Yes	Yes
Academic year FEs	Yes	Yes	Yes	Yes
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. 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. College transcripts of students that completed high school in Milan, Italy; 1958-1968.

Table A.6: STEM, Courses with Low and High Student-Faculty Ratio

	Bottom Quartile	Top Quartile	Difference
	(1)	(2)	(3)
Tenured professor	0.691	0.214	0.477***
Institute director	0.709	0.233	0.476***
Female professor	0.073	0.252	-0.179***
Grades (18-31)	24.382	22.924	1.458***
Not in B cv	0.653	0.713	-0.060

Notes: Only compulsory courses. Below (Above) Median are courses with pre-1964 student-faculty ratio below (above) median. “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. Means computed for academic years 1958-1960. *** 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. College transcripts of students that completed high school in Milan, Italy; 1958-1968.

Table A.7: Quality of Education, Additional Results and Robustness Checks

	Academic Years 1961-1964		Academic Years 1965-1968		
	Coeff.	Average Effect	Coeff.	Average Effect	Obs.
	(1)	(2)	(3)	(4)	(5)
STEM Majors					
Only type A students	-0.004***	-0.108	-0.004***	-0.100	52,815
	(0.001)	(0.031)	(0.001)	(0.031)	
Augmented student-faculty ratio	-0.001***	-0.098	-0.001***	-0.084	62,417
	(0.000)	(0.029)	(0.000)	(0.028)	
Restricted Majors	-0.003	-0.009	-0.002	-0.007	39,357
	(0.003)	(0.008)	(0.003)	(0.008)	
Non-Restricted Majors	-0.000	0.023	0.001	-0.070	12,888
	(0.002)	(0.095)	(0.002)	(0.096)	

Notes: Each row shows coefficients and average effects estimated from a different regression. The specification is $g_{cmt} = \alpha + \beta \frac{E^{pre}_{cm}}{fac^{pre}_{cm}} + \sum_t \gamma_t A_t + \delta_1 \left(\frac{E^{pre}_{cm}}{fac^{pre}_{cm}} \times \text{Post } 61_t \right) + \delta_2 \left(\frac{E^{pre}_{cm}}{fac^{pre}_{cm}} \times \text{Post } 61_t \right) + \zeta Z_{cmt} + \psi_m + t_{mt} + u_{cmt}$. g_{cmt} are standardized grades , and $\frac{E^{pre}_{cm}}{fac^{pre}_{cm}}$ is the pre-existing student-faculty ratio (average over 1958-64). Column (1) shows the estimator of δ_1, while column (3) shows the estimator of δ_2. "Average Effect" is the effect of an average change in the student-faculty ratio between 1961 and 1968: +24.3 in STEM (+84.3 in case of the augmented student-faculty ratio), +2.8 in restricted majors, -61 in non-restricted majors. Robust standard errors in parentheses, *** p<0.01, ** p<0.05, * p<0.1.
Sources: College transcripts of students that completed high school in Milan, Italy; 1958-1968.

Table A.8: 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 A.9: 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***	-0.212***	-0.202***	-0.210***	9,857
	(0.020)	(0.022)	(0.024)	(0.024)	(0.022)	
Type B vs Type A						
Not in B cv_c x $TypeB_i$	-0.279***	-0.294***	-0.293***	-0.277***	-0.279***	63,522
	(0.022)	(0.022)	(0.022)	(0.021)	(0.021)	
Student controls	No	Yes	Yes	Yes	Yes	
Course controls	No	No	Yes	Yes	Yes	
Academic year FEs	No	No	No	Yes	Yes	
Student FEs	No	No	No	No	Yes	

Notes: Each cell shows the coefficient from a separate regression. Coefficient measure the average grades of type B students in college courses that were not included in their high school cv, compared with other courses. The first row shows coefficients from $g_{ict} = \alpha + \beta Not\ in\ B\ cv_c + \zeta Z_{ict} + \psi_t + u_{ict}$, using data on type B only. g_{ict} are standardized scores in STEM compulsory exams. $Not\ in\ B\ cv_c$ is 1 if a course is included in the high school cv of type B. Z_{ict} is a set of student and course characteristics and ψ_t are academic year fixed effects. The second row shows coefficients from $g_{ict} = \alpha + \beta Not\ in\ B\ cv_c + \gamma \text{Type B}_i + \delta Not\ in\ B\ cv_c \times \text{Type B}_i + \zeta Z_{ict} + \psi_t + u_{ict}$, using data on type A and B. Robust standard errors in parentheses, *** p<0.01, ** p<0.05, * p<0.1.
Sources: College transcripts of students that completed high school in Milan, Italy; 1958-1968.

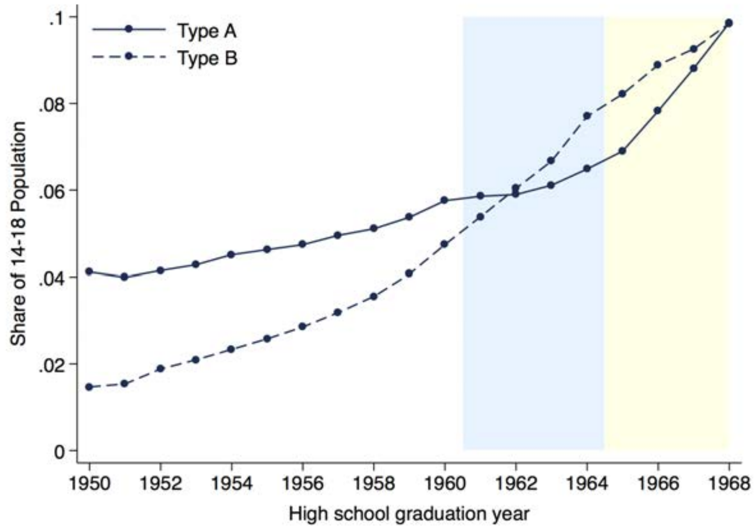
Table A.10: STEM, Courses in the High School Cv of Type B Students

	In type B cv	Not in type B cv	Difference
	(1)	(2)	(3)
Tenured professor	0.261	0.433	-0.172**
Institute director	0.500	0.356	0.144
Female professor	0.152	0.212	-0.060
Grades (18-31)	22.864	23.444	-0.580***
Student-faculty ratio	18.469	17.906	0.563

Notes: Only compulsory courses. "Not in type B cv" are courses not included in the curricula of type B high schools. "Grades" are non-standardized grades. "Tenured", "Institute director", "Female professor" are characteristics of the professor assigned to the course. Means computed for academic years 1958-1960. *** p<0.01, ** p<0.05, * p<0.1.

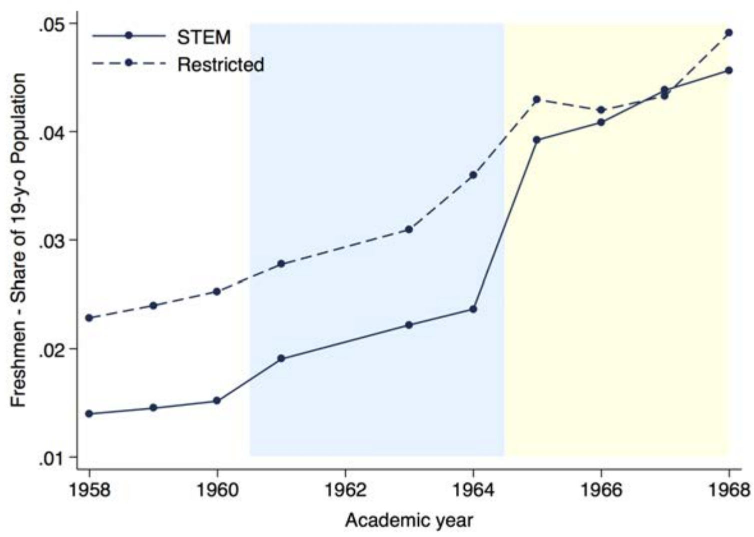
Sources: Annals of Università Statale di Milano, Politecnico di Milano, and Università Cattolica del Sacro Cuore. College transcripts of students that completed high school in Milan, Italy; 1958-1968.

Figure A.1: Type A and Type B Students



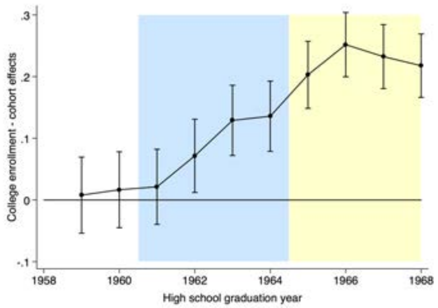
Notes: Type A and type B students as a share of the 14- to 18-year-old population.
Sources: Annuario Statistico dell'Istruzione Italiana, 1958-1968, Istituto Nazionale di Statistica.

Figure A.2: Freshmen in STEM and Restricted Majors

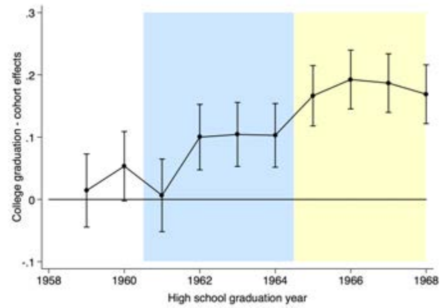


Notes: Freshmen enrolled in STEM and restricted majors as a share of the 19-year-old population.
Sources: Annuario Statistico dell'Istruzione Italiana, 1958-1968, Istituto Nazionale di Statistica.

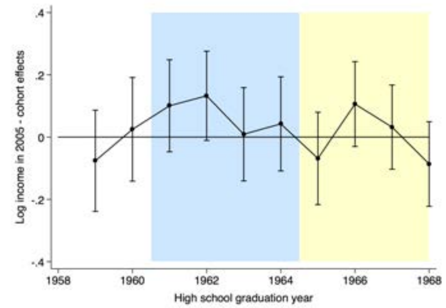
Figure A.3: Cohort Effects, Education and Income of Type B Students



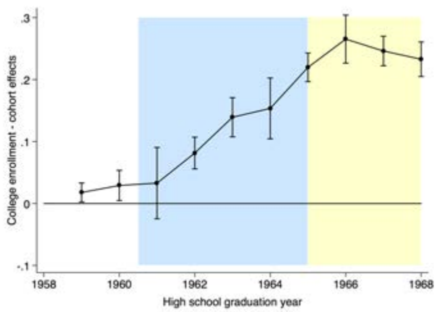
A. Enrollment - No pre-collegiate ability



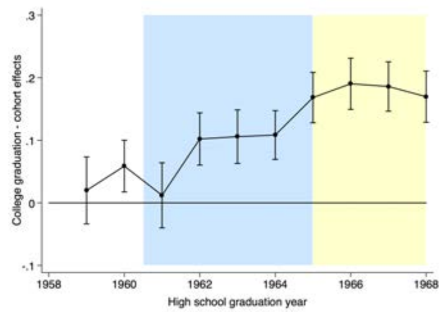
B. Graduation - No pre-collegiate ability



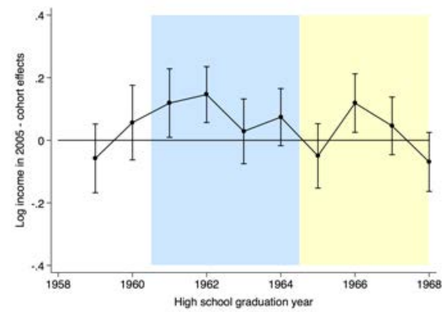
C. Log Income - No pre-collegiate ability



D. Enrollment - Clustered SEs



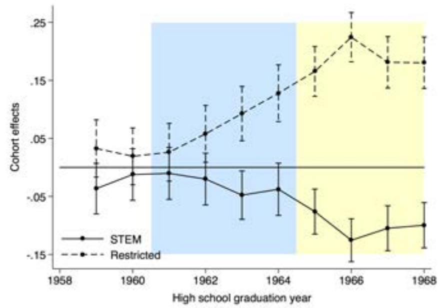
E. Graduation - Clustered SEs



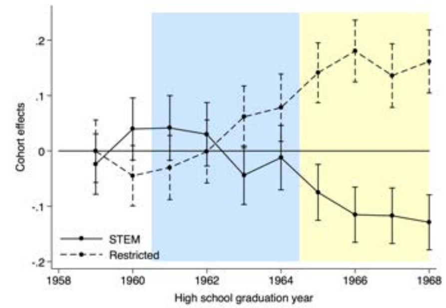
F. Log Income - Clustered SEs

Notes: Robustness checks of the marginal cohort effects on education choices and income of type B students from equation (1). The bars represent 95 percent confidence intervals. The first row shows marginal cohort effects computed without including measures of pre-collegiate ability on college enrollment (Panel A), on college graduation (Panel B), and on log income in 2005 (Panel C). The second row shows marginal cohort effects with standard errors clustered by high school and graduation year on college enrollment (Panel D), on college graduation (Panel E), and on log income in 2005 (Panel F).
Sources: School data of type B students that completed high school in Milan, Italy; 1958-1968.

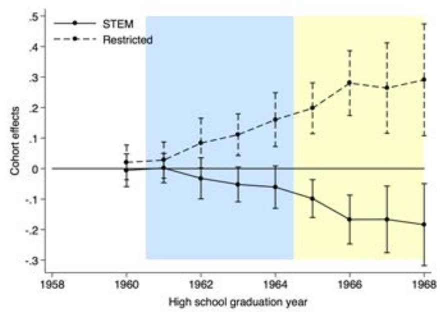
Figure A.4: Multinomial Logit, Major Choice of Type A Students



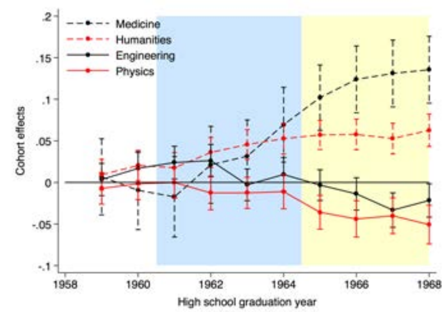
A. No pre-collegiate ability



B. Males



C. Conditional Multinomial Logit

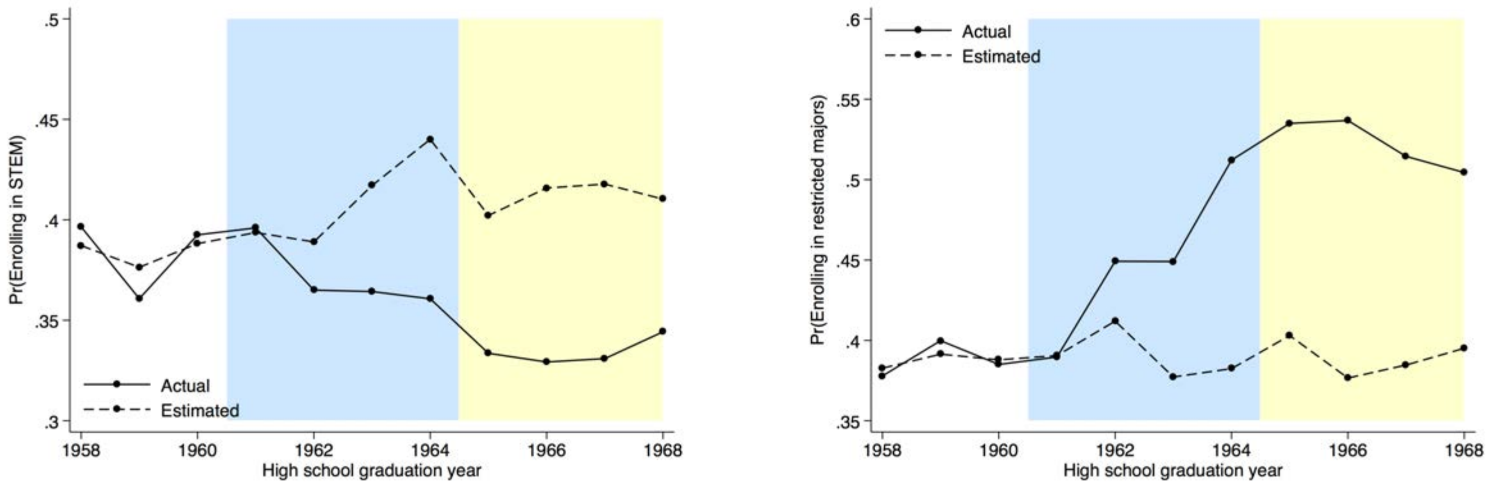


D. Disaggregated Major Choice

Notes: Robustness checks of the marginal cohort effects on the major choice of type A students from equation (13). The bars represent 95 percent confidence intervals. Panel A shows marginal cohort effects without controlling for measures of pre-collegiate 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 that completed high school in Milan, Italy; 1958-1968.

Figure A.5: Predicted and Actual Probability of Type A Enrollment

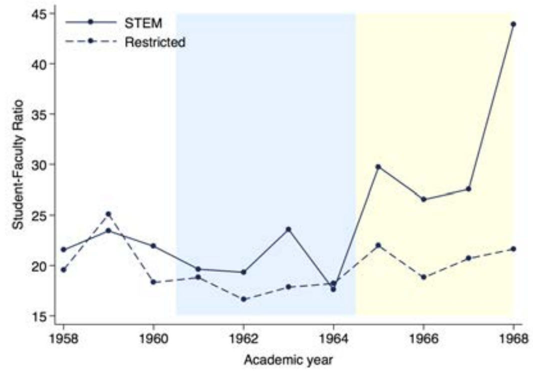


A. STEM Majors

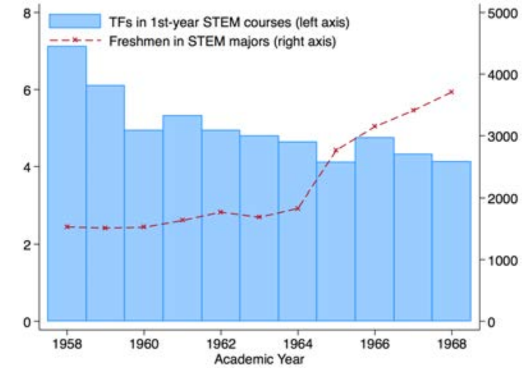
B. Restricted Majors

Notes: Actual and predicted probabilities of type A students enrolling in STEM and restricted majors. The predicted probabilities are constructed estimating the equation (13), using only the cohorts that completed high school until 1961.
Sources: School data of type A students that completed high school in Milan, Italy; 1958-1968.

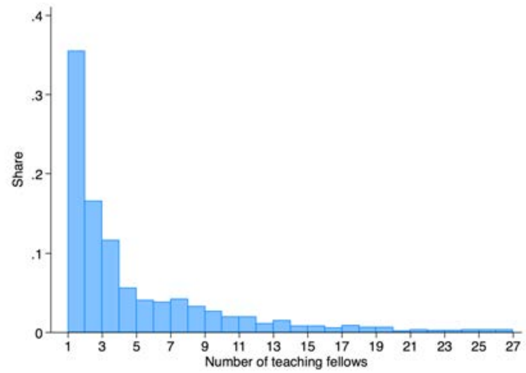
Figure A.6: Trends of Students-Faculty Ratios



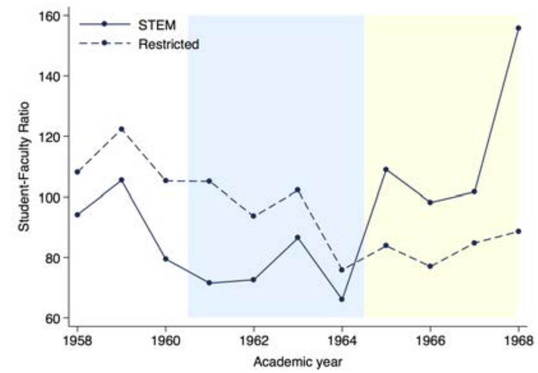
A. Mean Student-Faculty Ratio of First-Year Courses



B. Students Enrolled and Number of Teaching Fellows



C. Distribution of Teaching Fellows in Compulsory STEM Courses



D. Mean Augmented Student-Faculty Ratio

Notes: The student-faculty ratio of course c in academic year t is the number of enrolled students that completed high school in Milan between 1958 and 1968 over the number of teaching fellows. The teaching fellows is the sum of the professor and the teaching assistants assigned to each course.
Sources: Annals of Università Statale di Milano, Politecnico di Milano, and Università Cattolica del Sacro Cuore; 1958-1968. College transcripts of students that completed high school in Milan, Italy; 1958-1968.

B Data Collection

The data collection targeted the population of high school students that 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 B.2, 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 the college transcripts for the same group of students from the archives of the three local universities (Appendix Figure B.2, 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 college 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 college 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 college 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 much 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 Ministry of the 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 data. For this reason, 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 college 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 college courses at Università Cattolica.

Tables and Figures

Table B.1: Test of Means, Students With and Without an Income Observation

	Matched (1)	Not Matched (2)	Difference (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	

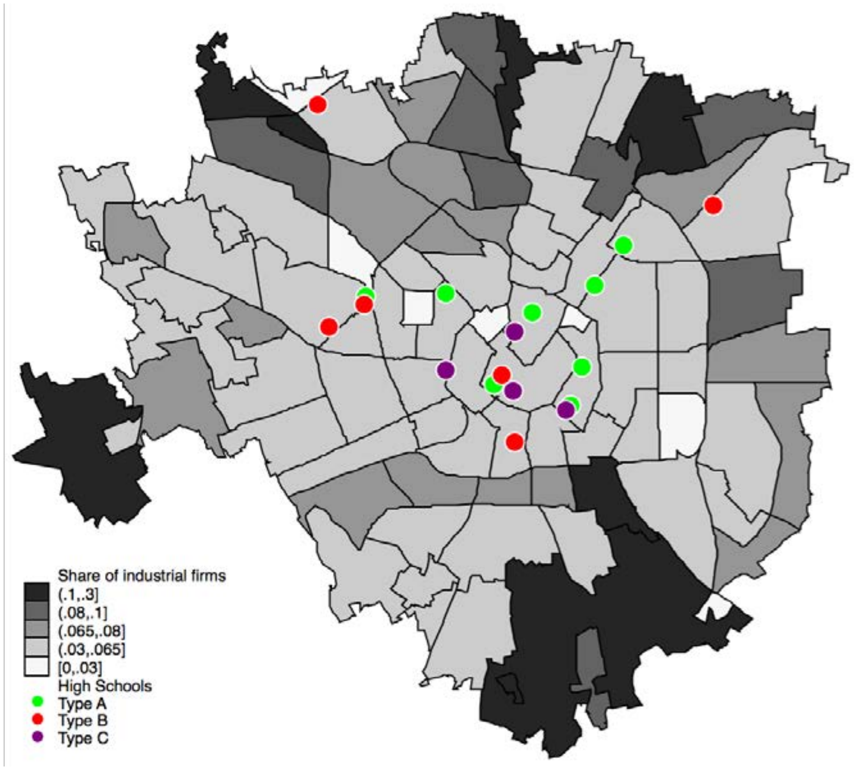
Notes: *Matched* are the 22,579 students that are matched with an income earner from the complete list of personal income tax returns in 2005. Similarly, *Not Matched* are the remaining 4,657 students that did not have a correspondence in the list of income tax returns. *** p<0.01, ** p<0.05, * p<0.1.

Table B.2: OLS, Test for Selective Attrition in Income

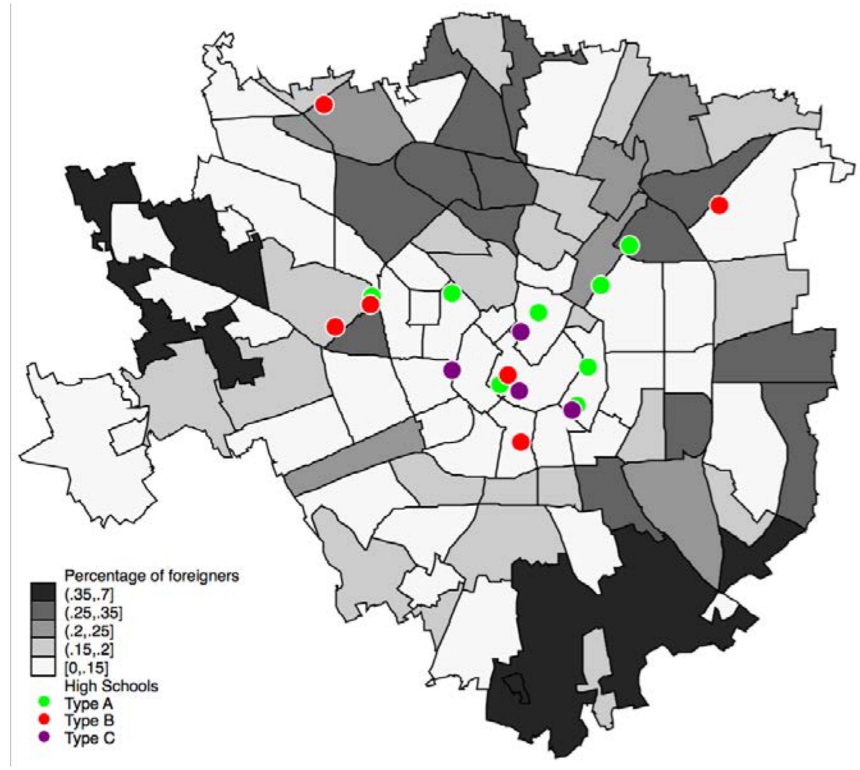
	Matched (1)	Matched (2)
F-test on student controls	19.43	14.80
	(<0.001)	(<0.001)
F-test on cohort FEs	5.87	0.074
	(<0.001)	(0.689)
F-test on high school FEs	2.38	2.32
	(0.001)	(0.002)
F-test on cohort FEs x Male	Not included	0.91
		(0.527)
F-test on cohort FEs x HS score	Not included	0.16
		(0.999)
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 B.1: Location of High Schools in Milan



A. Share of Industrial Firms (2011)

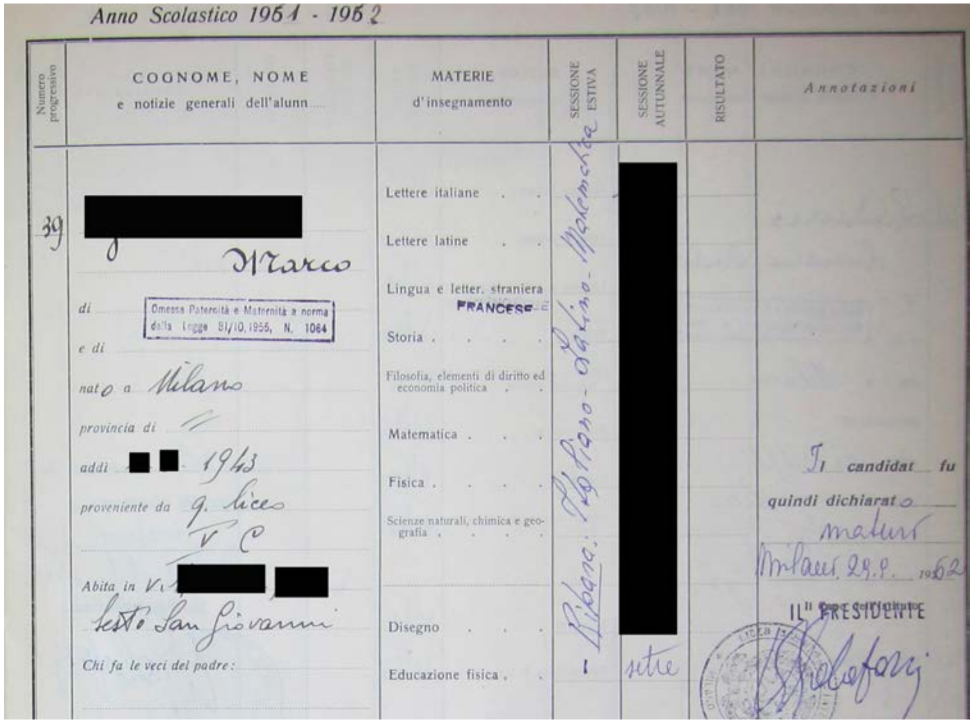


B. Share of Foreigners (2011)

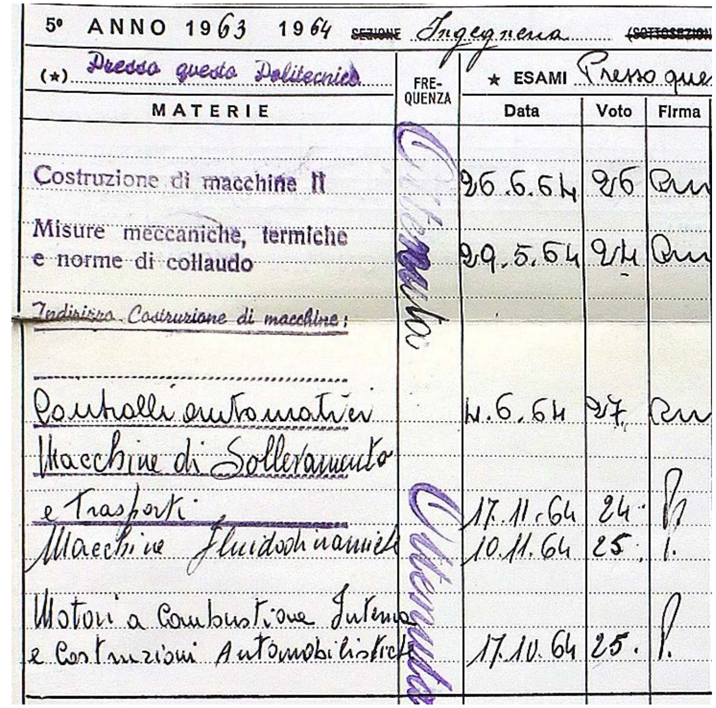
Notes: Map of Milan, Italy. The boundaries identify the 88 neighborhoods (*Nuclei di Identità Locale*). The share of industrial firms is the percentage of industrial firms out of all economic activities in the neighborhood (Panel A). The percentage of foreigners is the share of non-nationals over total 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 B.2: Examples of High School and College Data



A. High School Data



B. College 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 college transcript. Specifically, it shows courses attended during the fifth year of engineering with exam dates and outcomes for one college student.

C 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 that 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 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 C.1 (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 C.1 (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 C.1 (Panel B) represents this problem graphically. So, 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. These representative dataset of the Italian population contain 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 C.1 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.[54] 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 college degree and earned lower incomes. These differences could be due to the fact that Milan has higher returns to college 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 college 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):

[54] This is the subsample that is closer (both in characteristics and in time) to the group of income earners in 2005 that completed high school in Milan between 1958 and 1968. The estimating sample, however, is bigger (it includes 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 C.1.

$$p_{abt} = F(\text{age}_{abt}\,,\ \text{age}^2_{abt}\,,\ \text{age}^3_{abt}\,,\ \text{male}_{abt}\,,\ \text{college}_{abt}\,,\ \text{age}_{abt} \cdot \text{male}_{abt}\,,\ \text{age}_{abt} \cdot \text{college}_{abt}\,,\ u_t\,,\ B_b) \tag{C.1}$$

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 C.2). Male individuals with a college 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 college 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 that are not retired and below 65 years old, I estimate the following income equation

$$y_{abt} = \alpha + \beta \text{age}_{abt} + \sum_b \gamma_b B_b + \delta u_t + \epsilon_{abt} \tag{C.2}$$

separately for men (1) with and (2) without a college degree, and women (3) with and (4) without a college 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 college graduates and by €524 for female college 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:

$$\tilde{y}_b = \hat{\pi}_b \cdot \left[\frac{\hat{\%}_b \cdot y_b}{\hat{\text{replace}}_b} + (1 - \hat{\%}_b) \cdot y_b \right] + (1 - \hat{\pi}_b) \cdot \left[y_b + \hat{\beta} \text{age}_b \right] \tag{C.3}$$

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 in 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 (C.2) are above the actual means, but the difference is not significant for men with and without a college degree. For women with a college 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 percentile to limit the influence of outliers on the analysis. Figure C.3 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 IV estimator of β_1 in equation (2), which measures the returns to a STEM education for type B students that enrolled after 1961 (Table C.4). For each set of assumptions, I also plot the average adjusted income (Figure C.4). This exercise shows that the different assumptions lead to results that are qualitatively similar.

A. Period effects do not exist

I estimate equations (C.1) and (C.2) 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 IV estimator of β_1 is very close to the baseline. It is negative (-0.258) but not statistically significant.

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 (C.1) and (C.2) replacing the unemployment rate with these set of modified survey year dummies. The average adjusted income follows the same path. The IV estimator of β_1 is positive (0.079), but not statistically significant. The point estimate lies in the 95 percent confidence interval of the baseline estimate.

C. Quadratic age effects on income

In this case, I estimate equation (C.2) 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 (C.2) 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 IV estimator of β_1 is smaller, compared with the baseline. The point estimate is negative (-1.940) and statistically significant.

D. Age effects change with cohorts

I estimate equations (C.1) and (C.2) 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 IV estimator of β_1 is negative (-0.886) and statistically significant.

Tables and figures

Table C.1: 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
College graduates	0.25	0.35	-0.10***	0.24
College 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 - college graduates	42,038	43,403	-1,365	41,584
Income - male college 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 C. 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 C.2: Marginal Age Effect on Income

	College 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 (C.2), separately estimated on men (1) with and (2) without a college degree, and women (3) with and (4) without a college degree. *** p<0.01, ** p<0.05, * p<0.1.
Sources: Bank of Italy's SHIW; 1991-2012 waves.

Table C.3: Income Adjustment, Actual and Predicted Age Effects

	Actual	Predicted	Difference
	(1)	(2)	(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 college degree	43,892	46,078	-2,186
Male with HS diploma	28,694	28,885	-191
Female with college degree	21,556	24,651	-3,095***
Female with HS diploma	17,114	17,948	-834***

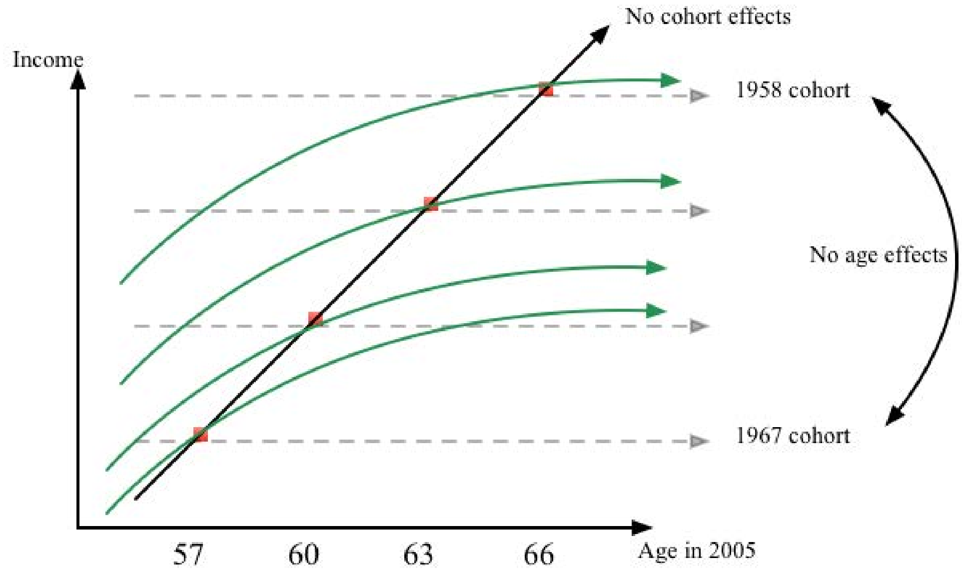
Notes: For a random subsample (25 percent) of the SHIW dataset, the table compares the actual means to those predicted with the coefficients estimated using the remaining 75 percent of the sample. The probability of being retired, the ratio of pensions to total income, and the replacement rate are estimated using equation (C.1). Income is estimated using equation (C.2). *** p<0.01, ** p<0.05, * p<0.1.
Sources: Bank of Italy's SHIW; 1991-2012 waves.

Table C.4: Sensitivity Analysis of Assumptions on Age-Cohort-Period Effects

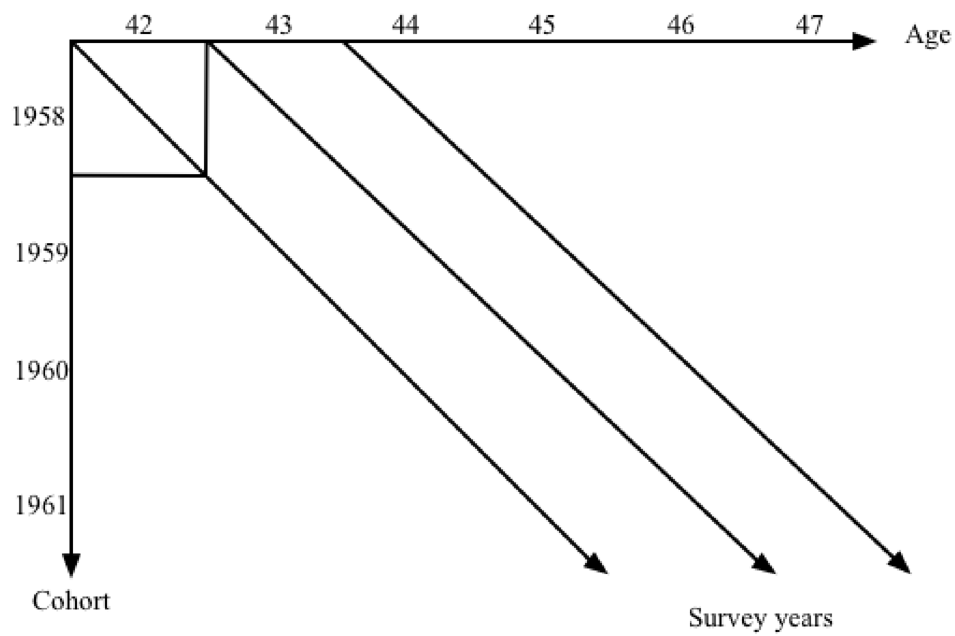
Baseline	Not Adjusted
-0.238	-1.208***
(0.238)	(0.457)
A. No Period Effects	B. Normalization
-0.258	0.079
(0.237)	(0.230)
C. Quadratic Age Effects	D. Changing Age Effects
-1.940***	-0.886***
(0.472)	(0.198)

Notes: The table shows the IV estimator of β_1 in equation (2), which measures the returns to a STEM education for type B students that enrolled after 1961. Different estimators use different assumptions about the structure of age - cohort - period effects, as described in section C. *** $p<0.01$, ** $p<0.05$, * $p<0.1$.
Sources: School data of type B students that completed high school in Milan, Italy; 1958-1968.

Figure C.1: Identification of Age and Cohort Effects



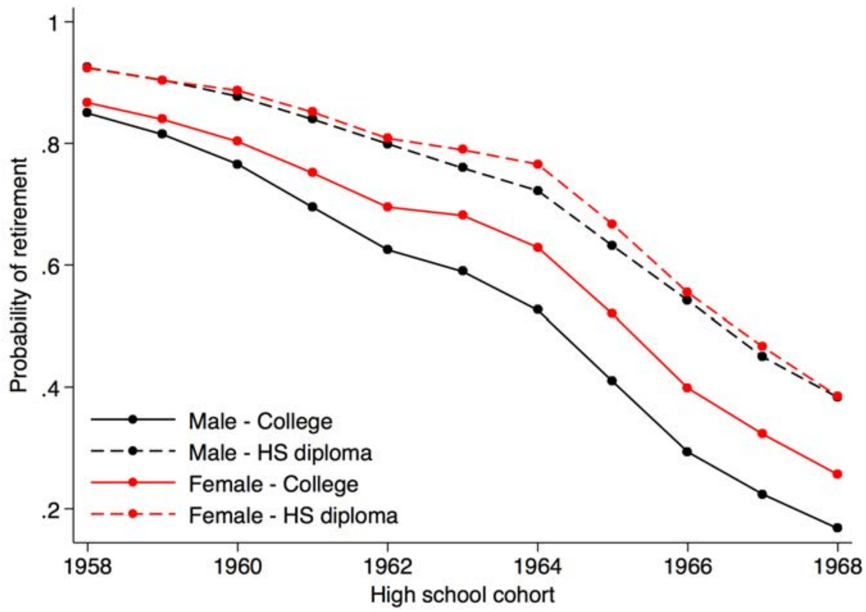
A. Age and Cohort Effects in One Cross-Section



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

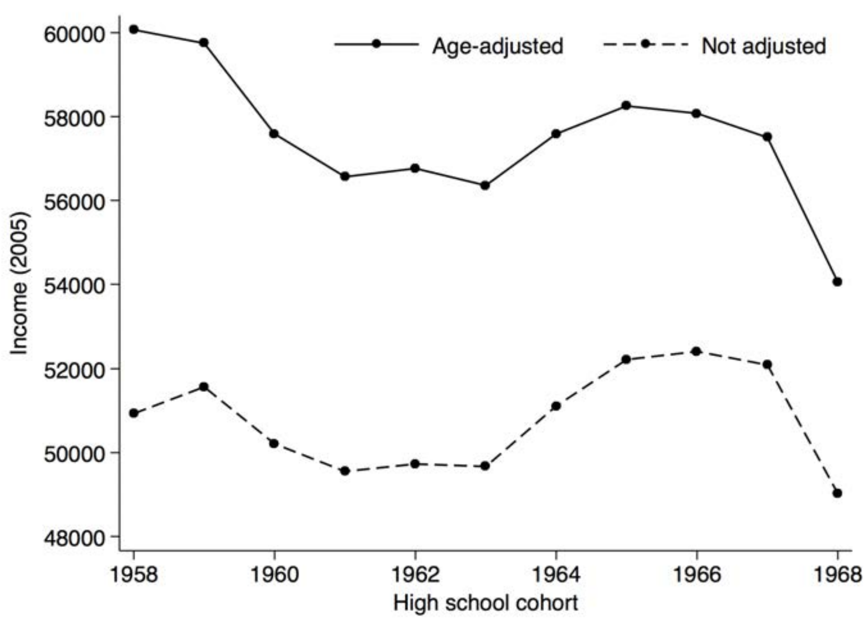
Notes: Panel A shows how age and cohort effects are confounded in a single cross-section. Panel B shows that in repeated cross-sections there is an additional dimension (period or survey year), which is a linear combination of cohort and age. Therefore, the simultaneous estimation of all three is not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

Figure C.2: Estimated Probability of Retirement



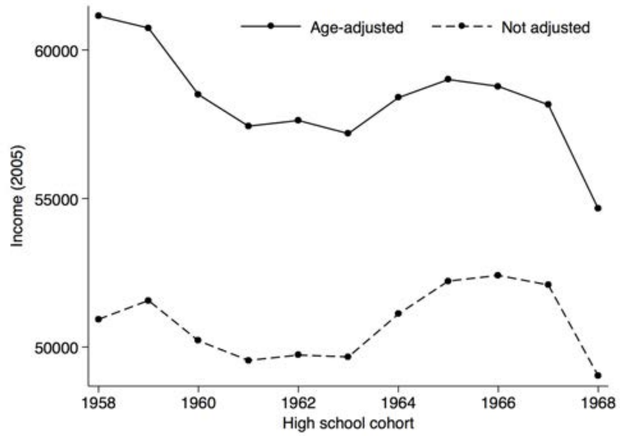
Notes: Estimated probability of retirement by completed education and gender. Probability of retirement is estimated by equation (C.1).
Sources: Bank of Italy's SHIW; 1991-2012 waves.

Figure C.3: Age-adjusted vs. Unadjusted Income

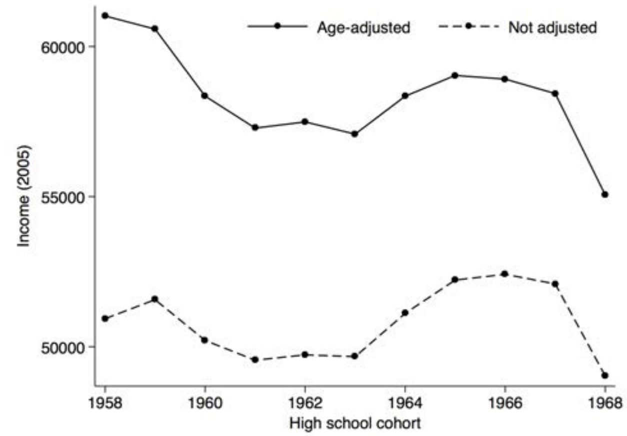


Notes: The figure shows means by high school cohorts. Income is adjusted to account for age effects using equation (C.3).
Sources: Bank of Italy's SHIW; 1991-2012 waves.

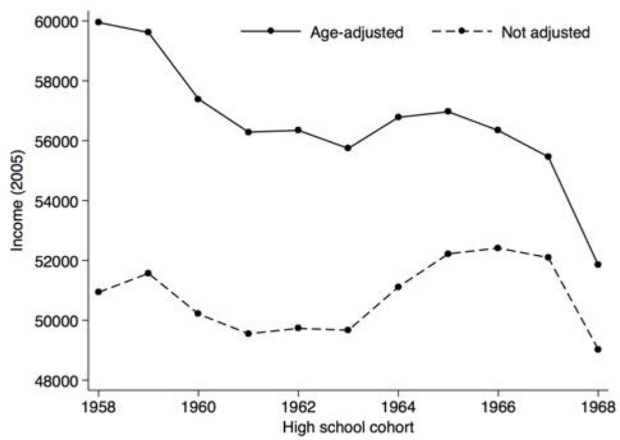
Figure C.4: Age-adjustment: Different Assumptions on Age-Cohort-Period Effects



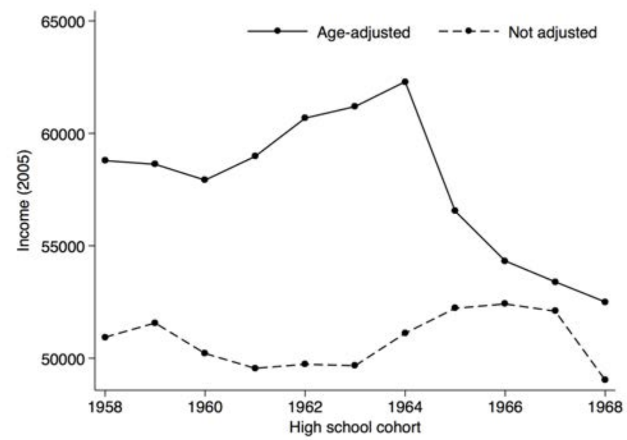
A. No Period Effects



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 C.

D Parental Characteristics of High School Graduates

The administrative data collected in school and college archives do not contain any information about the family background. To run some additional tests on the changing characteristics of high school graduates, I use out-of-sample observations from the Bank of Italy's Survey of Household Income and Wealth (SHIW). Differently from what done in Appendix C, I restrict the sample to the four more recent waves of the SHIW (2006, 2008, 2010, 2012), because previous waves contain only information about the highest level of completed education and the high school diploma of college graduates could not be observed.

Tables

Table D.1: Parents' Characteristics and High School Choice

	Father			Mother		
	Type A (1)	Technical (2)	Difference (3)	Type A (4)	Technical (5)	Difference (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.

Table D.2: Change of Father's Characteristics

	Technical students			Type A		
	Pre 1965 (1)	Post 1965 (2)	Post-Pre (3)	Pre 1965 (4)	Post 1965 (5)	Post-Pre (6)
Education						
No education	0.075	0.074	-0.001	0.067	0.029	-0.038
Low education	0.727	0.791	0.064	0.507	0.592	0.085
High education	0.198	0.135	-0.063*	0.425	0.379	-0.046
Occupation						
Low income	0.411	0.403	-0.008	0.301	0.209	-0.092
Middle income	0.448	0.485	0.037	0.414	0.499	0.085
High income	0.141	0.112	-0.029	0.285	0.292	0.007
Sector						
Public servant	0.242	0.216	-0.026	0.270	0.296	0.026

Notes: See table D.1, *** p<0.01, ** p<0.05, * p<0.1.
Sources: Bank of Italy's SHIW; 2006, 2008, 2010, 2012 waves.

Table D.3: Change of Mother's Characteristics

	Technical students			Type A		
	Pre 1965 (1)	Post 1965 (2)	Post-Pre (3)	Pre 1965 (4)	Post 1965 (5)	Post-Pre (6)
Education						
No education	0.130	0.119	-0.011	0.105	0.065	-0.040
Low education	0.773	0.856	0.083**	0.658	0.663	0.005
High education	0.097	0.026	-0.071***	0.237	0.272	0.035
Occupation						
Low income	0.860	0.841	-0.019	0.810	0.857	0.047
Middle income	0.132	0.153	0.021	0.153	0.135	-0.018
High income	0.008	0.006	-0.002	0.037	0.008	-0.029
Sector						
Public servant	0.145	0.114	-0.031	0.478	0.501	0.023

Notes: See table D.1, *** p<0.01, ** p<0.05, * p<0.1.
Sources: Bank of Italy's SHIW; 2006, 2008, 2010, 2012 waves.

E Quality of Education

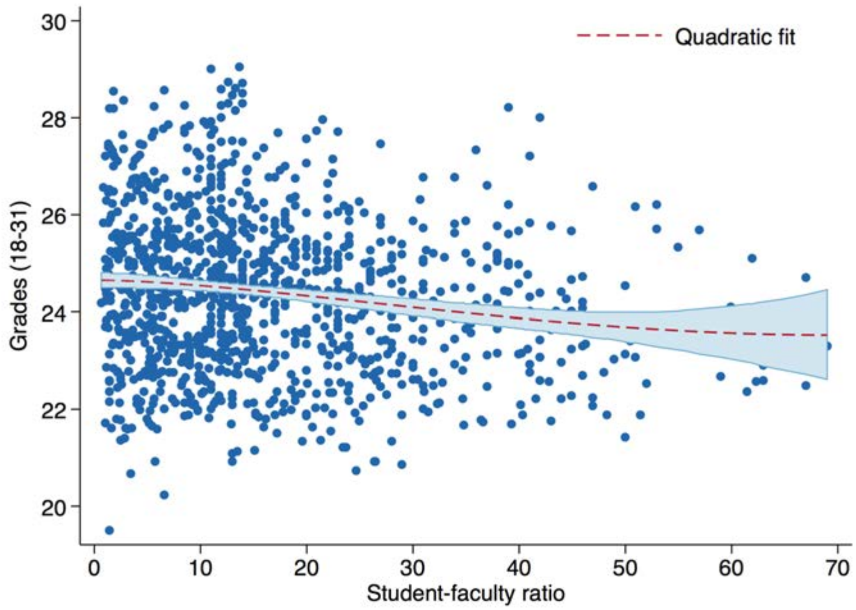
In section V.B, I assumed that quality of education in each college course is a function of the student-faculty ratio ($Q_c = f\left(\frac{E_c}{fac_c}\right)$). The function f is decreasing ($f' < 0$) and either linear and concave ($f'' \leq 0$). This function implies that the decrease in quality that follows a marginal increase in enrollment is larger in courses with a higher pre-existing student faculty ratio. This assumption is consistent with contexts in which congestion costs increase as quality decreases. It is, however, not consistent with the existence of a steep quality decrease at low levels of the student-faculty ratio. The empirical relationship between quality of education (average grade) and student-faculty ratio in STEM compulsory courses is fairly linear and decreasing (Figure E.1). Importantly, it does not show evidence of a steep quality decrease at low levels of the student-faculty ratio.

The entry of type B students and a fixed number of professors and assistants (this element exacerbates the effect of the policy, but is not necessary) imply that the quality of education decreases more in courses with a higher pre-existing student-faculty ratio. Let's consider two compulsory courses in the same major. In 1960, course I has a student-faculty ratio of $\frac{20}{1}$. After an enrollment increase of 20, the student-faculty ratio becomes $\frac{40}{1}$. In a different course in the same major, course II, the student-faculty ratio in 1960 is equal to $\frac{20}{2}$. After the same enrollment increase (the courses are in the same major), the student-faculty ratio increases by 10 to $\frac{40}{2}$. As a result, quality decreases more in course I than in course II.

The assumption on function f is consistent with a scenario in which students compete to access fixed school inputs and courses with low pre-existing student-faculty ratio have unused capacity. Let's consider the same two courses and assume that professors have two weekly tasks: (1) preparing and giving lectures, and (2) receiving students during office hours. Professors have a fixed amount of time every week (20 units): lectures take 10 units of time and the remaining units can be allocated to office hours. Assistants have 20 units of time for office hours, in case the professor cannot accommodate all students. Finally, all students want to attend office hours and each student takes one unit of time. Before the enrollment increase, 10 students could not attend office hours in course I (the professor had only 10 units for 20 students) and 0 in course II (the professor had 10 units and the assistant provided the additional 10). After the enrollment increase, 30 students could not attend office hours in course I and 10 in course II. In the latter course, there was unused capacity which allowed to accommodate more students after the enrollment increase. As a result, quality of education decreased more in course I. This is only a practical example of how the assumption on f can be rationalized.

Tables and Figures

Figure E.1: Empirical Relationship Between Quality of Education and Student-Faculty Ratio



Notes: Each point is a course-academic year combination. Sample includes all compulsory courses in STEM majors with more than 10 students enrolled.
Sources: Annals of Università Statale di Milano, Politecnico di Milano, and Università Cattolica del Sacro Cuore; 1958-1968. College transcripts of students that completed high school in Milan, Italy; 1958-1968.