DISCUSSION PAPER SERIES

IZA DP No. 12691

The Pathways to College

Lisha Agarwal
Giorgio Brunello
Lorenzo Rocco

OCTOBER 2019

IZA – Institute of Labor Economics

Initiated by Deutsche Post Foundation

DISCUSSION PAPER SERIES

IZA DP No. 12691

The Pathways to College

Lisha Agarwal
University of Padova

Giorgio Brunello
University of Padova and IZA

Lorenzo Rocco
University of Padova and IZA

OCTOBER 2019

ISSN: 2365-9793

IZA – Institute of Labor Economics

Schaumburg-Lippe-Straße 5–9
53113 Bonn, Germany

Phone: +49-228-3894-0
Email: publications@iza.org

www.iza.org

ABSTRACT

The Pathways to College*

We estimate the effect of the high school curriculum (or track) on the returns to college using data from the Italian PLUS (Participation Labour and Unemployment Survey) survey. We find that college graduates with vocational high school are less likely to be employed than graduates with academic high school. When employed, they earn 7.3 percent less per hour but work 3.8 percent more hours per week. They are less likely to fill high ranked occupations and more likely to find their first job quickly after school completion than other graduates. The wage penalty associated to vocational education in high school is larger for females than for males and for those born in the less economically developed Southern regions.

JEL Classification: J24

Keywords: high school curriculum, returns to college, Italy

Corresponding author:
Giorgio Brunello
Department of Economics and Management
University of Padova
via del Santo 33
35123 Padova
Italy

E-mail: giorgio.brunello@unipd.it

* We are grateful to Carmen Aina, Marco Bertoni, Eric Bettinger, Simona Comi and to the audiences in Novara (AIEL), LSE (CVER) and Padova for comments and suggestions. The usual disclaimer applies.

1. Introduction

The earnings of college graduates vary considerably. In 2016, for instance, the coefficient of variation of hourly earnings for college graduates aged 30 to 40 and working full time was close to 50 percent of the mean in the US and close to 47, 42 and 64 percent of the mean in Germany, France and Italy.[1]

A considerable body of literature has investigated the many factors driving this variation, including individual ability, the quality of institutions and the college major (see for instance Altonji, Blom and Meghir, 2012; Lemieux, 2006; Autor et al, 2008; Arcidiacono, 2004; Deming et al, 2014; Chevalier, 2011; Lindley and McIntosh, 2015 and Kirkeboen et al, 2016).

One factor that has not been sufficiently explored so far is the educational pathway or bundle of qualifications attained before entering and completing college.[2] Most studies of the returns to college have focused only on the highest attained qualification. Yet, since skill formation is characterized by dynamic complementarities (see Cunha and Heckman, 2007), these returns are likely to vary with the set of skills accumulated before entering college.

In several countries, education systems offer alternative pathways to college education. Both in Europe and in Asia, high school curricula are organized in academic and vocational tracks, with the differentiation occurring at different ages but before college entry (see for instance Brunello and Checchi, 2007). Although some tracks do not lead to

[1] The coefficients of variation are computed using the US Current Population Survey and the European Union Statistics on Income and Living Conditions (EU-SILC).
[2] Some contributions in this literature are Altonji, 1995; Levine and Zimmermann, 1995; Rose and Betts, 2004 and Joensen and Nielsen, 2008.

college, in several countries access to college is open to graduates with both an academic or vocational curriculum.[3]

Understanding the importance of high school qualifications for the returns to college is often difficult due to the lack of data. Labour force surveys, both in Europe and in the US, collect information on the highest attained degree but are silent on the educational pathways leading to college. The Survey on Adult Skills (PIAAC), a key survey designed to understand skill formation and use, provides data on pathways only for respondents who are currently in education at the time of the survey (see Allen, Massing, Schneider and van der Velden, 2017).

In Italy, access to college has been progressively liberalized since the 1960s (see Bianchi, 2018). In 2016, the share of engineering graduates who had completed an academic and a vocational high school was equal to 64.6 and 35.4 percent respectively. For law graduates, the difference was much larger (76.7 versus 23.2 percent). In the same year, engineers with an academic high school degree earned on average 37,247 euro (gross), significantly more than engineers with a vocational high school degree (29,854 euro). In a similar fashion, law graduates with an academic high school degree earned more than graduates with a vocational high school degree (29,507 versus 25972 euro).[4]

Is this gap driven by the fact that college graduates with the same major but different high school education belong to different ability groups, or

[3] High school in France includes both a general and a technological path leading to a baccalaureate and allowing access to college. In the UK, access to college is open both to A levels and to BTEC, a vocational track. In Italy, students completing both an academic-oriented *Liceo* and a vocationally-oriented *Istituto Tecnico* can access university.

[4] These data are drawn from the Italian Participation Labour Unemployment Survey (PLUS).

should we instead consider the possibility that there are different degrees of complementarity between high school curricula and college education? In this paper, we address this question and contribute both to the relatively small literature investigating the effects of the school curriculum on college returns and to the literature – mainly European – that explores the economic consequences of secondary school tracks (see for instance Dustmann et al, 2017; Biewen and Tapalaga, 2016, and Brunello and Rocco, 2017)

Our empirical analysis relies on data drawn from the Italian PLUS (*Participation Labour and Unemployment Survey*) survey, which provides information both on the highest completed education and on the intermediary degrees attained between compulsory education and the highest degree. We focus on the economic returns to college, which we measure with the probability of employment, hourly earnings and weekly hours worked, occupation, type of contract, time to the first job and the probability of receiving training.

We estimate the effect of the high school curriculum (or track) on the returns to college using selection on observables and a rich set of individual controls, which include measures of cognitive skills and individual attitudes before the start of high school. We show that these estimates are both robust to omitted un-observables - using the test devised by Oster (see Oster, 2016) - and to the use of alternative estimation methods, including the inverse-probability weighted regression adjusted estimator (IPWRA) (see Cattaneo, 2010) and entropy balancing.

We find that college graduates with vocational high school are less likely to be employed than graduates with academic high school. When employed, they earn 7.3 percent less per hour but work 3.8 percent more hours per week. They are less likely to fill high ranked occupations and more likely to find their first job quickly after school completion than other graduates. The wage penalty associated to vocational education in high school is larger for females than for males and for those born in the less economically developed Southern regions.

Assuming that wages are proportional to productivity, and that selection into treatment has been adequately controlled for, our results suggest that the vocational curriculum in Italian high school is a weaker complement to college education than the academic curriculum, and that the latter is a better pathway to college in terms of expected economic returns. Although it is difficult to say whether they hold in different economic and educational contexts, they also have implications for the current tendency of making vocational education more academic (or "academic drift" – see Green et al, 1999): increasing the academic content of vocational education may payoff in the labour market by increasing its complementarity with college education.

The paper is organized as follows. Section 2 provides a brief description of the Italian education system and Section 3 introduces the data. The empirical approach is discussed in Section 4 and results are presented in Section 5. Conclusions follow.

2. Upper Secondary Education in Italy

In Italy, upper secondary education lasts between three and five years, typically starting at age 14 upon completion of junior high school, and is organized in academic and vocational curricula or tracks. The vocational track comprises three and five-year high schools with a predominant technical training (*scuole professionali, istituti tecnici e commerciali*), and the academic track consists of four or five-year high schools with a more general education, which focuses on classical, scientific or linguistic and pedagogical studies (*licei* and *scuole magistrali*). The two tracks differ both in their orientation and in the learning objectives. In addition to specialising in their particular fields (humanities, the arts or science), the *licei* also include subjects such as maths, chemistry, physics and biology, history, geography and Italian language and literature. Vocational schools are more geared instead towards technical and practical subjects such as technology, informatics, engineering, construction and accounting, and focus on developing industrial and administrative skills.

Access to tracks is based on individual and/or parental choice (see Checchi and Flabbi, 2007, for a comparison with Germany). In our sample, 61.8 percent of the individuals who have completed at least upper secondary education have graduated from a vocational school. Typically, students from vocational high schools are much less likely to complete college than those from an academic track (13.1 versus 53.7 percent in 2016). While vocational high school graduates have lower final scores than academic high school graduates (77.1/100 versus 79.9/100, where 100 if the maximum score), those enrolling in college have similar scores (83.3/100 versus 83.3/100).

3. The Data

We draw our data from the Italian PLUS (Participation Labour Unemployment Survey) survey, run by INAPP (National Institute for the Study of Public Policies) between 2005 and 2016.[5] This survey, based on a stratified nationally representative sample of more than 50 thousand individuals aged 18 to 74, has two key advantages with respect to more standard sources such as the Labour Force Survey.

First, it includes information both on the highest and on intermediate qualifications, which allows us to trace the education pathways from the end of junior high school (at age 14) to college. Second, it contains relevant information on school evaluations at the end of junior high school, when students take a national school leaving exam, as well as data on individual attitudes at age 13 and detailed information on parental education, occupation and the region were the individual grew up and went to school before college.

Since some of this information is only available in the recent waves, we focus on the years 2010, 2011, 2014 and 2016. Our working sample consists of 78,358 individuals aged 23 to 59, who are not full-time students, were born in Italy between 1947 and 1991 and have completed at least junior secondary education.

Average labour market outcomes and characteristics of college graduates by type of upper secondary education are shown in Table 1. The table consists of three columns, the first for college graduates with vocational high school, the second for college graduates with academic high school and the third for individuals with at least lower secondary education who have not completed college and who might have completed vocational or academic high school education.

[5] The survey was implemented in 2005, 2006, 2008, 2010, 2011, 2014 and 2016.

College graduates with a vocational education have lower average hourly wages and work longer average weekly working hours than those with an academic education; they have similar employment probabilities and require on average less time to find their first job; they are about as likely to have completed a STEM major and to have graduated from a high ranked university; they are less likely to be females, to have completed junior high education with top grades and typically have a less privileged parental background; their attitudes at age 13 were mainly for sports, compared to sports and literature & arts for those with an academic education.

We classify the educational choices made from age 14 – at the end of lower secondary education – onwards into three potential treatments: 1) NC: a binary variable indicating either no completion of high school or college, or attainment of vocational or academic high school but no college degree (82.5 percent of the sample); 2) V: a binary variable for vocational high school plus college completion (4.9 percent); 3) G: a binary variable for academic high school plus college completion (12.6 percent). This classification assigns individuals with some college but no college degree and with some high school but no degree to group NC. An alternative classification – that we consider as a robustness exercise – breaks down group one into three groups, one for those with no high school degree, and the remaining two for those with a vocational or an academic high school degree as their highest degree.

We study the effects of high school curriculum on the following labour market outcomes for college graduates: log hourly wages; log weekly hours worked; the employment probability; the probability of having received training after completing school and during the past three years; the time interval between finishing school and starting the first job, measured as a binary variable equal to 1 if the time is at or above the median and to 0 otherwise; the occupation, measured as a binary variable equal to 1 if the reported occupation belongs to one of the three top three categories in the ISCO classification (managers, professionals, technicians and associate professionals), and to 0 otherwise; the type of contract, a binary variable equal to 1 if the contract is open ended and to 0 otherwise.

Due to the poor reliability of wage data for the self-employed, we consider two alternative measures of earnings and hours, one that excludes this group from the sample and the other that includes it. We compute log hourly wages and log weekly hours using the approximation $lnx = \ln(1 + x)$, and consider either the employed only or both employed and not employed, by assigning x=0 to the latter.

4. The Empirical Approach

It is well known that the comparison of average outcomes between college graduates with vocational and academic curricula is not informative of the difference in economic returns because of selection bias: students with a vocational curriculum are less likely to enrol in college than students with an academic curriculum, and may have different abilities.

Our identification strategy is selection on observables and relies on the "conditional independence assumption" (CIA), stating that each treatment T=t – t=V (college plus vocational high school), t=G (college plus academic high school) and t=NC (no college) - is as good as randomly assigned conditional on the vector of controls X.

This assumption requires that the key determinants of educational choice are captured by X, so that any residual variation in education is either random or due to factors that do not influence the outcomes of interest. Formally, this implies that $\{Y(V), Y(G), Y(NC)\} \perp T|X$.

We believe that our data is sufficiently rich to support CIA. Apart from the standard information on parental education and occupation, we have three variables that help us capture differences in ability and attitudes before treatment: 1) the final score in the compulsory exit exam taken at the end of junior high school, a qualitative indicator with four possible outcomes (excellent, very good, good and sufficient). We take into account the fact that grading may be area specific (more generous in the South than in the North) by interacting this indicator with a dummy for Southern regions; 2) self-reported area of personal interest (or attitude) at age 13, which includes music, sports, math and science, reading and the arts; 3) the region where the individual grew up before college age.

We recognize that selection into high school type is likely to depend not only on individual ability and parental background but also on peer effects and local demand and supply. We capture peer effects with the regional share of pupils enrolled in vocational high school at age 14, local demand effects with the local unemployment rate at age 14, and local supply with the number of college courses offered in the region when the individual was 19 (see Rizzica, 2013). Additional covariates in vector X include the presence of parents in the household at the end of lower secondary education, whether parents grew up in the country or abroad, a fourth order polynomial in year of birth, gender and year dummies.

The causal parameters (ATT: average treatment effects on the treated) are defined as $E(Y(t) - Y(t')|T = t)$ for t,t′ = V,G,NC. While there are many possible ATT, we are interested in this paper in $E(Y(V) - Y(G)|T = V)$, or the average effect of vocational education relative to academic education for individuals treated with t=V.

When CIA holds, ATT are obtained as a weighted averages of ATT(x) - the ATT conditional on X=x or $E(Y(t) - Y(t')|T = t, X = x)$, computed over all possible values x of vector X, using as weight the size of cells defined by X=x, conditional on T=t. The ATT(x) for the cell characterised by X=x can be computed only if it contains individuals with all possible treatments (common support or overlapping). In our empirical estimates, we always verify that overlap holds, and retain in the sample only the cells containing individuals with at least a one percent probability of being assigned to any possible treatment.[6]

We estimate ATT using the following regression

$$Y_{it} = X_{it}'\alpha + \beta_V V_i + \beta_{NC} NC_i + \varepsilon_{it} \quad (1)$$

where Y is the outcome; V is a binary variable equal to one for vocational high school plus college and to zero otherwise; NC is a binary variable equal to one for individuals without college and to zero otherwise; the treatment G is in the constant term and the indices i and t are for the individual and time.[7]

We verify whether the estimates obtained from (1) differ qualitatively from those based on two alternative approaches, that also rely on

[6] This probability is estimated parametrically using a multinomial logit model.
[7] The estimates based on regression (1) provides a close approximation of ATT, not the ATT. This approximation is closest to the true value when the empirical model is saturated with respect to X, which is unfeasible in most settings, including ours.

selection on observables, the inverse probability weighted regression adjusted method - briefly IPWRA – and entropy balancing, briefly EB.

IPWRA imputes to each individual the missing potential outcomes by exploiting information on individuals with similar characteristics who received alternative treatments, and consists of three steps. First, the inverse predicted probabilities of treatment are derived for all possible treatments using a multinomial logit model. Second, potential outcomes are predicted for all individuals and each treatment *t*. Predictions rely on the parameters of the outcome model, which is estimated by weighted least squares, with weights given by the normalised inverse probability of treatment. By weighting all groups with their inverse probability of treatment in the outcome model, IPWRA makes sure that the covariates are balanced across treatments. Finally, the ATT parameters are obtained by comparing the sample means of the predicted potential outcomes.[8]

Entropy balancing, proposed by Hainmueller and Xu, 2013, is based on a maximum entropy re-weighting scheme that fits weights satisfying a potentially large set of balance constraints. This method re-balances the covariates in groups T=G and T=NC to reproduce the distribution observed in group T=V. Once re-balancing has been attained, average treatment effects are estimated by standard regressions using re-weighted data.

[8] An attractive property of IPWRA is that it is doubly robust to misspecification. The method is a RA (regression-adjusted) estimator that uses inverse-probability weights obtained from the treatment model to correct estimates when the outcome model is incorrectly specified. If the outcome model is correctly specified, the weights do not affect the consistency of the estimator.

5. Results

5.1 Main findings

In our baseline specification, we focus on treatment V (college plus vocational high school) and two counterfactual treatments (G: college plus academic high school and NC: no college (with or without high school). We estimate Eq. (1) and report our estimates in Table 2 (for both the employed and the not employed) and Table 3 (for the employed only).

In the former table, we present the estimates of the binary variable V (or the effect of vocational high school relative to academic high school for college graduates) on the following outcomes: employment status, log hourly earnings and weekly hours by either including or excluding the self-employed, the probability of investing in training and the time to the first job. In the latter table, we consider as outcomes: log hourly earnings and weekly hours (either including or excluding the self-employed); a binary variable for high ranked occupations; the type of contract and training probability.

Tables A1-A3 in the Appendix show – only for employment and log hourly wages - how these estimates vary as we progressively add blocks of controls, starting with age and gender and adding local demand and supply factors and peer effects, junior high school final test scores and individual attitudes at age 13 and parental background variables.

When we consider both employed and not employed individuals (Table 2), we find that college graduates with vocational education have both a 2 percent lower probability of employment than graduates with academic high school and significantly lower hourly earnings (-11.1% when we include the self-employed and -12.2% when we exclude them).

They also work shorter weekly hours, mainly because of their lower probability of employment. On the one hand, having a vocational high school degree reduces the time required to find the first job. On the other hand, it also reduces the probability of investing in training.

When we focus on the employed only (Table 3), we find that college graduates with a vocational high school degree earn about 7 percent lower hourly earnings than those with an academic high school degree, work longer weekly hours, are less likely to fill a high ranked occupation and are almost as likely to be on an open ended contract and to receive training.[9]

Our identification strategy is likely to fail if the treatment is correlated with un-observables. We investigate this possibility by applying the test recently developed by Oster, 2016, which establishes bounds to the true value of the relevant parameters in two polar cases. In the first case, there are no un-observables and Equation (1) is correctly specified. We denote as $\hat{R}$ the estimated R squared. In the second case, there are un-observables but observables and un-observables are equally related to the treatment. When un-observables are included, we conservatively assume that the R squared is equal to $R_{max} = \min(1.3\hat{R}, 1)$. If zero can be excluded from the bounding set, accounting for un-observables does not change the direction of our estimates.

The bound associated to the presence of un-observables is shown in the last column of Tables 2 and 3. When we consider both the employed and the not employed, it always has the same sign and is close to estimated

[9] The effect of treatment V on weekly hours is negative when we consider both the employed and not employed and positive when we consider the employed only, reflecting the fact that college graduates with vocational high school are less likely to be employed and more likely to work zero hours.

β_V. This is also the case when we focus only on the employed, with the notable exception of training and high ranked occupation, suggesting that for these outcomes the estimated effect of treatment V can be spurious.

5.2 Results based on Entropy Balancing and IPWRA

We verify whether our results are robust to alternative estimation methods by using either entropy balancing or IPWRA. In the former case, we impose pre-estimation balancing on the first and second moments of the covariate distributions in the treatment and the re-weighted control group. Since our treatment is multi-valued, we proceed by first restricting the sample to groups with T=V and T=NC – therefore excluding the group with college and academic education – and second by focusing on the groups with T=V and T=G – excluding the group with no college. By so doing, we reproduce the distribution of covariates prevailing in group T=V in the other two groups.

Our results, presented in Table 4 both for the employed and not employed and only for the employed,[10] confirm the qualitative findings shown in Tables 2 and 3.

5.3 Heterogeneous Effects

We replicate our baseline estimates by distinguishing by gender, age and by region of birth (North and Centre versus South), and by focusing only on the employed. We find (see Table 5) that female college graduates with a vocational high school education have a larger negative earnings gap than males with respect to academic high school graduates (-7.9% versus -5.3%), a larger negative training gap (-4.1% versus 1.5%), a larger

[10] In this and the next tables, including those in the Appendix, we do not report the estimates for hourly wages and weekly hours that include the self-employed.

positive gap in weekly hours (7.6% versus 1.2%), and are less likely to fill high ranked occupations (-9.3% versus -2%).

We find small differences across age groups (23 to 39 and 40 to 59) – see Table 6 – and evidence (see Table 7) that college graduates with vocational high school who were born in the South have a larger negative earnings gap than those with an academic high school who were born in the North and Central areas of the country (-13.5% versus -4.3%).

5.4 Sensitivities

The negative employment and wage payoff to college graduates with vocational high school degrees could be driven by the fact that we are pooling together higher level technical and commercial schools – where many graduates go to college - and lower level professional schools, where only few graduates complete tertiary education. To verify whether this is the case, we re-arrange treatments by assigning the former to treatment V and the latter to the treatment NC. Results in Table A4 are qualitatively similar to those in Tables 2 and 3.

In our baseline specification, we have assigned college dropouts to the treatment NC. In an alternative specification (see Table A5), we assign to treatments V and G individuals with vocational and academic high school and some college (rather than with a completed college degree). When we do so, we find that the negative employment, earnings and training gap are somewhat larger than in Tables 2 and 3.

Our findings are also qualitatively robust to changes in the number of potential treatments, as shown by Table A6, where we consider five treatments by classifying those without college into three groups: no college and no high school, no college and vocational high school and no college and general high school.

5.5 Decomposing ATET by college field and rank

The effect of high school education on the economic returns to college could be due to the fact that vocational high school graduates are more likely to select college majors that offer lower payoffs, or to complete similar college majors as graduates from academic high schools but in less prestigious universities. Let the expected returns to college for a vocational high school graduate be

$$E[y|V] = E[y_a|V] * P_{Va} + E[y_b|V] * (1 - P_{Va}) \quad (2)$$

where y is the outcome, V is for vocational high school, the indices *a* and *b* are for the college major or the university rank and P_{Va} is the probability that a vocational high school graduate completes college major of type *a* and *b*. In a similar fashion,

$$E[y|G] = E[y_a|G] * P_{Ga} + E[y_b|G] * (1 - P_{Ga}) \quad (3)$$

are the returns associated with both college and academic high school. The difference in expected returns $\Delta = E[y|V] - E[y|G]$ is given by

$$\Delta = \Delta_a P_{Va} + \Delta_b (1 - P_{Va}) + residual \quad (4)$$

where the residual is defined as

$$residual = (E[y_a|G] - E[y_b|G]) * (P_{Va} - P_{Ga}).$$

Consider for instance the case when the indices *an* and *b* are for scientific college majors (STEM) and for other majors respectively. Then Eq. (4) indicates that the average earnings gap for college graduates with different high school education can be decomposed into the gap for graduates with a STEM major (Δ_a), weighted with the probability that vocational high school graduates complete a STEM major in college, the gap for graduates with other majors (Δ_b), also weighted with the probability of graduating, and the difference in the probability that graduates with different high school curricula complete a STEM major, weighted by the earnings gap between STEM and other majors for those with academic high school.

We perform the decomposition in Eq. (4) by estimating a five – treatments model (no college, vocational high school and STEM major, vocational high school and other majors, academic high school and STEM major, academic high school and other majors). Table 8 illustrates the results when we consider only the employed. We find that the negative wage gap and positive gap in weekly hours reported in Table 3 is larger for vocational high school graduates who have completed a non-STEM major. There is also evidence that the negative gap in the probability of filling a high ranked occupation is driven mainly by non-STEM graduates.

Table 9 presents instead the decomposition by university rank. The ranking is calculated using the CENSIS classification of Italian Universities: universities are first classified in five homogeneous groups according to size; for each group, they are awarded a rank depending on indicators of service quality, availability of scholarships and contributions, presence of a website and internationalization. We assign a university to high rank if its CENSIS rank is above the group-specific median and to low rank if it is at or below the median. We find that the negative gap in hourly wages and the positive gap in weekly hours experienced by college graduates with a vocational high school degree are larger for those who completed a low ranked college.

Finally, we decompose the estimated effects in Table 3 according to the final college score. This score ranges between 66 and 111 and its median value is 105.[11] Although graduates with vocational education have a mean score that is close to that attained on average by graduates with academic education (101 versus 104), the proportion completing their degree with a score higher than median is significantly lower (43 versus 57%). Table 10 shows that the negative earnings gap and positive gap in weekly hours is somewhat larger for vocational high school graduates who completed college with higher final scores.

Conclusions

A key feature of skill formation is dynamic complementarity (Cunha and Heckman, 2007): skills produced at one stage raise the productivity of investment at subsequent stages. An example of this complementary is the one involving skills accumulated in high school and college. If the skills accumulated in an academic high school have higher complementarity with the skills learned at college than the skills learned in a vocational high school, and if wages are proportional to productivity, the economic returns to college should be higher for college graduates completing academic high schools.

In this paper, we have investigated the effect of high school type (academic versus vocational) on college returns by using data from the Italian PLUS survey, which contains information both on the highest attained degree and on intermediate degrees.

We have found that college graduates with vocational high school are less likely to be employed than graduates with academic high school.

[11] The score 111 corresponds to 110 (the maximum score) cum laude.

Since the employed with a vocational high school degree earn 7.3 less per hour but work 3.8 percent more hours per week, the estimated gap in weekly earnings with respect to those with an academic high school degree is 3.5 percent.

We have also found that college graduates with a vocational education are less likely to fill high ranked occupations and more likely to find their first job quickly after school completion than other graduates. The wage penalty associated to vocational education in high school is larger for females than for males and for those born in the less economically developed Southern regions.

Although it is not clear from this study whether these results can be extended outside of Italy, we can speculate on the implications for educational choice. In some countries, vocational tracks in high schools do not allow access to college but prepare for early labour market entry. In other countries, students who complete vocational high school tracks can freely enrol in college.

If the dynamic complementarity between high school curriculum and college education is important, as this study documents, individuals who intend to complete tertiary education should enrol in academic high schools. By doing so, they can increase their expected earnings and the probability of receiving further training after completing their education. Vocational schools can perhaps address the gap in returns by increasing the academic content of their curriculum, thereby raising the complementarity of their skills with those developed at college. This is a process – called academic drift - already ongoing in some countries.

Table 1. Summary statistics

Variables	College and Vocational High School	College and Academic High School	No college
Hourly wages	16.03 (7.73)	17.78 (8.58)	12.27 (5.23)
Weekly hours worked	37.03 (10.80)	34.91 (11.51)	36.91 (10.41)
Employment probability	0.85	0.85	0.69
Training in the past three years	0.56	0.59	0.24
Time to first job below median	0.66	0.58	0.25
STEM major	0.43	0.42	-
High ranked university	0.60	0.62	-
Age	40.24(9.19)	40.75(9.88)	42.86(9.79)
Female	0.44	0.61	0.47
Top score in junior high	0.34	0.52	0.24
Highly educated mother	0.04	0.16	001
Highly educated father	0.07	0.22	0.02
Father high skilled job	0.15	0.29	0.07
Attitude at 13: music	0.06	0.08	0.04
Attitude at 13: sport	0.24	0.20	0.23
Attitude at 13: math and science	0.11	0.12	0.06
Attitude at 13: literature and arts	0.13	0.17	0.10
Attitude at 13: other	0.04	0.02	0.18

Notes: standard errors within parentheses

Table 2. Average treatment effects on the treated. College with vocational relative to college with academic high school education. OLS estimates.

	ATT	Oster's Beta
Employment Probability	-0.020***	-0.031
	(0.006)	
Log Hourly Wages	-0.122***	-0.151
	(0.022)	
Log Weekly Hours	-0.045*	-0.094
	(0.027)	
Log Hourly Wages (including self-employed)	-0.111***	-0.145
	(0.020)	
Log Weekly Hours (including self-employed)	-0.042*	-0.091
	(0.022)	
Probability of Training	-0.020*	-0.008
	(0.011)	
Time to First Job	0.055***	0.036
	(0.010)	

Notes: standard errors within parenthesis; ***, **, * for statistical significance at the 10, 5 and 1 percent level of confidence. Ordinary least squares include the following controls: region of birth dummies, year dummies, the interactions of junior-high school final scores with area dummies, dummies for attitudes at age 13, the interactions of mother and father education with area dummies, fathers' occupation, a dummy for the presence of parents at the end of junior high school, a dummy for parents born in the country, a gender dummy, a quartic in the year of birth, the local unemployment rate at age 14, the regional share of students enrolled in vocational education at age 14 and the number of college course available in the region of birth at age 19.

Table 3. Average treatment effects on the treated. Employed individuals. College with vocational relative to college with academic high school education.

	ATT	Oster's Beta
Log Hourly Wages	-0.073*** (0.012)	-0.065
Log Weekly Hours	0.038*** (0.009)	0.026
Log Hourly Wages (including self-employed)	-0.071*** (0.015)	-0.075
Log Weekly Hours (including self-employed)	0.029*** (0.016)	0.016
High Ranked Occupation	-0.051*** (0.011)	0.071
Type of Contract	-0.014 (0.013)	-0.003
Training	-0.004 (0.013)	0.013

Notes: see Table 2

Table 4. Average treatment effects on the treated. College with vocational relative to college with academic high school education.

	IPWRA	EB
All Individuals		
Employment Probability	-0.006	-0.017**
Log Hourly Wages	-0.059***	-0.101***
Log Weekly Hours	0.021	-0.030
Training	-0.032**	-0.042***
Time to First Job	0.057***	0.029***
Only for employed		
Log Hourly Wages	-0.054***	-0.075***
Log Weekly Hours	0.045***	0.045***
High Ranked Occupation	-0.086***	-0.091***
Type of Contract	0.006	0.017*
Training	-0.032**	-0.039***

Notes: ***, **, * for statistical significance at the 10, 5 and 1 percent level of confidence. IPWRA: Inverse Probability Weighted Regression Adjusted. EB: Entropy Balancing.

Table 5. Heterogeneous average treatment effects on the treated based on gender. College with vocational relative to college with academic high school education. OLS estimates. Employed individuals only.

	Male	Female
Log Hourly Wages	-0.053***	-0.079***
Log Weekly Hours	0.012	0.076***
High Ranked Occupation	-0.020	-0.093***
Type of Contract	-0.015	0.005
Training	0.015	-0.041**

Notes: see Table 2

Table 6. Heterogeneous average treatment effects on the treated based on age. College with vocational relative to college with academic high school education. OLS estimates. Employed individuals only.

	Young	Old
Log Hourly Wages	-0.066***	-0.071***
Log Weekly Hours	0.017	0.052***
High Ranked Occupation	-0.033**	-0.065***
Type of Contract	0.020	-0.022
Training	-0.020	-0.010

Notes: see Table 2

Table 7. Heterogeneous average treatment effects on the treated based on region of birth. College with vocational relative to college with academic high school education. Employed individuals only.

	North & Centre	South
Log Hourly Wages	-0.043***	-0.135***
Log Weekly Hours	0.031***	0.055***
High Ranked Occupation	-0.043***	-0.059***
Type of Contract	0.006	-0.022
Training	-0.006	-0.022

Notes: see Table 2

Table 8. Decomposing average treatment effects on the treated. By college major: STEM major versus other majors. Employed individuals only.

	ATT	STEM	NON-STEM	Residual
Log Hourly Wages	-0.073	-0.057	-0.079	-0.004
Log Weekly Hours	0.038	-0.002	0.07758	-0.005
High Ranked Occupation	-0.051	-0.0090	-0.117	0.020
Type of Contract	-0.004	0.003	0.018	-0.015
Training	-0.014	-0.008	-0.034	0.009

Table 9. Decomposing average treatment effects on the treated. By university rank: higher than median ranked versus median or lower than median ranked university. Employed individuals only.

	ATT	Higher rank	Lower rank	Residual
Log Hourly Wages	-0.073	-0.036	-0.109	0.007
Log Weekly Hours	0.038	-0.005	0.098	-0.023
High Ranked Occupation	-0.051	-0.044	-0.077	0.013
Type of Contract	-0.004	0.011	0.019	-0.020
Training	-0.014	-0.031	-0.046	0.026

Table 10. Decomposing average treatment effects on the treated. By final college score: higher than median college score versus median or lower than median score. Employed individuals only.

	ATT	Higher rank	Lower rank	Residual
Log Hourly Wages	-0.073	-0.060	-0.049	-0.019
Log Weekly Hours	0.038	0.040	0.025	0.006
High Ranked Occupation	-0.051	-0.072	-0.030	-0.003
Type of Contract	-0.004	-0.002	0.001	-0.008
Training	-0.014	0.001	-0.045	0.012

Appendix

Table A1. Average treatment effects on the treated by blocks of controls. College with vocational relative to college with academic high school education. OLS estimates with different sets of controls. All individuals

	Employment Probability			
ATT	-0.030***	-0.039***	-0.030***	-0.020***
Observations	70,009	70,009	70,009	70,009
R-squared	0.095	0.150	0.157	0.168
Age and gender	Yes	Yes	Yes	Yes
Local demand, supply and peer effects	No	Yes	Yes	Yes
Final scores and attitudes	No	No	Yes	Yes
Family background	No	No	No	Yes

Notes: ***, **, * for statistical significance at the 10, 5 and 1 percent level of confidence

Table A2. Average treatment effects on the treated by blocks of controls. College with vocational relative to college with academic high school education. OLS estimates with different sets of controls. All individuals

	Log Hourly Wages			
ATT	-0.143***	-0.165***	-0.139***	-0.122***
Observations	60,469	60,469	60,469	60,469
R-squared	0.124	0.185	0.193	0.206
Age and gender	Yes	Yes	Yes	Yes
Local demand, supply and peer effects	No	Yes	Yes	Yes
Final scores and attitudes	No	No	Yes	Yes
Family background	No	No	No	Yes

Notes: ***, **, * for statistical significance at the 10, 5 and 1 percent level of confidence

Table A3. Average treatment effects on the treated by blocks of controls. College with vocational relative to college with academic high school education. OLS estimates with different sets of controls. Employed individuals (excluding the self-employed)

	Log Hourly Wages			
ATT	-0.087***	-0.089***	-0.082***	-0.073***
Observations	32,858	32,858	32,858	32,858
R-squared	0.150	0.154	0.158	0.165
Age and gender	Yes	Yes	Yes	Yes
Local demand, supply and peer effects	No	Yes	Yes	Yes
Scores and attitudes	No	No	Yes	Yes
Family background	No	No	No	Yes

Notes: ***, **, * for statistical significance at the 10, 5 and 1 percent level of confidence

Table A4. Average treatment effects on the treated. College with vocational education (excluded professional education) relative to college with academic high school education. Employed individuals only

	ATT	Standard error
Log Hourly Wages	-0.072***	(0.013)
Log Weekly Hours	0.037***	(0.010)
High Ranked Occupation	-0.053***	(0.012)
Type of Contract	-0.001	(0.014)
Training	-0.012	(0.013)

Notes: standard errors within parentheses; ***, **, * for statistical significance at the 10, 5 and 1 percent level of confidence. See Table 2.

Table A5. Average treatment effects on the treated. College with vocational relative to college with academic high school education. College dropouts classified in the college groups. Employed individuals only.

	ATT	Standard errors
Log Hourly Wages	-0.104***	(0.009)
Log Weekly Hours	0.038***	(0.006)
High Ranked Occupation	-0.135***	(0.010)
Type of Contract	0.027***	(0.010)
Training	-0.067***	(0.010)

Notes: see Table A4

Table A6. Average treatment effects on the treated. Five treatments. College with vocational relative to college with academic high school education. OLS estimates.

	ATT	Standard errors
Log Hourly Wages	-0.090***	(0.013)
Log Weekly Hours	0.040***	(0.010)
High Ranked Occupation	-0.087***	(0.012)
Type of Contract	-0.012	(0.014)
Training	-0.038***	(0.014)

Notes: see Table A4. We classify individuals without college (NC) into three groups: no college and no high school (NCNH), no college and vocational high school (NCV) and no college and general high school (NCG). To verify common support, we retain only individuals with a probability of being assigned to treatment of at least 1 percent, for each treatment.

References

Allen, J., Massing, N., Schneider, S. and Van der Velden, R., 2017, Proposal for the Revision of the Background Questionnaire on Education and Training for the 2nd Cycle of PIAAC

Altonji J. 1995. The effects of high school curriculum on education and labor market outcomes. *Journal of Human Resources*. 30(3):409-38

Altonji, J.G., Blom, E. and Meghir, C., 2012. Heterogeneity in human capital investments: High school curriculum, college major, and careers. *Annu. Rev. Econ.*, *4*(1), pp.185-223.

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

Autor, D., Katz, L., Kearney, M., 2008, Trends in US wage inequality: revising the revisionists, *The Review of Economics and Statistics*, 300-323

Bianchi, N., 2018. The indirect effects of educational expansions: Evidence from a large enrollment increase in STEM majors. Available at SSRN 3037247.

Biewen, M. and Tapalaga, M., 2017, Life cycle educational choice in a system with early tracking and "second chance" option, *Economics of Education Review*.

Brunello G. & Rocco L., 2017. The Labor Market Effects of Academic and Vocational Education over the Life Cycle: Evidence Based on a British Cohort, *Journal of Human Capital*, University of Chicago Press, vol. 11(1), pages 106-166.

Brunello, G. & Checchi, D. 2007. Does school tracking affect equality of opportunity? New international evidence, *Economic Policy*, vol. 22, pages 781-861

Cattaneo, M. D. 2010. Efficient semiparametric estimation of multi-valued treatment effects under ignorability. *Journal of Econometrics* 155: 138–154.

Checchi, D. and Flabbi, L., 2007. Intergenerational mobility and schooling decisions in Germany and Italy: The impact of secondary school tracks, IZA Discussion Paper.

Chevalier, A., 2011. Subject choice and earnings of UK graduates. *Economics of Education Review*, *30*(6), pp.1187-1201.

Cunha, F. and Heckman, J., 2007. The technology of skill formation. *American Economic Review*, *97*(2), pp.31-47.

Deming, D. & Hastings, J. & Kane, T. & Staiger, T., 2014. School Choice, School Quality, and Postsecondary Attainment, *American Economic Review*, American Economic Association, vol. 104(3), pages 991-1013

Dustmann, C. & Patrick A. Puhani & Uta Schönberg, 2017. The Long-term Effects of Early Track Choice, *Economic Journal*, Royal Economic Society, vol. 127(603), pages 1348-1380, August.

Green, A, Wolf, A and Leney, T, 1999, *Convergence and Divergence in European Education and Training Systems*, London.

Hainmueller, J. and Xu, Y., 2013. Ebalance: A Stata package for entropy balancing. *Journal of Statistical Software*, 54(7).

Joensen, J.S. and Nielsen, H.S., 2009. Is there a causal effect of high school math on labor market outcomes? *Journal of Human Resources*, 44(1), pp.171-198.

Kirkeboen, L.J., Leuven, E. and Mogstad, M., 2016. Field of study, earnings, and self-selection. *The Quarterly Journal of Economics*, 131(3), pp.1057-1111.

Lemieux, T., 2006, Increasing residual wage inequality: composition effects, noisy data or rising demand for skills?, *American Economic Review*, 461-98

Levine, P.B. and Zimmerman, D.J., 1995. The benefit of additional high-school math and science classes for young men and women. *Journal of Business & Economic Statistics*, 13(2), pp.137-149.

Lindley, J and McIntosh, S., 2015, Growth in within graduate wage inequality: the role of subjects, cognitive skills dispersion and occupational concentration, *Labour Economics*, 101-111

Rizzica, L., 2013. Home or Away? Gender differences in the effects of an expansion of tertiary education supply. *Bank of Italy Occasional Paper*, (181).

Rose, H. and Betts, J.R., 2004. The effect of high school courses on earnings. *The Review of Economics and Statistics*, 86(2), pp.497-513.