The Pathways to College

Lisha Agarwal (Padova)

Giorgio Brunello (Padova and IZA) ^

Lorenzo Rocco (Padova and IZA)

Acknowledgements: we are grateful to Marco Bertoni and to the audience in Padova for comments and suggestions. The usual disclaimer applies.

^ Corresponding author: Giorgio.brunello@unipd.it

1. Introduction

College earnings vary considerably across individuals. In 2014, the coefficient of variation of gross earnings for college graduates aged 30 to 40 and working full time in Germany, France, the UK and Italy was equal to 70.6, 48.9 and 65.4 and 54.3 percent of the mean (equal to 46,409, 35,347, 47,006 and 29947 euro respectively). There are many factors driving this variation, including individual ability and the college major (see Altonji, Blom and Meghir, 2012).

One factor is the educational pathway or the bundle of qualifications attained before entering and completing to college. High school quality affects both enrolment and college completion (Deming et al, 2014). In several countries, both in Europe and in Asia, high school curricula are organized in academic and vocational tracks, with the differentiation occurring at different ages (Brunello and Checchi, 2007). While some tracks do not lead to college, in several countries access to college is open to graduates with an academic or vocational curriculum.[1]

In Italy, access to college has been progressively liberalized during 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 56.8 and 34.2 percent. For law graduates, these percentages were 77.2 and 22.8 respectively. In the same year, engineers with an academic high school degree earned on average 37,353 euro (gross), significantly more than engineers with a vocational high school degree (29,870 euro). In a similar fashion, law graduates with an academic high school degree earned significantly more than similar graduates with a vocational high school degree (29,507 versus 26,209 euro).[2]

[1] 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.

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

Understanding the importance of high school qualifications for the returns to college is often precluded by 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 only partial data on pathways, for respondents who are currently in education at the time of the survey (see Allen, Massing, Schneider and van der Velden, 2017).

In this paper, we investigate whether economic returns to college, measured by probability of employment, hourly and weekly earnings, time to the first job and probability of receiving training, vary significantly with high school curriculum, using data from the Italian PLUS (Participation Labour and Unemployment Survey) for the period 2010 to 2016. The key feature of this data is that the include information both on the highest completed education and the type of completed high school. They also contain a fairly rich set of covariates, such as measures of ability before high school, region of birth and family background and occupation.

We believe that our research contributes both to the relatively small literature investigating the effects of the school curriculum on college returns, which has been recently reviewed by Altonji, Blom and Meghir, 2012,[3]and to the literature – mainly European – that explores the economic consequences of secondary school tracks (recent contributions including Dustmann et al, 2017, Biewen and Tapalaga, 2016, and Brunello and Rocco, 2017)

We estimate the causal impact of choosing a vocational or an academic high school curriculum (or track) on the returns to college using both the inverse-probability weighted regression adjusted estimator (IPWRA), which imputes to each individual the missing potential outcomes by exploiting information on

[3] Relevant contributions in this literature are Altonji, 1995; Levine and Zimmermann, 1995; Rose and Betts, 2004 and Joensen and Nielsen, 2008.

individuals with similar characteristics who received alternative treatments (Cattaneo, 2010)), and entropy balancing, a re-weighting scheme that specifies for each selected covariate a set of balance constraints to equalize the moments of the covariate distribution between the treatment and the reweighted control group (see Hainmueller and Xu, 2013).

We find that college graduates with vocational high school are about as likely to be employed as graduates with academic high school but earn between 6.5 and 10.1 percent less per hour. However, since they work 5.7 to 6.1 percent more hours per week, the gap in weekly earnings is much lower, between 0.8 and 4 percent, depending on the estimation method. We also find that college graduates with a vocational education are less likely to receive training after completing their education and more likely to find their first job quickly after school completion. The wage and training penalty associated to vocational education in high school is significantly larger for females than for males.

Although graduates from vocational high schools are much less likely to complete college than other graduates, their share a similar probability of completing a STEM field or a high ranked university. We decompose the estimated returns gap between college graduates with vocational and academic high school into the gap associated both to college major (STEM versus non-STEM) and to college rank. We find that the gap in hourly earnings and training is larger for those completing low ranked or non-STEM universities than for other graduates.

We argue that these findings are driven neither by selection into college nor by the fact that graduates from academic and vocational high school types choose different college majors or college quality. They suggest instead that the academic curriculum in high school is a stronger complement to college education than the vocational curriculum.

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

2. 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) not continuously between 2005 and 2016.[4] The 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 paths from the end of junior high school (at age 14) to college and its completion. 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 detailed information on parental education and 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 72,703 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.

3. 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

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

classical, scientific or linguistic and pedagogical studies (*licei* and *scuole magistrali*).

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 who graduate from a vocational school are much less likely to enrol in college than those who complete an academic track (13.4 versus 53.4 percent). While vocational high school graduates have lower final scores than academic high school graduates (76.9/100 versus 79.6/100, where 100 if the maximum score), those enrolling in college have similar scores (83.2/100 versus 82.9/100).

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.

College graduates with a vocational education have lower average hourly wages and 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 inclination at age 13 was mainly for sports, compared to sports and literature and the arts for those with an academic education.

From the viewpoint of educational choices from age 14 – at the end of lower secondary education – onwards, we consider three possible states, or potential treatments, at age t, where t is between 23 and 59: 1) no completion of high school or college, vocational or academic high school graduate with no college degree (82.5 percent); 2) vocational high school and college graduate (4.9 percent); 3) academic high school and college graduate 12.6 percent). This classification assigns individuals with some college but no college degree and with some high school but no degree to group 1.

An alternative classification – that we also consider as a robustness exercise – breaks down group one into three groups, one for those with no high school degree, another for those with a vocational high school degree and one for those with an academic high school degree.

We study the effects of high school curriculum on the labour market outcomes of college graduates. These outcomes include: log hourly wages; log weekly hours worked; employment probability; the probability of having received training after completing school and in 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.

Due to the poor reliability of wage data for the self-employed, we exclude this group from the sample. We compute log hourly wages and hog weekly hours using the approximation $lnx = \ln(1 + x)$, which allows us to compute also a broader measure of earnings which assigns zero wages and hours to individuals not employed.

4 The Empirical Approach

It is well known that the comparison of average outcomes between college graduates with different curricula is not informative of the average treatment effect – or the difference between average outcomes in the event of treatment and the counterfactual outcomes in the event of no treatment – because of selection effects. Students with a vocational curriculum in high school are less likely to enrol in college than students with an academic curriculum, and may have different abilities.

In this paper, we address selection into treatment using a selection on observables strategy. This strategy is based on the “conditional independence assumption” (CIA), which states that, conditional on a set of pre-determined variables, potential treatments are as good as randomly assigned to individuals. This assumption requires that the factors that determine the choice of education should be sufficiently captured by observed pre-determined variables, so that any residual variation in education, conditional on these characteristics, is either random or due to factors that do not influence the outcomes of interest.

We believe that our data are sufficiently rich to support this assumption. A part 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 at 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 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.

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 include the presence of parents 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 identification of causal effects requires that the treatment can be considered as good as randomly assigned. When this happens, covariates are balanced across treatment and control groups. We use two alternative approaches to covariate balancing, the inverse probability weighted regression adjusted method - briefly IPWRA – and entropy balancing.

IPWRA imputes to each individual the missing potential outcomes by exploiting information on individuals with similar characteristics who received alternative treatments. Under the conditional independence assumption, potential outcomes for all possible treatments t are given by

$$E[Y_i(t)|X] = E[Y_i(t)|t_i(t) = 1, X] \tag{1}$$

where X is a vector of controls, the condition $t_i(t) = 1$ defines the set of individuals assigned to treatment *t*, and the average treatment effect (ATE) of treatment *t* with respect to *t'* is computed by comparing the average outcome of the individuals treated with *t* and the average outcome of the individuals treated with t'

$$ATE(t, t'|X) = E[Y_i(t) - Y_i(t')|X] = E[Y_i(t)|t_i(t) = 1, X] - E[Y_i(t')|t_i(t') = 1, X] \tag{2}$$

For the ATE to be correctly defined, the cell characterised by X = x, for all possible x, must include both subjects assigned to treatment *t* and subjects assigned to treatment t'. This condition is called common support or overlapping. In our empirical estimates, we always verify that overlap holds, and retain in our final sample only those individuals with a predicted probability of at least 1 percent for each possible treatment.

IPWRA 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 both groups with their inverse probability of treatment in the outcome model, IPWRA ensures that the covariates are balanced across the treatment and control groups. Finally, ATEs are obtained by comparing the sample means of the predicted potential outcomes. In our study, we bootstrap the three-step procedure to obtain consistent point estimates.[5]

Entropy balancing, proposed by Hainmueller and Xu, 2013, is based on a maximum entropy re-weighting scheme that fits weights that satisfy a potentially large set of balance constraints. Instead of checking for covariate balance after pre-processing the data, this method is based on the specification of a desired level of covariate balance – on the first, second or higher moments of the covariate distributions in the treatment and the re-weighted control group - using a set of balance conditions. Once re-balancing has been attained, average treatment effects are estimated by standard regressions using re-weighted data.

5. Results

5.1 Main findings

In our baseline specification, we consider two treatment groups (college plus vocational high school and college plus academic high school) and a control group (junior high school graduates who have not completed college). The estimates reported in Table 2 are based on IPWRA. We find that college graduates with vocational and academic education have similar employment probabilities. Graduates with a vocational high school education, however, have significantly lower hourly earnings (-6.5 or -7.3 percent, depending on whether we include the non employed in the sample). They work longer hours (+5.7 percent) and are less likely to be training recipients (-2.8 percent) during the last

[5] An attractive property of IPWRA is that it is doubly robust to misspecification. Indeed, this is an 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.

three years. On the other hand, they are much more likely (+7.1 percent) to find their first job below the median time required. Assuming that utility is increasing in weekly earnings but decreasing in weekly hours, and that preferences are similar across groups, these results suggest that college graduates with vocational high school have similar weekly earnings but lower utility than graduates with an academic high school degree.

We verify whether the re-weighting scheme implied by IPWRA guarantees covariate balancing by regressing each covariate on treatment dummies using as weights the normalised inverse probability of treatment. Results reported in Table 3 show that for most covariates we cannot reject balancing, because the coefficients associated to the two dummies are not statistically different from zero. There are, however, a few exceptions that include gender, region of birth and junior high school grades.

We therefore replicate our estimates using entropy balancing and imposing pre-estimation balancing on the first and second moment of the covariate distributions in the treatment and the re-weighted control group. Since our treatment is multi-valued, we impose these constraints by restricting the sample first to the group with college and academic high school education and the control group (no college) – therefore excluding the group with college and vocational education – and second the group with college and vocational high school education and the control group (no college) –excluding the group with college and academic high school education.

Our results using entropy balancing - reported in Table 4 - qualitatively similar to those reported in Table 2 using IPWRA. The estimated hourly wage gap, however, is larger (-10.1 percent versus -6.5 percent). Since the positive gap in weekly hours is similar, the negative gap for weekly wages is also larger, both when we only consider employees (-4 rather than -0.8 percent) and when we assign zero wages to the non-employed (-7.7 versus -3.5 percent).

Using both methods, we compare the effects for males and females in Table 5. In the case of females, we find that having completed vocational high school before completing college carries a penalty with respect to having completed an academic high school for employment, wages and training. For males, there is a penalty only for wages, and it is considerably smaller than the one for females. There is also evidence that the negative gap in hourly wages is much higher for individuals who have grown up in Southern regions (Table A1).

5.2 Sensitivities

These findings are qualitatively robust to changes in the number of potential treatments. Tables A2 to A3 in the Appendix present the estimates in the case of five treatments, where we classify 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; and two treatments, that only consider college graduates. They are also broadly unchanged when we allocate college dropouts, who have done some college, to the group with a college degree (see Table A4).

5.3 Decomposing ATE 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 that they complete similar college majors as graduates from academic high schools but in less prestigious universities. Let the expected returns from college for a vocational high school graduate be

$$E[y|V] = E[y_a|V] * P_{Va} + E[y_b|V] * P_{Vb} \quad (3)$$

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} and P_{Vb} are the probabilities that a vocational high school graduate completes college major or type a and b. In a similar fashion,

$$E[y|G] = E[y_a|G] * P_{Ga} + E[y_b|G] * P_{Gb} \quad (4)$$

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

$$\Delta = \Delta_a P_{Va} + \Delta_b P_{Vb} + residual \quad (5)$$

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 *a* and *b* are for scientific college majors (STEM) and for other majors respectively. Then Eq. (5) 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 enrolment, and the difference in the probability that graduates from different high school types enrol in a STEM major, weighted by the earnings gap between STEM and other majors for those with academic high school.

Table 6 illustrates the decomposition when we distinguish between STEM and non-STEM majors, based on the IPWRA estimates of the five – treatments model reported in Table A5 in the Appendix (no college, vocational high school and STEM college, vocational high school and non - STEM college, academic high school and STEM college, academic high school and non-STEM college). We find that the biggest share of the gap reported in Table 2 for hourly wages, weekly hours and probability training is the gap in non-STEM majors, weighted with the probability of enrolment by vocational high school graduates.

Table 7 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, universities are awarded a rank depending on indicators of service quality, availability scholarships and contributions, presence of a website and internationalization. We assign a university to high rank if its CENSIS rank is above the median and to low rank if it is below the median. As in the previous table, we compute the decomposition by estimating a five – treatments model (no college, vocational high school and high ranked university, vocational high school and low ranked university, academic high school and high ranked university, academic high school and low ranked university). We find that the gap in hourly wages and probability of training is due mainly to the gap associated to low ranked universities (weighted for the probability of enrolment by vocational high school graduates).

We conclude from this that the negative wage and training penalty associated to having completed a vocational high school is larger when the college major is non-STEM and the university is low ranked, suggesting that the complementarity between high school and college skills in favour of academic high school skills is weaker in these types of colleges.

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 curriculum have higher complementarity with skills learned during college than the skills learned in a vocational high school curriculum, the economic returns to college are likely to differ according to the type of high school education.

In this paper, we have used Italian data, which include information both on the highest attained degree and on intermedia degrees, as well as a rich set of covariates which allow us to control for cognitive ability before the start of high school education, to investigate the effect of high school type (academic versus vocational) on college returns.

We have addresses endogenous selection into high school type using inverse probability weighted regression analysis and entropy balancing. Both methods re-weight the data so as to attain covariate balancing between treatment and control groups.

We have found that having completed a vocational high school before college generates a negative premium in hourly earnings and in the probability of training (after college). The estimated hourly wage penalty is sizeable and ranges between 6.5 and 10.1 percent. In order to attenuate this penalty, college graduates with a vocational high school degree need to work longer weekly hours. This gap is particularly pronounced among females. The negative penalty is common to different college fields (STEM versus non-STEM) and is larger in absolute value for individuals attending low ranked universities.

These results have 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 so doing, they can increase their expected earnings and the probability of receiving further training after completing their education.

Table 1. Summary statistics

Variables	College and Vocational High School	College and Academic High School	No college
Hourly wages	15.91 (7.56)	17.67 (8.53)	13.52(5.56)
Weekly hours worked	37.11 (10.69)	35.00 (11.50)	36.74(10.16)
Employment probability	0.85	0.84	0.76
Training in the past three years	0.57	0.60	0.34
Time to first job below median	0.66	0.58	0.33
Prob. college	0.13	0.53	
STEM major	0.41	0.40	-
High ranked university	0.60	0.63	-
Age	40.31 (9.18)	40.81 (9.89)	41.56(10.02)
Female	0.43	0.61	0.49
Top score in junior high	0.33	0.52	0.27
Highly educated mother	0.08	0.19	0.07
Highly educated father	0.11	0.25	0.09
Father high skilled job	0.16	0.28	0.12
Attitude at 13: music	0.08	0.10	
Attitude at 13: sport	0.41	0.34	
Attitude at 13: math and science	0.20	0.20	
Attitude at 13: literature and arts	0.23	0.29	
Attitude at 13: other	0.07	0.07	

Notes: standard errors within parentheses

Table 2. Average treatment effects. College with vocational relative to college with academic high school education. Three potential treatments. IPWRA.

	ATE	Standard Error
Employment Probability	-0.003	0.006
Log Hourly Wages	-0.065***	0.014
Log Weekly Hours	0.059***	0.010
Training	-0.028**	0.012
Time to First Job	0.071**	0.013
Log Hourly Wages (Zero Wages for Non-employed)	-0.073**	0.029
Log Weekly Hours (Zero Hours for Non-employed)	0.065***	0.011

Notes: ***, **, * for statistical significance at the 10, 5 and 1 percent level of confidence. Selection into treatment is estimated using a multinomial specification with 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. Balancing tests using the re-weighted sample from IPWRA

Covariate	College & Academic High School Dummy	College & Vocational High School Dummy
Year of birth	-0.078	0.151
Year of birth squared	1.592	6.783
Year of birth to the cube /1000	0.184	0.260
Year of birth to the fourth/1000000	9,901	9.728
Region of birth=2	-0.019***	0.002
Region of birth=3	0.001	-0.000
Region of birth=4	-0.001	-0.006
Region of birth=5	0.008**	0.000
Region of birth=6	-0.003	0.001
Region of birth=7	0.001	-0.001
Region of birth=8	-0.001	0.001
Region of birth=9	0.005	0.000
Region of birth=10	0.003	-0.002
Region of birth=11	0.007	0.007
Region of birth=12	0.002	-0.001
Region of birth=13	0.004	-0.004
Region of birth=14	-0.003	0.005
Region of birth=15	0.001	0.004
Region of birth=16	-0.000	-0.003
Region of birth=17	-0.004	-0.000
Region of birth=18	-0.000	-0.001
Year=2011	-0.000	0.003
Year=2014	-0.014	0.000
Year=2016	0.002	-0.000
Score=2 area=1	0.001	0.003
Score=3 area=1	0.003	0.007
Score=4 area=1	-0.003	-0.001
Score=1 area=2	0.011*	0.002
Score=2 area=2	-0.012	-0.010
Score=3 area=2	-0.009	-0.007
Score=4 area=2	0.006	0.004
Personal interest=2	-0.000	0.001
Personal interest=3	0.001	0.000
Personal interest=4	0.003	-0.005
Personal interest=5	0.000	-0.002
Personal interest=6	0.008**	-0.002
Personal interest=7	0.005	0.001
Personal interest=8	0.001	0.001
Mother education=1	-0.002	-0.014
Mother education=2	0.000	-0.002
Mother education=3	-0.004	0.004
Mother education=4	-0.003	0.002
Mother education=5	-0.000	0.005
Mother education=6	0.002	0.001
Mother education=7	-0.001	0.000
Mother education=8	-0.000	-0.001
Father education=1	-0.007	-0.013
Father education=2	0.000	-0.001
Father education=3	-0.004	0.008

Table 3 (continued)

Father education=4	-0.003	0.004
Father education=5	0.004	0.005
Father education=6	0.003	0.000
Father education=7	0.001	0.000
Father education=8	0.001	-0.003
Father occupation=2	-0.001	0.000
Father occupation=3	-0.003	-0.004
Father occupation=4	-0.000	-0.002
Father occupation=5	-0.004	0.004
Father occupation=6	0.009	0.000
Father occupation=7	-0.002	-0.004
Father occupation=8	0.007	-0.007
Father occupation=9	0.001	-0.002
Father occupation=10	-0.009	0.003
Father occupation=11	-0.003	0.004
Father occupation=12	0.005	0.005
Parents at home at age 14	-0.007	-0.003
Unemployment rate at 14	-0.000	0.000
Mother or father born abroad	0.003	0.002
Female	-0.006	-0.023**
Regional share enrolled in VHS	-0.003*	-0.001
Number of regional college courses	0.061	0.019

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

Table 4. Average treatment effects. College with vocational relative to college with academic high school education. Three potential treatments. Entropy balancing.

	ATE	Standard Error
Employment Probability	-0.007*	0.004
Log Hourly Wages	-0.101***	0.005
Log Weekly Hours	0.061***	0.004
Training	-0.034***	0.004
Time to First Job	0.056***	0.004
Log Hourly Wages (Zero Wages for Non-employed)	-0.094***	0.013
Log Weekly Hours (Zero Hours for Non-employed)	0.017**	0.014

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

Table 5. ATE by gender. IPWRA and entropy balancing.

	Males IPWRA	Females IPWRA	Males Entropy B.	Females Entropy B.
Employment Probability	0.004	-0.024*	0.007	-0.020***
Log Hourly Wages	-0.042*	-0.102***	-0.077***	-0.116***
Log Weekly Hours	0.051***	0.066***	0.048***	0.065***
Training	0.004	-0.055***	0.000	-0.062***
Time to First Job	0.074***	0.070***	0.051***	0.059***
Log Hourly Wages (Zero Wages for N.E.)	-0.001	-0.166***	0.005	-0.169***
Log Weekly Hours (Zero Hours for N.E.)	0.054***	0.079***	0.061***	-0.031

Notes: ***, **, * for statistical significance at the 10, 5 and 1 percent level of confidence.
N.E: non employed

Table 6. Decomposing average treatment effects. By college major: STEM versus other majors. IPWRA.

	ATE	Δ_a	Δ_b	$\Delta_a P_{Va}$	$\Delta_b P_{Vb}$	*Residual*
Employment Probability	-0.003	-0.022	-0.002	-0.009	-0.001	0.007
Log Hourly Wages	-0.065	-0.045	-0.071	-0.018	-0.041	-0.005
Log Weekly Hours	0.057	0.019	0.056	0.008	0.033	0.016
Training	-0.028	-0.007	-0.042	-0.003	-0.024	-0.000
Time to First Job	0.071	0.047	0.057	0.019	0.032	0.018
Log Hourly Wages (Zero Wages for N.E.)	-0.073	-0.077	-0.056	-0.032	-0.033	-0.008
Log Weekly Hours (Zero Hours for N.E)	0.038	-0.07	0.025	-0.029	0.015	0.052

Note: N.E.: not employed. The subscripts a and b in $\Delta_a P_{Va}$ and $\Delta_b P_{Vb}$ are for STEM and non-STEM majors.

Table 7. Decomposing average treatment effects. By college major: higher than median ranked versus median or lower than median ranked university.

	ATE	Δ_a	Δ_b	$\Delta_a P_{Va}$	$\Delta_b P_{Vb}$	Residual
Employment Probability	-0.003	-0.022	-0.002	-0.009	-0.001	0.007
Log Hourly Wages	-0.065	-0.045	-0.071	-0.018	-0.041	-0.005
Log Weekly Hours	0.057	0.019	0.056	0.008	0.033	0.016
Training	-0.028	-0.007	-0.042	-0.003	-0.024	-0.000
Time to First Job	0.071	0.047	0.057	0.019	0.032	0.018
Log Hourly Wages (Zero Wages for n.e.)	-0.073	-0.077	-0.056	-0.032	-0.033	-0.008
Log Weekly Hours (Zero Hours for n.e.)	0.038	-0.07	0.025	-0.029	0.015	0.052

Note: N.E.: not employed. The subscripts a and b in $\Delta_a P_{Va}$ and $\Delta_b P_{Vb}$ are for high and low rank colleges.

Table A1. ATE by area. IPWRA

	North	South
Employment Probability	-0.003	0.004
Log Hourly Wages	-0.035*	-0.110***
Log Weekly Hours	0.042***	0.083***
Training	-0.020	-0.023
Time to First Job	0.074***	0.074***
Log Hourly Wages (Zero Wages for Non-employed)	-0.047*	-0.081
Log Weekly Hours (Zero Hours for Non-employed)	0.044***	0.105***

Notes: see Table 2

Table A2. Average treatment effects. Five potential treatments. College with vocational relative to college with academic high school education.

	ATE	Standard error
Employment Probability	0.000	0.007
Log Hourly Wages	-0.055***	0.019
Log Weekly Hours	0.055***	0.012
Training	-0.024*	0.015
Time to First Job	0.073***	0.016
Log Hourly Wages (Zero Wages for Non-employed)	-0.059**	0.026
Log Weekly Hours (Zero Hours for Non-employed)	0.061***	0.010

Notes: see Table 2

Table A3. Average treatment effects. Two potential treatments. College with vocational relative to college with academic high school education.

	ATE	Standard error
Employment Probability	-0.016**	0.007
Log Hourly Wages	-0.073***	0.014
Log Weekly Hours	0.045***	0.009
Training	-0.024*	0.012
Time to First Job	0.050***	0.010
Log Hourly Wages (Zero Wages for Non-employed)	-0.105***	0.028
Log Weekly Hours (Zero Hours for Non-employed)	-0.019	0.027

Notes: see Table 2

Table A4. Average treatment effects. College dropouts classified in the college group. College with vocational relative to college with academic high school education.

	ATE	Standard error
Employment Probability	-0.004	0.006
Log Hourly Wages	-0.085***	0.011
Log Weekly Hours	0.066***	0.008
Training	-0.058***	0.012
Time to First Job	0.012	0.011
Log Hourly Wages (Zero Wages for Non-employed)	-0.082***	0.021
Log Weekly Hours (Zero Hours for Non-employed)	0.072***	0.008

Notes: see Table 2

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., Blom, E. and Meghir, C., 2012, Heterogeneity in Human Capital Investments: High School Curriculum, College Major and Careers, NBER Working Paper 17985

Bianchi, N., 2018, The Indirect Effects of Educational Expansions: Evidence from a Large Enrollment Increase in STEM Majors, mimeo

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, CEPR;CES;MSH, 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. & Flabbi, L., 2007. Intergenerational Mobility and Schooling Decisions in Germany and Italy: The Impact of Secondary School Tracks," IZA Discussion Papers 2876

Cunha, F. and Heckman, J., 2007, The Technology of Skill Formation, The American Economic Review.

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

Christian Dustmann & 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.

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.

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.

Rizzica, L., 2013. Home or Away? Gender differences in the effects of an expansion of tertiary education supply.

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